UNIVERSITY OF CALIFORNIA SAN DIEGO

Engineering Progressively Deimmunized and Redosable CRISPR-Cas Gene Therapies

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Biology

by

Nathan David Palmer

Committee in charge:

Professor Prashant G. Mali, Chair
Professor Justin R. Meyer, Co-Chair
Professor Vineet Bafna
Professor Matthew Daugherty
Professor Ananda Goldrath

2022

The Dissertation of Nathan David Palmer is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

DEDICATION

This work is dedicated to my loving and supportive family, David, Sonja, Jonathan, Jason, Scott, and Kaden who have always inspired me to pursue scientific discovery and without whom this work would have been impossible. And to my superlative colleagues in the Mali lab with whom it has been an absolute privilege to work and interact with throughout my time here.

TABLE OF CONTENTS

DISSERTATION APPROVAL PAGE iii

DEDICATION iv

TABLE OF CONTENTS v

LIST OF FIGURES vii

LIST OF TABLES viii

ACKNOWLEDGEMENTS ix

VITA xii

ABSTRACT OF THE DISSERTATION xiii

CHAPTER 1 – THE STATE OF CRISPR-CAS GENE THERAPY 1

1.1 THERAPEUTIC DEVELOPMENT CASCADE 2
1.2 DEVELOPING CRISPR-CAS THERAPEUTICS 5
1.2.1 Payload 5
1.2.2 Delivery 7
1.2.3 Administration 15
1.2.4 Pre-IND and pre-clinical development of CRISPR-Cas therapeutics 21
1.2.5 Challenges 35
1.3 LOOKING FORWARD 40
1.4 ACKNOWLEDGEMENTS 41

CHAPTER 2 – THE ADAPTIVE IMMUNE RESPONSE TO AAV-CRISPR-CAS 42

2.1 RESULTS 43
2.1.1 Immune response to AAV and Cas9 43
2.1.2 Immune barriers to effective gene editing 48
2.2 DISCUSSION 52
2.3 METHODS 55
2.3.1 Computational Methods 55
2.3.2 Experimental Methods 56
2.4 ACKNOWLEDGEMENTS 61

CHAPTER 3 – IMMUNE ORTHOGONAL ORTHOLOGS ENABLE REDOSING OF AAV-CRISPR-CAS 62

3.1 RESULTS 64
3.1.1 Identifying immune-orthogonal proteins 64
3.1.2 Confirming humoral immune orthogonality among Cas9 proteins 89
3.1.3 Confirming broad immune cross-reactivity among AAV serotypes 92

3.1.4 Overcoming immune barriers to effective gene editing 94
3.2 DISCUSSION 98
3.3 METHODS 101
3.3.1 Computational Methods 101
3.3.2 Experimental Methods 103
3.4 ACKNOWLEDGEMENTS 107

CHAPTER 4 – APPLYING LORAX TO PROGRESSIVELY DEIMMUNIZE CAS9 108

4.1 RESULTS 111
4.1.1 LORAX 111
4.1.2 Library design 114
4.1.3 Readout 114
4.1.4 Immunogenicity scoring 115
4.1.5 Library construction 119
4.1.6 Screening 119
4.1.7 Hit selection 121
4.1.8 Validations 125
4.2 DISCUSSION 131
4.3 METHODS 132
4.3.1 Computational methods 132
4.3.2 Experimental methods 134
4.4 ACKNOWLEDGEMENTS 142

CHAPTER 5 – CONCLUSIONS 142

5.1 SUMMARY OF THE WORK 143
5.2 LIMITATIONS 144
5.3 BROADER APPLICATIONS 145
5.4 ACKNOWLEDGMENTS 146

REFERENCES 147

LIST OF FIGURES

Figure 1.1: Steps towards enabling a gene therapy via CRISPR-Cas. 5

Figure 1.2: CRISPR-Cas therapeutic strategies towards disease-causing mutations. 7

Figure 1.3: Delivery modalities for CRISPR-Cas therapeutics. 8

Figure 2.1: Treatment efficacy and humoral response. 45

Figure 2.2: Experimental validation of a MHCII peptide predictions via IFN-γ ELISPOT.... 46

Figure 2.3: Cas9-specific splenocyte clearance in vivo.. 48

Figure 2.4: Redosing in immunized mice. 50

Figure 3.1: Assessing immune orthogonality among CRISPR effectors and AAV capsids.... 86

Figure 3.2: Cas9 immune orthogonal cliques. 87

Figure 3.3: In silico analyses of immunogenicity of Cas9 and AAV orthologues. 88

Figure 3.4: Humoral immune orthogonality among Cas9 and AAV proteins. 90

Figure 3.5: Confirming immune orthogonality of C. jejuni Cas9 to Sp- and SaCas9. 92

Figure 3.6: Major human AAV serotype groups. 93

Figure 3.7: Time course of multiple dosing with immune orthogonal orthologues..... 95

Figure 3.8: Exploring pre-existing CRISPR immunity. 99

Figure 4.1: LORAX protein engineering methodology to screen progressively deimmunized Cas9 variants. 112

Figure 4.2: LORAX screen design and results..... 117

Figure 4.3: Validations of LORAX screen identified Cas9 variants. 122

Figure 4.4: Validation of LORAX screen identified Cas9 variants for genome editing and de-immunization..... 124

Figure 4.5: Characterization of Cas9 variants V3 (2 mutations), V4 (7 mutations), and V5 (8 mutations) across genome and epigenome targeting assays. 126

Figure 4.6: Predicted and experimentally confirmed deimmunization across Cas9 epitopes.128

Figure 4.7: Delivery of de-immunized Cas9 variants as icRNAs for genome and epigenome targeting..... 130

LIST OF TABLES

Table 1.1: In vivo CRISPR-Cas strategies developed for functional rescue in animal models. 30

Table 3.1: CRISPR effectors analyzed for immune orthogonality. 65

Table 3.2: AAV VP1 orthologs analyzed for immune orthogonality. 81

Table 4.1: qPCR primers 142

ACKNOWLEDGEMENTS

I have been fortunate to receive outstanding guidance, support, and assistance throughout the work described herein. I would like to thank my advisor, Dr. Prashant Mali, for his expertise and insight which was absolutely critical in helping to design and guide all the research within this dissertation, and for his continual feedback, support, and attention to scientific rigor which pushed me to bring my work to an entirely new level. In addition, I would like to thank my committee members, Dr. Justin Meyer, Dr. Matthew Daugherty, Dr. Vineet Bafna, and Dr. Ananda Goldrath for their advice, encouragement, and feedback throughout this process. I feel very fortunate to have been mentored by such a wonderful, friendly, and expert group of scientific minds.

I would also like to acknowledge my colleagues in the Mali lab, especially my coauthors Ana Moreno and Aditya Kumar, as well as others in the lab including but not limited to Kyle Ford, Dhruva Katrekar, Amir Dailamy, Daniella McDonald, Udit Parekh, Michael Hu, Amanda Suhardjo, James Yen, Debbie Chen, Andrew Portell, Yichen Xiang, and Rebecca Panwala. Thank you for your tireless collaborations, unflagging patience and incredible insight. It has been an honor and a pleasure to work with all of you.

Chapter 1, in part, is a reprint of the material as it appears in The CRISPR Journal 3(4): 253–275 (2020). Tay, Lavina Sierra; Palmer, Nathan; Panwala, Rebecca; Chew, Wei Leong; Mali, Prashant. The dissertation author was the primary researcher and author of this paper. This work was generously supported by NIH grants (R01HG009285, RO1CA222826, RO1GM123313) and by the Agency for Science, Technology, and Research under its Industrial Alignment Fund (Pre-Positioning) (H17/01/a0/012).

Chapter 2, in part, is a reprint of the material as it appears in Nature Biomedical Engineering 3, 806–816. (2019). Moreno, Ana; Palmer, Nathan; Alemán, Fernando; Chen, Genghao; Pla, Andrew; Jiang, Ning; Chew, Wei Leong; Law, Mansun; Mali, Prashant. The dissertation author was the primary researcher and author of this paper. We thank members of the Mali laboratory for advice and help with experiments and the Salk GT3 viral core for help with the production of AAVs. This research was supported by UCSD Institutional Funds, the Burroughs Wellcome Fund (1013926), the March of Dimes Foundation (5-FY15-450), the Kimmel Foundation (SKF-16-150), and NIH grants (R01HG009285, RO1CA222826, RO1GM123313, R01AI079031 and R01AI106005). A.M.M. acknowledges a graduate fellowship from CONACYT and UCMEXUS. W.L.C. acknowledges the IAF-PP grant (H17/01/a0/012).

Chapter 3, in part, is a reprint of the material as it appears in Nature Biomedical Engineering 3, 806–816. (2019). Moreno, Ana; Palmer, Nathan; Alemán, Fernando; Chen, Genghao; Pla, Andrew; Jiang, Ning; Chew, Wei Leong; Law, Mansun; Mali, Prashant. The dissertation author was the primary researcher and author of this paper. We thank members of the Mali laboratory for advice and help with experiments and the Salk GT3 viral core for help with the production of AAVs. This research was supported by UCSD Institutional Funds, the Burroughs Wellcome Fund (1013926), the March of Dimes Foundation (5-FY15-450), the Kimmel Foundation (SKF-16-150), and NIH grants (R01HG009285, RO1CA222826, RO1GM123313, R01AI079031 and R01AI106005). A.M.M. acknowledges a graduate fellowship from CONACYT and UCMEXUS. W.L.C. acknowledges the IAF-PP grant (H17/01/a0/012).

Chapter 4, in part, is a reprint of the material as it appears in bioRxiv (2022). Aditya Kumar, Nathan Palmer, Katelyn Miyasaki, Emma Finburgh, Yichen Xiang, Andrew Portell, Amir Dailamy, Amanda Suhardjo, Wei Leong Chew, Ester J. Kwon, & Prashant Mali. The dissertation author was the primary researcher and author of this paper. Thanks to members of the Mali lab for discussions, advice and help with experiments. This work was generously supported by UCSD Institutional Funds, NIH grants (R01HG009285, R01CA222826, R01GM123313), Department of Defense Grant (DODPR210085), and a Longevity Impetus Grant from Norn Group. This publication includes data generated at the UC San Diego IGM Genomics Center utilizing an Illumina NovaSeq 6000 that was purchased with funding from a National Institutes of Health SIG grant (S10 OD026929). Some schematics were created using BioRender.

Chapter 5, in part, is a reprint of the material as it appears in The CRISPR Journal 3(4): 253–275 (2020). Tay, Lavina Sierra; Palmer, Nathan; Panwala, Rebecca; Chew, Wei Leong; Mali, Prashant. The dissertation author was the primary researcher and author of this paper. This work was generously supported by NIH grants (R01HG009285, RO1CA222826, RO1GM123313) and by the Agency for Science, Technology, and Research under its Industrial Alignment Fund (Pre-Positioning) (H17/01/a0/012).

VITA

2016 Bachelor of Science in Molecular Biosciences and Biotechnology, Arizona State University

2022 Doctor of Philosophy in Biology, University of California San Diego

PUBLICATIONS

Kumar, A*, **Palmer, N***, Miyasakia, K, Emma Finburgh, E, Xiang, Y, Portell, A, Dailamy, A, Suhardjo, A, Chew WL, Kwon, EJ, Mali, P (2022). Extensive in vitro and in vivo protein translation via in situ circularized RNAs ***bioRxiv***.

Tay LS*, **Palmer N***, Panwala R, Chew WL, Mali P. (2020) Translating CRISPR-Cas therapeutics: Approaches and challenges. ***The CRISPR Journal***. 3(4), 253-275. PMID: 32833535.

Moreno AM*, **Palmer N***, Alemán F, Chen G, Pla A, Jiang N, Chew WL, Law M, Mali P. (2019). Immune-orthogonal orthologues of AAV capsids and of Cas9 circumvent the immune response to the administration of gene therapy. ***Nature biomedical engineering***. 3(10), 806-816. PMID: 31332341.

Ford KM, Panwala R, Chen DH, Portell A, **Palmer N**, Mali P. (2021) Peptide-tiling screens of cancer drivers reveal oncogenic protein domains and associated peptide inhibitors. ***Cell Systems.*** https://doi.org/10.1016/j.cels.2021.05.002

Moreno AM, Alemán F, Catroli GF, Hunt M, Hu M, Dailamy A, Pla A, Woller SA, **Palmer N**, Parekh U, McDonald D, Roberts AJ, Goodwill V, Dryden I, Hevner RF, Delay L, Gonçalves Dos Santos G, Yaksh TL, Mali P. (2021) Long-lasting analgesia via targeted in situ repression of NaV1.7 in mice. ***Science Translational Medicine***. 13(584):eaay9056. PMID: 33692134.

Katrekar D, **Palmer N**, Xiang Y, Saha A, Meluzzi D, Mali P. (2020). Interrogation of the ADAR2 deaminase domain for engineering enhanced RNA base-editing activity, functionality and specificity. ***eLife*** 2022; 11:e75555 https://doi.org/10.7554/eLife.75555.

Petrie KL, **Palmer ND**, Johnson DT, Medina SJ, Yan SJ, Li V, Burmeister AR, Meyer JR. (2018). Destabilizing mutations encode nongenetic variation that drives evolutionary innovation. ***Science***. 359(6383), 1542-1545. PMID: 29599247.

ABSTRACT OF THE DISSERTATION

Engineering Progressively Deimmunized and Redosable CRISPR-Cas Gene Therapies

by

Nathan David Palmer

Doctor of Philosophy in Biology

University of California San Diego, 2022

Professor Prashant G. Mali, Chair
Professor Justin R. Meyer, Co-Chair

The advent of genomic editing technology offers the tantalizing promise of precisely treating a variety of disease states at the genetic level. CRISPR-Cas has democratized this capability in laboratory settings across the world. However, developing these exciting tools into therapies appropriate for clinical use involves several challenges including the potential recognition of foreign components such as the Cas effector proteins by the adaptive immune system which may limit the effectiveness of gene therapy. Here I approach this problem by first

characterizing the immune response to AAV-delivered CRISPR-Cas in a mouse model targeting the PCSK9 protein, known to play a major role in atherosclerotic disease. Next, I propose and test a method to enable redosing by leveraging immune orthogonal components of AAV-CRISPR-Cas therapeutics, systematically defining the capability of this approach to avoid immune-mediated inhibition of therapeutic efficacy. Finally, in an effort towards building a single deimmunized Cas9 protein usable across diverse contexts while circumventing pre-existing immunity, I develop a scalable and portable long-range multiplexed protein engineering platform to progressively de-immunize target proteins by abolishing the most immunogenic MHC-restricted epitopes without disrupting protein function. By applying this technique to Cas9, I identify a Cas9 variant with 7 simultaneously deimmunized epitopes that retains near wild-type functionality for both direct editing and gene activation/repression. Taken together, this work represents a meaningful step towards unlocking the potential to precisely edit genomes and gene expression in the clinical setting.

CHAPTER 1 – THE STATE OF CRISPR-CAS GENE THERAPY

Gene therapy holds the promise of treating genetic, epigenetic, and transcriptomic diseases at their roots. The first several decades of research and development in this field were met alternatingly with eager exploration and substantial setbacks.[1] Even so, persistence has finally begun to pay off with the first generation of gene therapies which replace missing or defective genes in conditions such as transfusion-dependent β-thalassemia, spinal muscular atrophy and inherited retinal disease.[2–4]9/1/22 12:28:00 PMFortuitously, the long-awaited emergence of these drugs coincides with the recent advent of genome and transcriptome engineering tools based on the clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR-associated (Cas) RNA-guided nuclease systems that have transformed our ability to precisely manipulate nucleic acids. CRISPR-Cas has expanded our toolkit, enabling precision alteration, replacement, and regulation of genes to therapeutically target disease states and potentially confer disease resistance. Beyond genetic diseases, CRISPR-Cas editing therapies have been used to target acquired diseases such as cancer, cardiovascular disease, infectious disease, and chronic ailments such as Alzheimer's disease (AD) in animal models.[5–8] The nexus of these two fields has created extraordinary excitement for the potential of CRISPR-based gene therapies to revolutionize the treatment of many of the most important ailments currently impacting human health.

While that potential has yet to be fully realized, pre-clinical studies of CRISPR-based therapies have exploded in recent years, and pioneering CRISPR-Cas gene therapy clinical trials have begun treating rare monogenic diseases and severe cancers. Still, broad application of CRISPR Cas therapies has a long way to go and hinges on multiple factors including long term studies to demonstrate safety without off-target activity that could potentially cause serious side effects, advancement in precise and effective delivery systems, immune system interactions,

efficient and cheap manufacturing, and effective regulation to democratize access to these therapies. In this review, we first broadly cover the steps necessary to take CRISPR-based gene therapies from the idea, through drug design and pre-clinical testing, to clinical trials. This process is complex and may be iterative with many decision points and important milestones along the way. We then focus in on the most critical challenges facing the development of these therapies, especially in delivery, administration, safety, design of appropriate pre-clinical studies, drug manufacturing, and economic and regulatory issues, describing the obstacles they present and how studies to date have approached these issues. We hope to offer both a window into the current state of the field and a map to orient thought and resources towards solving the most significant challenges that may impede the realization of the potential of CRISPR-based gene therapy.

1.1 THERAPEUTIC DEVELOPMENT CASCADE

Development of a gene therapeutic is a complex endeavor,[5,9] with the path towards enabling a new therapy following a multifaceted development cascade (Figure 1). We briefly summarize this process below.

First, is the selection of payload, which in turn is governed by the nature of the edit desired, and considerations of enabling activity at the target locus while ensuring high specificity. In addition, delivery constraints such as vehicle size capacity and interaction with native human biology, such as immunity, availability of appropriate PAM sites, and the dynamic range of therapeutic effect also affect this choice.

Second, is the selection of a vehicle for *in vivo* delivery. Adeno-Associated Viruses (AAVs) are commonly used in this regard as they deliver the payload without insertion into the host genome by maintaining the packaged DNA in an episome. AAVs work well in post-mitotic cells and can be made tissue-specific via use of tissue-specific promoters or selection of

appropriate serotypes.[10] Efforts have been made to enhance the safety profile of AAVs through investigating how AAVs interact with and evade host immunity and engineering new capsids to mitigate the immune response.[11] And while cargo space of AAV limits the size of packaged constructs to ~5kb, recent studies have uncovered Cas proteins that are considerably smaller than *Streptococcus pyogenes* Cas9, some of which show similar functional promise.[12,13] Notably, CRISPR-Cas can also be delivered via non-viral vehicles, with purified ribonucleoproteins and lipid nanoparticles showing promise for therapeutic translation.

Third, is determining the route of therapeutic delivery. Treatments delivered to the brain, spinal cord, and eye benefit from their status as immune-privileged, and therefore are usually more tolerant to foreign antigens used in gene therapies.[14] For this reason and also relative ease of delivery, some of the first clinical trials of CRISPR-Cas9 therapies have been initiated against *ex vivo* hematopoietic cells, or ocular disorders where treatment is delivered via subretinal AAV injection (NCT03872479). Nonetheless, it is important to note that this immune privilege can also be breached when the blood-brain barrier is compromised in certain pathological states. The timing of the administration is also key, as studies have shown that the earlier the treatment is given in the course of disease progression, the higher the potential for efficacy.[15] Associated with this, are also studies to assess the distribution, persistence and clearance of the delivery vector and the expression of the therapy in the target tissues.[16] A challenge to gene therapies can be the occurrence of off-target tissue effects, where expression occurs in unintended tissues or organs (contrasted with off-target genetic effects where edits occur at unintended genomic locations within the target cells). This can be minimized when therapies are administered to confined spaces, such as the eye or joints[17] and through the incorporation of tissue specific promoters to drive gene expression.[18]

Fourth, is the selection of an appropriate animal model to obtain clinically relevant insights into the efficacy of the therapeutic. In this regard, there are a variety of challenges associated with testing gene therapies in animal models, including replication of human diseases, response to therapy, and differing immune reactions. However, as a result of advances in genome editing technologies, many animal models with specific alterations able to replicate clinical phenotypes have been generated.[19] In addition, the use of human immune system mouse models has improved the translatability of animal studies to the human setting.[20]

Fifth, is the crafting of the overall study design. There are a number of factors that need to be considered when planning a preclinical study, including appropriate controls, number of animals per condition, testing across differing dosing and the time points at which samples are taken.[21] When utilizing viral vectors for delivery, gene expression can begin after one to two weeks, and extend for a period of months, during which samples are assessed.[22] Associated with the above is the setting of endpoints and measurement of clinically relevant phenotypes in order to assess the safety and efficacy of the treatment. To determine safety, toxicology studies are performed to establish potential local and systemic toxicities, safety and feasibility of the delivery system and procedure, immune response against the vector and delivered construct, and the potential for reproductive toxicity. To examine efficacy, biodistribution studies are performed, and clinical signs relevant to the disease measured.[16]

Finally, once preclinical experiments are run and analyzed, and safety and efficacy of the therapy are demonstrated, an IND application is prepared. The IND application contains information about the animal pharmacology and toxicology, manufacturing and clinical components, and materials and protocols[21], and is submitted to regulatory agencies as part of advancing into clinical trials.

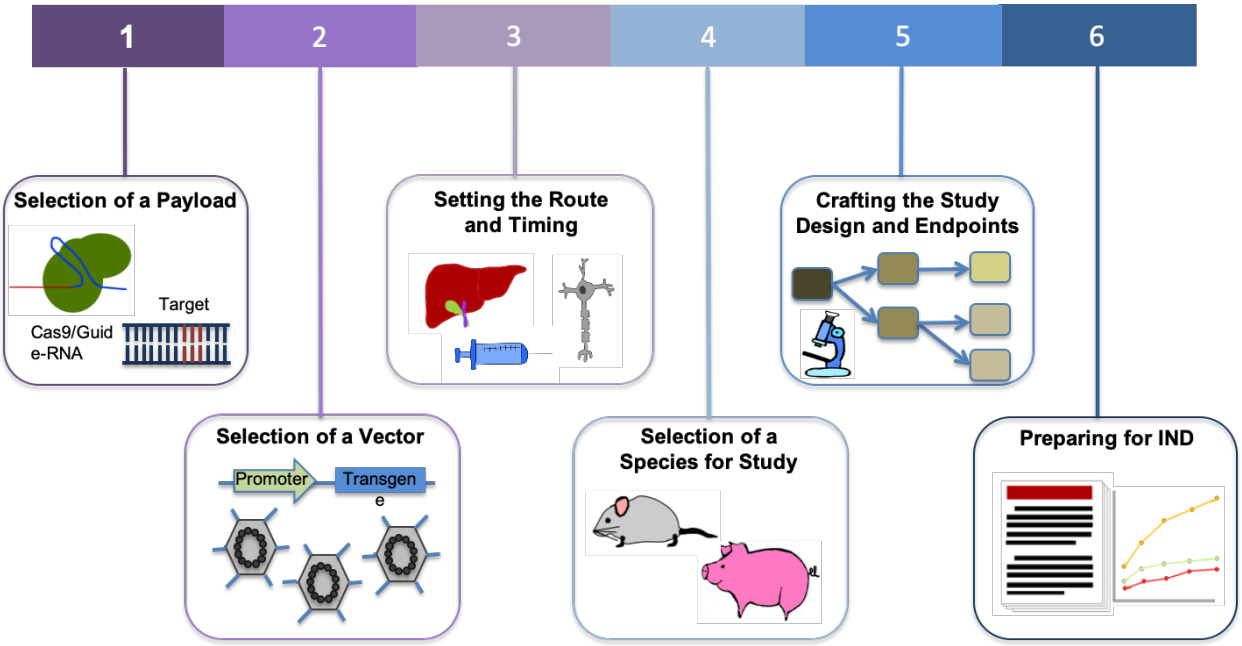


Figure 1.1: Steps towards enabling a gene therapy via CRISPR-Cas. For a disease target, the choice of Cas effector and corresponding guide-RNA (gRNA) is the first step. Next is to determine CRISPR-Cas delivery options via viral or non-viral methods. Associated with delivery system is also the route of administration. This typically depends on the organ or tissue of treatment and ideally should be limited to that location. Studies are then performed to assess the distribution, persistence and clearance of the vector and the expression of the therapy in target tissues. Relevant clinical insight is also dependent on the selection of species for study and a preclinical study design is crafted. Clinical endpoints are set based on safety and toxicology studies. Proof of safety and efficacy in preclinical studies will then lead to an IND application.

Below we discuss CRISPR-Cas based gene therapeutics development in the context of this cascade. We highlight the progress made so far, as well as pending challenges and discuss potential approaches to address some of these.

1.2 DEVELOPING CRISPR-CAS THERAPEUTICS

1.2.1 Payload

Among the most likely candidates for translating CRISPR-Cas components into functional therapeutics are the RNA-guided DNA nucleases Cas9 and Cas12a, the RNA-guided RNA

nucleases RCas9 and Cas13, and variants of these proteins. Different therapeutic modalities within these Cas families unlock a wide variety of clinical applications. The native nuclease function of Cas9 or Cas12a can be directly applied to genome editing in the form of gene knockouts or gene replacement. The resultant double-stranded DNA break is repaired by native cellular machinery, typically through the non-homologous end-joining (NHEJ) pathway[23–26] or homologous recombination with an endogenous sequence or an exogenously provided sequence. Further therapeutic approaches towards genome editing with Cas9 have been demonstrated with base-editing and prime-editing, with the former currently demonstrated and more efficiently used for single-base transitions while the latter demonstrated for a wider range of localized edits albeit with lower efficiencies. Additionally, targeting diseases at RNA level can be appealing as it prevents any permanent off-target effects, making it less risky for clinical applications. RNA-targeting RCas9 and Cas13 have been used for RNA knockdown with similar or comparable efficiency to RNAi.[27,28] Common RNA editing strategies rely on naturally occurring human adenosine deaminase acting on RNA (ADAR) enzymes. Modification of ADAR2 and fusing it to a catalytic inactive Cas13 known as RNA Editing for Programmable A to I Replacement (REPAIR) was able to recognize and correct disease mutations without any PAM sequence constraints.[29–31] In recent years there has been increased interest in developing Cas-mediated transcript editing into RNA therapeutics with applications in treating viral infection or alleviating symptoms in neurological disease. Together, these CRISPR-Cas modalities enable a broad spectrum of strategies towards disease treatment (**Figure 2**).

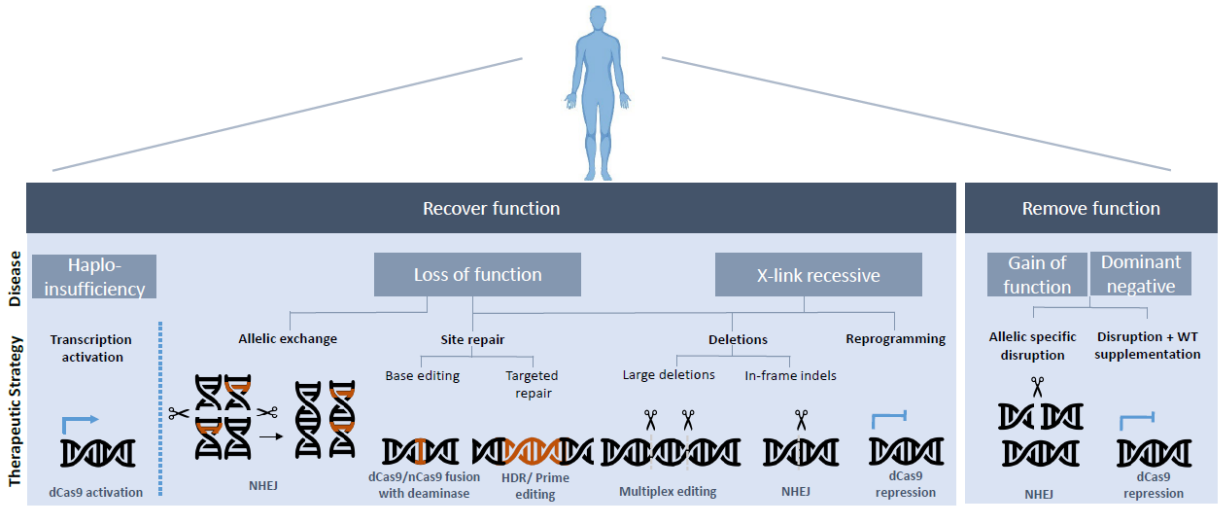


Figure 1.2: CRISPR-Cas therapeutic strategies towards disease-causing mutations.

1.2.2 Delivery

The CRISPR-Cas therapeutic cargo is important but not enough. Delivery represents one of the greatest challenges in developing a clinically viable CRISPR-Cas therapy and hence is a critical component of the therapeutic cascade. CRISPR-Cas systems can be delivered in three cargo formats, via DNA modalities encoding both the Cas protein and gRNA (such as plasmids, DNA viruses, nanoparticles), RNA modalities encoding the Cas and the gRNA (modified RNA, RNA viruses, *in vitro* transcribed RNA, nanoparticles), or direct delivery of the purified Cas protein complexed with the gRNA as a ribonucleoprotein (RNP) or nanoparticle. Each of these cargo formats have been used *ex vivo* (in cells outside the body) and *in vivo* (locally within specific tissues or systemically through the body) (**Figure 2**). Importantly, many of these delivery methods are built upon the framework of traditional gene therapy and hence share much of the technological underpinnings and desired properties. The ideal delivery system will likely encompass the following attributes: (1) non-integrating: limiting unintended integration events in the genome; (2) specific: achieving targeted tissue delivery; (3) low immunogenicity: enabling low toxicity

associated with administration and sustained therapeutic effect; (4) option to re-dose: allowing repeated regimens; (5) biodistribution: compatible with target tissue type and tissue accessibility, and (6) scalable: compatible with large-scale manufacturing for both common and orphan diseases. We describe below how these different delivery modalities have been used for CRISPR-Cas and how they might exhibit some of these ideal attributes.

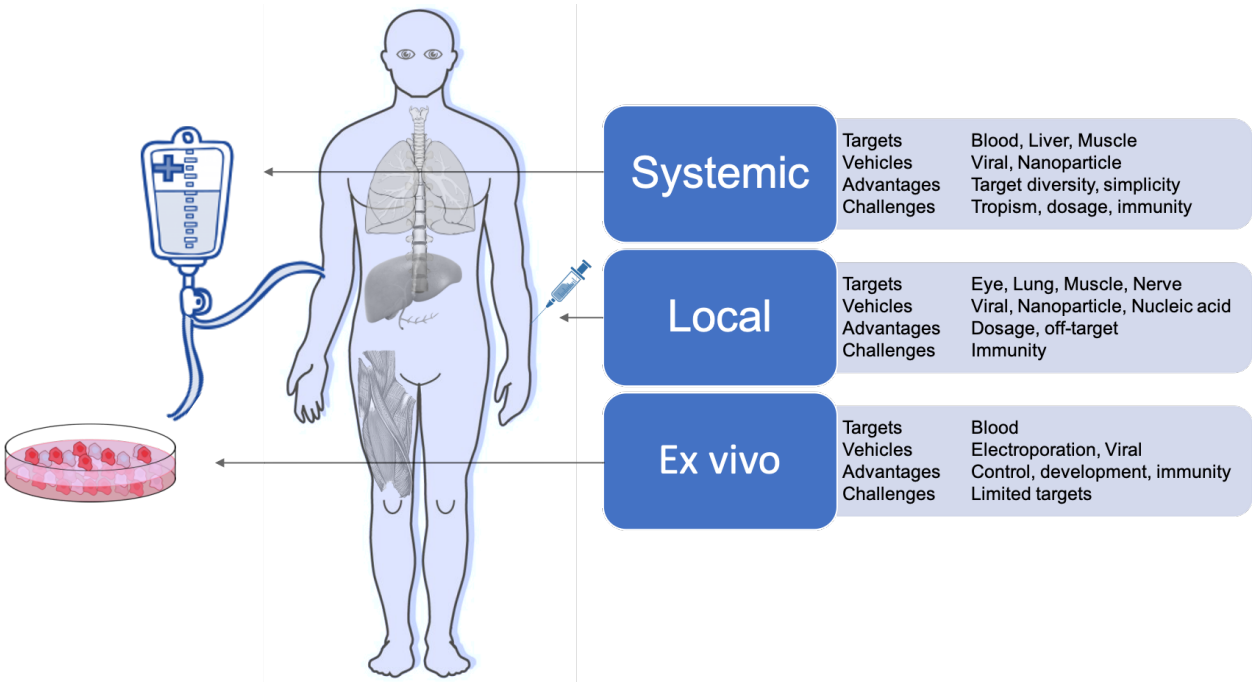


Figure 1.3: Delivery modalities for CRISPR-Cas therapeutics.

1.2.2.1 Ex vivo delivery modalities

Ex vivo gene-editing allows cell engineering and quality control regimes that would not be feasible within the human body and is most compatible with blood disorders owing to the fact that blood is much easier to safely remove, treat, and re-infuse than other tissues. A modality successfully used for *ex vivo* gene editing is Cas proteins complexed with sgRNAs *in vitro* as RNPs. Unlike DNA- or RNA- encoded CRISPR, RNPs do not rely on cell-driven expression and hence have the fastest activity onset upon administration. In addition, RNPs are not further

expressed and degrade over time, with this transient exposure window limiting potential off-target effects.[32] The most clinically advanced CRISPR-Cas editing strategy relies on *ex vivo* RNP-mediated editing followed by re-administration of the cells back to the donor.

Advances in the isolation and propagation of hematopoietic stem cells (HSCs) and immune cells has promoted cellular gene therapy as a viable option for the treatment of some monogenic diseases and malignancies. Early successes were complicated by uncontrolled insertional mutagenesis of the vector resulting in transcriptional activation of nearby proto-oncogenes, chromosomal translocations, and eventual leukemogenesis.[33] Unlike traditional gene therapy requiring sustained transgene expression, CRISPR-Cas systems can achieve therapeutic outcomes by transient expression allowing for less integrative risk and potentially safer methods of delivery. Current *ex vivo* CRISPR-Cas mediated clinical trials utilizes direct delivery of Cas9 RNP via electroporation (NCT03399448, NCT04035434 and NCT03655678).

Curative approaches for β-hemoglobinopathies includes allogenic hematopoietic stem cell transplant (HSCT) using HLA-matched stem cells derived from a donor or autologous transplantation of the patient's own HSCs after receiving the gene therapy *ex vivo*. Allogenic transplant is challenging due to dependence on finding an immune-matched donor and mitigating the effects of graft-versus-host response. Autologous transplantation with gene edited cells avoids much of this immune barrier. Although many indications require varying levels of conditioning treatment to reduce the population of native HSCs to create space for donor cells, autologous *ex vivo* gene therapy may also allow for a reduced intensity of this process.[34] The current approach towards HSC gene therapy uses electroporation of CRISPR-Cas mRNA or RNPs to edit HSC genomes.[35] If homology directed repair is required, the template is simultaneously electroporated or transduced by viral vectors.[36] Following a similar concept, the first two patients with β-

thalassemia and sickle cell disease were treated last year using CRISPR-Cas9-edited HSCs and autologous transplantation.

The RNP based approach has also been most successful in progressing towards clinical trials. For instance, Cas9 RNPs were electroporated into HSCs to disrupt BCL11A, a silencer of fetal hemoglobin, to generate CTX100 cells. Disruption of BCL11A de-represses fetal hemoglobin expression and rescues sickle cell defects.[37,38] A long-term follow-up study is underway to assess safety and efficacy of the treatment. Another therapeutic candidate, CTX110, utilized the multiplexing capability of CRISPR-Cas systems to optimize CAR-T cells at multiple levels. This approach results in replacement of the native TCR with an anti-CD19 CAR as well as knockdown of MHC I expression via targeting β2M (NCT04035434). Positioned as "off-the-shelf" therapy, using allogenic donor T-cells could reduce the cost of single batch production. Another ambitious study uses Cas9 RNPs to disrupt PDCD1, TRAC, and TRBC, and introduce transgenic NY-ESO-1-specific T-cell receptor (TCR) to generate engineered T-cells for a clinical trial last year.[5] The controlled environment of cell engineering and quality control (QC) relative to *in vivo* administration is the reason that *ex vivo* pipelines have been among the first to progress to the clinic, and will inform the uptake of genome editing for *in vivo* use.

1.2.2.2 In vivo delivery modalities

1.2.2.2.1 AAVs

Viral vectors are the most common *in vivo* delivery vehicles for CRISPR-Cas. Viruses have substantial differences in their safety profiles, with AAVs having one of the more favorable profiles. The recently approved gene therapies Luxturna[3] and Zolgensma[4] both deliver their therapeutic transgenes via AAVs. Soon after, the first *in vivo* CRISPR-Cas therapy entered clinical

trial, in which AAV5-CRISPR-Cas9 was delivered into the eye for safety and efficacy evaluation (NCT03872479). These clinical and commercial successes would likely continue to make AAVs attractive for upcoming therapeutic development pipelines.

AAVs are replication-defective viruses isolated from members of the Parvoviridae family and naturally infect humans and non-human primates (NHPs). Following infection, the DNA carried by AAVs achieves long-term episomal expression without genomic integration. AAVs are endemically found and have several variants which exhibit different tropism for various tissue types, such that a serotype can be selected to best target the tissue and disease indication of interest.[39] Significant work has been devoted to AAV capsid protein engineering. For example, rational capsid engineering has been used to increase neuronal transduction by mutating surface exposed tyrosine residues and phosphorylating threonine residues on the AAV2 capsid.[40] Altering vector capsids can also increase tropism towards tissues that are challenging to target, such as microglia.[41] Chimeras with new functions can be derived from peptide domain exchanges among different serotypes, such as AAV2.5 that has improved muscle transduction along with a different immune profile from both parental AAV1 and AAV2 serotypes, offering a potential strategy for immune bypass and therapeutic re-dosing.[42] More complex chimeric capsid libraries can be generated by DNA shuffling of capsid sequences of multiple AAV serotypes to generate greater variation and accelerate the evolutionary process.[43] Novel capsids have also been generated by using PCR-based mutagenesis and subsequent selection for specific attributes.[44] For instance, directed evolution has been an efficient approach to derive tissue tropisms in iPSCs[45] and human ciliated airway epithelium.[46]

Viral protein engineering has also been brought to bear on enabling safety of gene therapy vectors. Even though infection with AAVs is only mildly immunogenic and not associated with

disease, the remaining immunogenicity still represents a significant safety and efficacy challenge to overcome. The effective use of AAV vectors can be compromised by a pre-existing immune response from natural exposure to AAVs in early life or from prior AAV mediated therapy.[47,48] Adaptive immune response to AAVs leads to production of neutralizing antibodies (NAbs) that block transduction and clear the viruses. Adaptive cellular response to AAVs by CD8+ cytotoxic T lymphocytes leads to eradication of the transduced cells and can attenuate treatment efficacy and imposes safety risks. Various strategies have been developed to overcome these issues, one of which involves engineering the AAV capsid epitopes to avoid antibody and T-cell recognition. Immune reactive epitopes have been mapped for some serotypes and capsid mutagenesis can reduce binding and neutralization effects.[49,50] However, it is not always straightforward to engineer epitopes while maintaining desired functionality. Often such modifications can be limited and result in loss of function, as exemplified by another study where an AAV9-specific neutralizing epitope was mapped to capsid residues conferring liver tropism, and mutation on this did not provide a successful tradeoff in evasion from polyclonal antibodies.[51]

In addition to induced or pre-existing adaptive immune responses, AAVs trigger a natural innate immune response through recognition of pathogen-associated molecular patterns (PAMPs) inherent to AAV biology.[52] Multiple studies have elucidated some of the mechanisms of innate immune activation by AAVs including binding by TLR9 of unmethylated CpG dinucleotides in the AAV genome leading to activation of the MyD88/NF-κB pathway[53–55], and recognition by MDA5 of long dsRNAs created by bidirectional transcription from the AAV episome.[56] These pro-inflammatory signals result in substantial transcriptional changes, inducing an anti-viral state within the target cell, and expression of immune stimulating cytokines that allow for a robust adaptive immune response.[57] In the clinic, patients' cytokine levels are routinely monitored, and

sometimes modulated with corticosteroids.[58] In contrast to other viral vectors, AAVs have not demonstrated a clinically dangerous level of innate immune activation in humans, albeit certain cell types may be susceptible to damage reducing clinical efficacy.[59] Nevertheless, minimizing innate immune induction remains an important goal in AAV vector biology to maximize patient safety and limit adaptive immune induction.

More recently, AAV engineering has ventured beyond protein engineering towards chemical and biophysical means. Chemical modification introduces synthetic motifs to the surface of AAVs. An early chemical approach masked exposed arginine residues by a naturally occurring glycation reaction.[60] Very recently, a site-specific approach was developed using unnatural amino acids on AAV capsids to attach synthetic ligands onto these residues through highly specific and biocompatible small molecule chemical reactions termed "click chemistry". This strategy was used to tether ligands on the AAV capsid to achieve tissue-specific targeting and protection against antibodies.[61,62] Another approach involves encasing the AAV in a membrane exosome allowing for greater customization of the external layer of the drug. Exosome enveloped AAV vectors achieved enhanced transduction and protection against pre-existing neutralization antibodies.[63–66] A recent study demonstrated the stabilizing effects of tetraspanin CD9 increasing exosome production.[67] While it is an attractive solution, manufacturing methods for scalable production of exosomes will be challenging.

1.2.2.2.2 Nanoparticles

Non-viral modalities are also of significant potential. Various nanomaterials such as polymers, lipids, proteins and metals have been explored for *in vivo* delivery. Extensive research has been committed to lipid nanoparticles (LNPs) and there are currently more than ten FDA approved uses of LNPs for drug delivery.[68] Onpattro is a recently approved multi-dose gene

therapy utilizing LNP to deliver siRNA for treatment of polyneuropathies. LNPs typically comprise cationic lipids that facilitate endocytosis and release into cytoplasm and PEGylated lipids that facilitate stabilization and prevent non-specific interaction.[69] PEGylation also reduces immunogenicity and increases circulation time.[70] Recently, an LNP-based delivery system incorporating helper lipid and PEG-DMG to encapsulate Cas9 mRNA and modified sgRNA was able to effect durable editing of approximately 70% in the liver and benefited from a multiple dosing regimen.[71] Other earlier studies targeting the liver by LNPs showed robust knockdown[72] and a HDR editing rate of 6%.[73] Outside of the liver, however, application might be more limited as lipid particles are usually taken up and metabolized in the liver during systemic circulation,[74] thereby constraining the targetable tissue types for LNPs.

Another recent modality utilized gold nanoparticles, which can be used to bind Cas9 RNPs and conjugate with 5'-thio ssDNA for further hybridization with DNA repair templates.[75] The nanoparticle is then coated with negatively charged silica for cationic polymer PAsp(DET) encapsulation, following which cytoplasmic glutathione releases the contents from the gold nanoparticle.[76] Intramuscular injection of CRISPR Gold Cas9 RNP with template DNA was able to correct 5.4% of dystrophin gene. Intracranial injection of CRISPR Gold Cas9 RNP resulted in localized gene editing of 14.5% of the mGluR5 gene and effect 40-50% of the protein production in the brain and multiple tissue types, rescuing the effects of ASD in diseased mice.[77] The gold nanoparticles are well tolerated in the neurons, however, the long-term accumulation and elimination kinetics of these gold nanoparticles remain to be evaluated.

1.2.2.2.3 RNPs

Beyond use of delivery vehicles, established protocols for large-scale production of proteins makes direct delivery of the Cas9 RNP complex attractive for clinical translation. Many

polymers are being evaluated for coupling with the Cas9 RNP complex, which has a heterogeneous charge distribution. A mix of cationic and anionic monomers with imidazole can fully encapsulate the RNPs while facilitating endosomal escape. Glutathione cleavable link N,N′-bis(acryloyl)cystamine (BACA) around the RNP enables cytosolic release in the presence of glutathione, but in order to boost the efficacy of delivery and therefore editing, a further attachment of tissue-specific ligands to one end of the PEG is necessary.[78] While poly(aspartic) acid (PAsp) based polyplexes are biocompatible with limited toxicity, modified PAsp(DET) bearing 1,2-diaminoethane side chains showed significantly higher transfection efficiency.[79] Recently, PAsp(DET) assembled nanoparticles were able to effectively deliver Cas9 RNP complexes into muscle and brain tissues.[75,77] A screening effort was also able to identify poly(aspartic) acid polymer (PAsp) analogs to efficiently deliver Cas12 RNP in mice muscle fibers.[80]

1.2.3 Administration

1.2.3.1 Manufacturing and scale-up challenges

In general, manufacturing of pharmaceutical agents must conform to current Good Manufacturing Practice (cGMP). This often refers to a body of regulatory law set forth in US FDA Code of Federal Regulations (CFR) title 21, though other countries and organizations, particularly the WHO and EU have defined their own similar standards for cGMP. Although the application of GMP may vary depending on the wide range of pharmaceutical agents and production methods, generally it requires the use of dedicated clean production facilities, controlled starting materials and cell lines, and multiple product purification steps with accurate testing at each step. cGMP regulation is often focused on process controls and extensive documentation such that a wide variety of issues, should they arise, can be easily detected and corrected. The components of a

cGMP production system will vary depending on the agent being produced and must be evaluated on a case-by-case basis.

1.2.3.1.1 Nucleic acid manufacturing

The foundational starting material for generating CRISPR therapeutics is nucleic acid, either as synthetic gRNA, DNA template for gRNA transcription, plasmid encoding the Cas effector protein, or plasmid encoding delivery vehicle components, e.g., viral vectors. The manufacturing specifications required for these nucleic acids may differ depending on the use case but will generally need to conform at minimum to High Quality (HQ) grade manufacturing for vector plasmids and template DNA not used in the final therapeutic, and ideally to full cGMP grade for direct applications. HQ grade DNA can be produced in dedicated facilities using strains drawn from a research cell bank and utilizing two to three separate chromatography steps to eliminate linear, open, and chromosomal DNA and remove LPS and other cellular contaminants. The main differences when upgrading to cGMP production are the necessity of a GMP cell bank starter, a dedicated GMP facility, a validated quality assurance (QA) system with in-process control (IPC), and full documentation procedures. For a more thorough discussion of these procedures and QC validation requirements see ([81]).

Despite these stringent requirements, production of cGMP-compliant, therapeutic-grade nucleic acid is substantially easier than recombinant protein due to the fewer purification steps required as the reaction mixtures are much less complex than those involved in bacterial or mammalian cell culture. Additionally, like protein therapeutics, production of DNA- and RNA-based therapies have been under rapid development in the last two decades due to excitement surrounding the potential for nucleic acid vaccines and RNAi drugs. Utilizing many of these same facilities and procedures, multiple companies currently offer cGMP preparations of sgRNA for

clinical and preclinical CRISPR research, as well as HQ and cGMP viral vector preparations. The cost for these materials is manageable, except for a few cases requiring truly massive quantities of DNA such as low-yield AAV vectors needed for systemic delivery. Nucleic acid production with cGMP guidelines is a feasible and preferred option for more typical RNP and high-yield AAV serotype delivery modalities.

1.2.3.1.2 Manufacturing of non-viral RNPs

The current non-viral formulations for CRISPR-Cas range from simple to difficult for manufacturing, with each carrying its own particular set of manufacturing challenges for creating therapeutic-quality products. The most clinically advanced non-viral delivery modality is also one of the simplest to produce, with electroporated RNPs consisting of only two purified components, namely a gRNA and a Cas protein. The *in vitro* transcription or *de novo* synthesis of gRNAs is well-established. Cas proteins like Cas9, Cas12a, or engineered variants would be recombinantly produced via scalable fermentation in pipelines similar to thousands of other protein therapeutics. After production in fermentation vessels, the resultant product is run through multiple chromatography steps to concentrate the active product and remove impurities.

Much of the initial recombinant protein production for pharmaceuticals was dominated by classic producers such as *Escherichia coli* and *Saccharomyces cerevisiae,*[82] but since then most biopharmaceutical production has been deployed using mammalian cell lines, especially Chinese Hamster Ovary (CHO) cells or human cells, due to the superior protein-folding intracellular environment and post-translational modifications.[83] Still, microbial production retains an important niche owing to its advantages in rapid growth, low nutritional requirements, and ease of genetic manipulation.[84] Indeed, Cas effector protein production seems to fit nicely into this niche based on their bacterial origin and independence of complex folding cofactors or post-translational

modifications. Process improvements in bacterial production such as the deployment of an endotoxin-free *E. coli* strain[85], metabolic engineering efforts[86], and adaptive laboratory evolution[87], will yield additional improvements directly applicable to CRISPR-based protein therapeutics. Furthermore, new exploration of alternative manufacturing methods, such as co-production of Cas9 protein and sgRNA in *E. coli* may open up new avenues of inexpensive drug production.[88]

Nevertheless, current manufacturing methodologies are poised to implement recombinant Cas proteins at scale. Commercial recombinant Cas9 preparations, such as Aldevron's SpyFi™ Cas9 nuclease, have already been made available. Importantly, associated procedures should be equally amenable to production of other Cas9 effectors including dCas9, Cas12a, and variants thereof. Over the past decades, pharmaceutical companies have been under extreme incentives to optimize scaled protein production environments[89], including moving to continuous flow, rather than batch setups[90], due to lucrative demand for protein therapies, especially monoclonal antibodies (mAbs). Process advancements brought about by this economic landscape will translate well to CRISPR-based protein drugs.

The RNA component of the CRISPR-Cas RNP can be comprised of a synthetic crRNA and tracrRNA pair as in the native bacterial system[91], a synthetic combined sgRNA[92], or an *in vitro*-transcribed sgRNA[93]. Single gRNA constructs are preferable due to simplicity and elimination of the duplex formation step. Both synthetic and *in vitro*--transcribed gRNAs have associated benefits and drawbacks. Chemically synthesized gRNAs are amenable to modification of one or more of the nucleotide bases, which has been shown to provide advantages such as increased stability[94] and reduced immunogenicity[95]. New modifications to further optimize these parameters are currently being explored.[96,97] The potential advantages of *in vitro*-transcribed

gRNA production include simple and economical scale-up to large volumes, as generation of template can be done with PCR. Both *in vitro*-transcribed and chemically synthesized gRNAs are being evaluated in ongoing clinical trials.[5,98]

1.2.3.1.3 AAV manufacturing

AAVs are the prime viral delivery vehicle being explored for *in vivo* CRISPR therapies due to several beneficial properties including low immunogenicity, non-integration, different serotypes available for different tissue tropism, and a proven clinical record. Nonetheless, despite decades of research and development into AAVs, manufacturing challenges still exist.

Typically, AAVs are produced using a triple transfection method in which plasmids encoding 1) the recombinant genome to be delivered, 2) the viral packaging genes, and 3) adenoviral helper genes are transfected into cells, most commonly human embryonic kidney (HEK) 293 lines using chemical agents such as polyethylenimine (PEI). Many AAV products for clinical trials have been produced using these methods (NCT02122952, NCT25322757, NCT21031578, NCT27453480). Initially this was done using adherent HEK293 cells cultured in 2D on stacks of flasks. These formats can suit typical doses for early trial phases that range from 10^{11} to 10^{13} vector genomes (VG) per patient. However, achieving titers high enough for later phases or larger doses would mean hundreds of flask stacks and many labor months.[99] This obstacle is most pressing for systemic disease targets such as Duchenne muscular dystrophy (DMD), which require larger doses to achieve good efficacy. Current recommendations for clinical doses are as high as 10^{15}, even 10^{16} vg per patient. The product volumes required for clinical trials, especially critical Phase III trials with large patient cohorts, are simply impractical using traditional cell culture plasticware.

As a result, both academic and industrial facilities have also adopted more scalable platforms to produce AAVs, particularly in fixed-bed bioreactors with adherent HEK293 cells[100] and bioreactors with suspension HEK293 cells[101]. Both have provided scalability without a complete process overhaul. Still, yields using these methods have largely been unable to exceed 10^{14} VG per liter, necessitating thousands of liters for larger clinical trials. Additionally, scaling this process up requires non-trivial quantities of GMP-grade plasmids which, though easily obtained in smaller quantities, also face scalability problems[102]. One approach to the transfection scalability issue is to deliver the necessary genetic material using viral carriers, typically a herpes virus or baculovirus, rather than by chemical transfection[103]. This approach has led to a moderate increase (2-5 fold) in yield[104], and interestingly, also an increase in AAV product infectivity in the case of herpes virus production, potentially due to superior genome loading into AAV capsids in the presence of helper genes from the herpes virus.[105]

This yield challenge has also prompted the use of baculovirus-infected insect Sf9 cells for AAV production, which provides a much higher yield than HEK293 and has since been used in multiple clinical trials and production campaigns[106–108]. Nonetheless, a recent revelation that Sf9-produced AAVs differ qualitatively from HEK293-produced AAVs potentially creates a roadblock that needs to be further investigated[109].

Additional process developments, including continuous flow-based production methods, and optimization of producer cell lines may further increase AAV yields. Still, all AAV production methods require very significant downstream processing including multiple filtration steps, size exclusion chromatography, ion exchange chromatography, and ultracentrifugation to achieve the required purity and enrich for properly packaged, infectious vectors.[107,110] Even applying the many

technical advancements recently developed, scalability of AAV vector production remains a huge challenge, especially for relatively common diseases requiring large effective doses such as DMD.

1.2.4 Pre-IND and pre-clinical development of CRISPR-Cas therapeutics

Successful gene therapies require preclinical validation in animal models. Different preclinical models have been developed to collect data in three main areas: proof of concept for efficacy, determining dose ranges for the chosen route of administration, and evaluating toxicity and safety profiles. For AAV-CRISPR-Cas, four main organs have been most successfully and commonly targeted. This first wave of success builds on the decades of understanding in gene therapy and will itself guide future CRISPR therapeutic development.

1.2.4.1 Muscle

There is significant morbidity and mortality in hereditary disorders of the muscle, and gene therapy offers promising outcomes for these unmet needs. Moreover, studies in the last few decades have also demonstrated the effectiveness of using skeletal muscles as protein factories to express transgene, contributing to the first gene therapy product to pass regulation and be approved[111].

DMD is one of the most common hereditary muscular disease in children and is caused by mutations in dystrophin, the largest gene in the genome. Recently, progress in gene editing technologies has raised hopes for successful restoration of the dystrophin gene. Conventional gene therapy has been challenging due to the AAV packaging limit, hence efforts were focused on truncated versions of dystrophin for gene replacement, such as microdystrophin. The therapeutic strategy of AAV mediated gene therapy in muscles have shown to be safe in preclinical and clinical studies but limited therapeutic effects have been shown for the treatment of DMD[42]. Patients with

large frameshift deletions also risk T-cell directed destruction against the new epitopes[112,113]. The lack of long-term effects could also be due to constant muscle degeneration and regeneration, diluting the expression of the non-integrated transgene overtime. Another approach in the treatment of this lethal disease is the use of CRISPR-Cas gene editing to permanently restore endogenous dystrophin expression. Diverse CRISPR-Cas gene editing strategies have been employed to treat dystrophy mouse models, including exon skipping, exon deletion, adenine base editors and HDR (**Table 1.1**)[114–117].

Because DMD affects muscle groups throughout the entire body, the challenge is to develop a single injection sufficient for whole-body therapy. The approach to treat DMD systemically via intravascular injection is advantageous as it allows targeting of all muscles groups and has been shown to be less immunogenic[118]. Durable levels of dystrophin were detected in cardiac and skeletal muscles in systemically treated neonatal mdx mice with AAV-SaCas9[119]. Moreover, systemic delivery of AAV9-spCas9 and AAV9-sgRNA corrected frame-shifted mutant DMD restoring up to 90% dystrophin in a canine model of DMD[120]. However, intravascular administration will require high doses of AAV to reach therapeutic levels of gene editing in clinics. Current Sarepta Therapeutics clinical trial NCT03375164 efficacy dose is at $2x10^{14}$vg/kg, where a single treatment of a patient will require 10^{15} vg of AAV vectors. One approach to reduce AAV dose is the use of self-complementary AAV (scAAV), which in one study has shown up to 20-fold dose reduction[121], although the large payload requirements with CRISPR-Cas raise limitations on potential use of this strategy.

1.2.4.2 Eye

Gene therapy in the eye is attractive due to the anatomical compartmentalization allowing targeting of specific cells with minimal exposure to other organs and the immune system. Specific

cell types in the eyes such as retinal pigment epithelium, ganglion, and photoreceptors can be effectively transduced by AAVs, and surgical processes can provide access to the various structures in the eye[122]. Non-invasive methods have been established for evaluating therapeutic effects, while the contralateral eye can be used as a control for evaluating the effects of the gene therapy. Lastly, numerous genetic animal models have been generated for diseases in the retina, paving way for gene therapy development[123].

Leber congenital amaurosis (LCA) is a childhood-onset blindness caused by mutations in at least one of 18 different genes. Type II is the most studied form with mutations in RPE65 gene. Gene replacement of RPE65 in naturally occurring RPE-negative blind dogs demonstrated improved vision through simple behavioral tests and more importantly the therapy demonstrated cortex recovery and response to visual stimulation despite prolonged visual deprivation in RPE65 mutant dogs[124,125]. The results were indicative of the safety of AAV2 subretinal injection and some level of restoration of vision when RPE65 was first introduced in LCA patients[126–128]. Re-administration of the therapy in the second eye was also successful in large animal studies despite elevated serum NAb towards AAV2[129,130] and was predictive of the results in human clinical trial[131]. AAV2 transfer of RPE65 showed therapeutic efficacy in phase III randomized trials and has been marketed as Luxturna. Closely following the landmark success of Luxturna, the first *in vivo* CRISPR-Cas gene editing human trial program was developed to treat LCA type 10. In this case, subretinal injection of AAV5 delivering SaCas9 is driven by a human rhodopsin kinase (hGRK1) promoter to repair the splicing defect of CEP290. hGRK1 promoter is able to drive robust transgene expression in both cones and rods in NHPs and mice[132–134].

A wave of new treatments is being developed for inherited retinopathies. In particular, the CRISPR-Cas editing approach is attractive for retinitis pigmentosa (RP) as 25-30% of the cases

are due to autosomal dominant mutations[135]. CRISPR-Cas mediates allele specific editing by recognizing the unique PAM derived from point mutations in dominant disorders such as RP and Meesman epithelial corneal dystrophy[136,137]. However, many disease mutations like *RHO* P23H do not create a novel PAM. One study was able to improve allele discrimination in heterozygous *Rho* P23H mice by placing a single base mutation in the spacer region of the sgRNA coupled with the SpCas9 VRQR variant[138]. Interestingly, another study reported that increasing wild-type to mutant *Rho* expression was able to slow retinal degeneration[139]. This implies that therapeutic benefit can be attained despite incomplete mutant disruption and even with some levels of indiscriminate WT downregulation.

Chronic eye diseases such as age-related macular degeneration (AMD), the leading cause of blindness, have also emerged as a target for gene therapy. Current treatment requires regular anti-VEGF intravitreal injections, but improvements were limited in patients with advanced neovascular/wet age-related macular degeneration (Wet AMD). Previous clinical trials using AAV2-sFLT01, expressing soluble VEGF receptor in Wet AMD patients have reported variable results[140], therefore, other strategies have been developed to reduce angiogenesis. Currently underway clinical trial NCT03748784 uses intravitreal injection of AAV7 m8 expressing aflibercept, anti-VEGF protein. In addition, a recent CRISPR-Cas approach targeting *VEGF-A* using AAV8-spCas9 was able to successfully suppress 31% of laser-induced choroidal neovascularization (CNV) in mice[141]. Overall, the pathogenesis of AMD is complex when compared to monogenic disorders and a multimodal approach enabled by CRISPR technologies might be beneficial.

Successful gene therapy treatment is also dependent on effective administration at the target tissue and the resultant immune response. Subretinal injection has been shown to be efficient

for targeting photoreceptors and RPE while intravitreal (IVT) injections target cells in the inner retina such as ganglion cells[142]. While IVT administration of AAV provides a broader tissue application in the retina, there is an increased likelihood of generating NAbs against the AAV capsid and transgene as shown in NHPs[143]. In addition, T-cell responses against AAV2 capsid were also detected during IVT injection in a NHP toxicology study[144]. Altogether, studies show that IVT administration increases AAV exposure to the immune system[145]. Another consideration is the administration complexity and possible complications that can arise from it. IVT injection is a routine procedure for drug delivery while subretinal injection is more invasive, increasing risk of complications[146]. Therefore, multiple studies have tried targeting RPE using different AAV serotypes via IVT administration. Considerable research has been done to improve AAV transduction of photoreceptors by AAV capsid engineering or to reduce physical barriers by surgical methods such as vitrectomy or by peeling of the internal limiting membrane[147,148].

1.2.4.3 Liver

The liver is essential for many functions, one of which is to metabolize and synthesize intracellular and secreted proteins. Defects in lipid and protein metabolism or detoxification due to liver disease is often destructive to multiple organs causing complex disorders. Liver directed gene therapy has emerged as an attractive option in the treatment of metabolic diseases, demonstrating safety in acute intermittent porphyria (AIP) clinical studies as well as monogenic diseases such as hemophilia B[149,150]. Hepatocyte directed gene transfer is further enabled by easy access of the AAV vector through liver sinusoids as it is sufficiently small to pass through the fenestrae[151].

In the treatment of monogenic diseases such as hemophilia, AAV vectors are preferred for liver directed gene therapy as non-viral and retroviral delivery methods led to only transient

expression[152,153]. The earliest *in vivo* AAV vector used was AAV2 which has broad tissue tropism including hepatocytes. Subsequently, AAV8 was shown to be 10-100 fold more efficient in transducing hepatocytes[154–156]. Furthermore, humans have lower prevalence of NAb against AAV5 and AAV8 as compared to AAV2[157,158]. Hence there was a shift in preference to use AAV8 and AAV5 serotypes for liver directed gene therapies. Liver gene therapy clinical trials have been tested for a small number of diseases with Hemophilia B studies taking the center stage[159].

Ultimately, hemophilia patients need to maintain at least 1% plasma circulation for effective hemostasis in their lifetime. AAV mediated transfer of FIX or FIIIV in liver especially in pediatric patients will result in dilution of the transgene as hepatocytes divide. This problem can be mitigated by site-directed genomic integration. A gene replacement strategy using CRISPR-Cas to integrate the transgene into the albumin locus downstream of the endogenous albumin promoter produced sustained FIX and FIIIV expression in hemophilia mouse models[160,161]. AAV8-Cas9 mediated integration of human FIX-padua exon 2-8 in exon 2 of mFIX and a codon-optimized FIIIV into intron 13 of albumin, developed persistent FIX or FIIIV levels for 7-8 months with no complications. A similar strategy was efficacious for the treatment of ornithine transcarbamylase (OTC) deficiency an X-linked urea cycle disorder, where a codon-optimized human OTC was inserted into the intron 4 of mOTC using AAV8-SaCas9[162]. Conversely, Cas9 mediated HDR to correct the mutation had resulted in large deletions and affected endogenous OTC expression in the spf/ash heterozygous OTC mice[163]. Cas9 integration of OTC demonstrated both rapid and prolonged gene expression when compared with conventional gene replacement therapy[162].

Preclinical studies have shown efficacy in liver directed gene transfer for lysosomal storage diseases (LSD) such as Gaucher disease, Fabry disease, glycogen storage disease type Ib and II[164–

[166]. LSDs are mostly monogenic and include over 40 different metabolic diseases. The spectrum of disease severity also spans from disorders in the central nervous system (CNS) to systemic multi-organ pathology as the enzymes are ubiquitously expressed[167]. Current treatments such as enzyme replacement therapy (ERT) and HSCT have limited efficacy especially for those with CNS disease[168,169]. Prolonged ERT is associated with poorer quality of life, and intravenous ERT has been insufficient to remedy neurological manifestations[170], while HSCT is associated with transplant-related morbidity and mortality rates of 10%[171–173].

Interestingly, recent studies in LSDs were able to affect some neurological benefit using liver directed gene editing. The effects in CNS were likely due to enzymes crossing the BBB rather than gene editing in the brain as editing was driven by a liver specific promoter. This was demonstrated in AAV-SaCas9 mediated transgene integration into the albumin safe harbor locus in Sandhoff mice[174]. In addition, Sangamo Therapeutics, in a promising preclinical study, used an AAV2/8 ZFN editing strategy to insert α-L-iduronidase (IDUA) into the albumin locus and resulted in cognitive improvements in mice with Hurler syndrome, the most severe form of MPS I[175]. *In-vivo* correction of autosomal recessive Mucopolysacchraidosis type I (MPS I) has been explored by using CRISPR-Cas to insert the *Idua* gene in a murine safe harbor locus via HDR[176]. IDUA activity was detected in multiple tissues, including heart, lung, liver, kidney, and spleen with the use of cationic liposomes, but could not cross the blood-brain barrier (BBB). However, the BBB can be crossed with the use of specific AAV serotypes[177].

1.2.4.4 Central nervous system (CNS)

Neurological disorders are difficult to treat due to the complexity of the central nervous system (CNS). While conventional therapy has limited access to the brain, specific AAV serotypes are able to cross the BBB due to their unique tissue tropism[177]. Non-invasive mode of delivery by

intravenous injection of AAV9 is able to transduce the spinal cord including motor neurons and astrocytes in mice, large animals and NHP[178,179]. Intra-cerebrospinal fluid (CSF) delivery is also able to provide widespread distribution in the CNS and reduce transduction of periphery organs[180,181]. The third approach is to directly deliver to target regions, limiting therapy to the site of administration hence broader delivery will require multiple injections and surgery[182].

Zolgensma the approved CNS targeting gene therapy for spinal muscular atrophy (SMA) delivers scAAV9-SMN1 systemically to target motor neurons extending from the brain, spinal cord to muscles[4]. However, the high vector dose resulted in liver toxicity and increased cost of therapy. Alternatively, the payload can be reduced with intra-CSF administration by intracerebroventricular and intrathecal injections. Gene transfer in a severe SMA mouse model was able to reduce the dose by 10 times with intra-CSF administration to improve survival[183]. Safe and robust widespread transduction in the motor neurons of the spinal cord was also demonstrated by intrathecal injection of AAV9 in pigs[184]. As a result, intrathecal administration of scAAV9-SMN1 is currently being evaluated in clinical trials in SMA type 2 patients (NCT03381729).

Amyotrophic lateral sclerosis (ALS), similar to SMA, is characterized by motor neuron loss in multiple regions in the CNS[185,186]. Dominant SOD1 mutations are most common in familial ALS and CRISPR-Cas editing of mutant SOD1 using AAV9-SaCas9 improved locomotor functions and survival when administered intravenously or via intracerebroventricular injection[187,188]. More recently, CSF administered Cas9-cytidine base editor was able disrupt mutant SOD1 in adult ALS transgenic mice and improve survival[189].

Localized administration of CRISPR-Cas therapies in the brain is also essential for neurodegenerative diseases with neuronal loss in specific areas of the brain such as the striatum in HD, limiting potential off-target effects in other tissues. Intracranial delivery of AAV1-SaCas9

targeting the 5' end of exon 1 reduced mutant *HTT* expression in HD mouse model was able to improve motor function and benefit survival[190]. CRISPR-Cas gene editing therapy was also demonstrated in common progressive neurodegenerative disorders such AD. Intracranial delivery of Cas9 nanocomplex to disrupt beta-secretase 1 *Bace1* in AD mouse models reduced amyloid plaque accumulation and cognitive deficits[7]. Dominant familial AD caused by mutations in amyloid precursor protein (APP) the site of beta-secretase 1 cleavage, results in accumulation of amyloid plaques. AAV9-Cas9 treatment via hippocampus injection specifically disrupted mutant APP gene, reducing amyloid-β levels[191]. However, a possible limitation of the current treatment strategy is the lack of widespread targeting as neurons in multiple area of the brains are progressively affected in AD. Cas expression in the CNS can be driven by neuron specific promoters while leveraging AAV tropism to increase target cell specificity[191].

1.2.4.5 Animal Models

Several mouse models harboring particular disease-causing mutations have been generated for the study of CRISPR-Cas mediated gene editing (**Table 1.0**). Knock-out mouse models have been essential in the study of inherited genetic diseases including those affecting the CNS[192,193]. Similarly, transgenic models bearing sporadic mutations have been generated for AD and Parkinson's disease (PD)[194]. Mouse models enable testing of the editing strategy, but often do not fully recapitulate the extent of the disease. *Mdx* mice commonly used for example, have a much slower disease progression, shortened life span of 25% as compared to 75% in DMD patients, and cardiomyopathy is only present in older mice[195]. Likewise, in AD mice while the main pathology of the disease Aβ aggregation is detected, overt neurodegeneration is not captured nor the early symptoms which are often cognitive and behavioral[196]. In addition, it is harder to extrapolate

results of widespread AAV delivery in mice organs such as the brain to humans due to difference in distribution[180] and possibly the large disparity in sizes.

Overall, advanced human diseases are better recapitulated in large animal models. DMD pig model, for example, embodies severe disease progression such as premature death and cardiomyopathy[197]. Large transgenic animal models of HD also display severe phenotypes and early death when compared to rodent models[198]. Larger species such as canines have a more similar brain disease profile to humans and develop naturally occurring LSDs[199,200]. In addition, large species are more suitable for surgical manipulations especially for neurodegenerative diseases[201].

Complexity can arise in the choice of animal species when progressing from small to large animal studies. For instance, vector tropism can differ between animal models as shown in the successful transduction of AAV5 in the mouse brain but not in cat[202]. In addition, while AAV-PHP.B has an enhanced ability to cross the BBB and increase CNS transduction in mice, this was not replicated in NHPs[203]. Extending preclinical studies to larger animals has been essential in better understanding immune responses towards gene therapy. Canine models have been essential in uncovering possible severe immune response against muscle-directed transgene[204] and cellular response towards AAV2 and AAV6 capsid[205]. Importantly, large animal studies can also inform immunosuppressive regimen for better therapy outcomes[206]. Interestingly naturally AAV-infected rhesus macaques were unable to induce similar T cell responses to AAV capsids upon re-exposure during treatment regimen despite the close phylogenetic relationship to humans[207,208].

Table 1.1: In vivo CRISPR-Cas strategies developed for functional rescue in animal models.

	Human Disease	**Editing strategy**	**Disease model**	**Administration**	**Ref**
Eye	Retinitis Pigmentosa, autosomal recessive, *PDE6B* mutation	SpCas9/ RecA mediated HDR	*Pde6b* Y347X, nonsense mutation mice	SpCas9, sgRNA, RecA-MS2 Plasmid and ssDNA donor, subretinal & electroporation	[209]

Table 1.1: In vivo CRISPR-Cas strategies developed for functional rescue in animal models… (Continued)

	Human Disease	Editing strategy	Disease model	Administration	Ref
Eye	Retinitis Pigmentosa, autosomal dominant • P23H mutation (class 2) • Class 1	Knockdown Rho	*Rho*-P23H transgenic mice	Cas9-sgRNA plasmids, subretinal & electroporation	210
		Knockdown mutant *Rho*	*Rho*-P23H heterozygous mice	SpCas9-VRQR plasmid & sgRNA plasmid, subretinal & electroporation	138
			Rho S334ter rat	Cas9-sgRNA plasmid, subretinal & electroporation	136
		Knockdown of transcription factor Neural Leucine zipper (Nrl)	Rho-/-, Rd10, RHO-P347S mice	AAV8-RK-spCas9, AAV-sgRNA, subretinal	211
	Leber Congenital amaurosis, autosomal recessive • Type 10, IVS26 mutation in *CEP20* • Type 2, *RPE65* mutation	NHEJ in-frame indels or HR	Humanised CEP290 IVS26 knock-in mice, NHP	AAV5-sgRNA-GRK1-saCas9, subretinal	212
			Rpe65 nonsense mutation rd12 mice	AAV9-EFS-spCas9, AAV9-sgRNA-*Rpe65*-donor, subretinal	213
	Meesman epithelial corneal dystrophy, dominant negative, *KRT12* mutation	Knockdown, allele specific, Deplete mutant *KRT12*	*KRT12*-L132P humanized heterozygous mice	SpCas9 & sgRNA plasmids, Intrastromal ocular	137
	Neovascular age-related macular degeneration	Knockdown, NHEJ mediated disruption of *Vegfa or Hif1a*	Laser-induced choroidal neovascularization in mouse eyes	AAV8-CMV-spCas9, AAV8-sgRNA, subretinal	141
				AAV9-EFS-LbCpf1-crRNA, intravitreal	214
Neurological	Huntington's disease, autosomal dominant, polyQ (CAG trinucelotide) expansion mutant *HTT*	Knockdown *HTT*	HD140Q-knock-in mice, human *HTT* exon 1 with 140 CAG repeats (homozygous & heterozygous)	AAV-MECP2-Cas9, AAV-gRNA, intracranial	215
		Knockdown mutant *HTT*	R6/2 mice, transgenic mice, human *HTT*	AAV1hSyn--SaCas9-sgRNA, intracranial	190

Table 1.1: In vivo CRISPR-Cas strategies developed for functional rescue in animal models… (Continued)

	Human Disease	Editing strategy	Disease model	Administration	Ref
			exon 1 with 115-150 CAG repeats		
	Amyotrophic lateral sclerosis, autosomal dominant, *SOD1* mutations	Knockdown human *SOD1*	Transgenic human *SOD1* G93A mutation ALS neonatal mice	AAV9-CMV-SaCas9-sgRNA, intravenous or intracerebroventricular	187,188
		BE3 disruption of *SOD1*		AAV9-CAG-N-Int-CBE-U6-sgRNA, AAV9-CAG-C-Int-CBE-U6-sgRNA, intrathecal	189
	Fragile X syndrome	Knockdown m*GluR5*	*Fmr1* KO mice, FXS mice model	CRISPR Gold nanoparticle Cas9 RNP & donor DNA, intracranial	77
	Alzheimer's disease	Knockdown *Bace1*	Five familial Alzheimer's disease transgenic mice, *APP* knock-in Alzheimer's disease mice	Amphiphilic R7L10 peptide nanocomplex Cas9 RNP, intracranial	7
	Dravet Syndrome, haploinsufficiency, *SCN1A* loss of function mutation	dCas9-Vp64 mediated *Scn1a* gene activation (both WT & mutant)	*Scn1a* heterozygous Dravet Syndrome mice	AAV9-dCas9-VP64, AAV9-sgRNA, intracerebroventricular	216
	Severe obesity, haploinsufficiency, *SIM1* or *MC4R*	dCas9-Vp64 mediated gene activation of *Sim1*	*Sim1/ Mc4r* heterozygous mice	AAVDJ-CMV-dCas9-Vp64, AAVDJ-sgRNA, hypothalamic	217
Liver	Hemophilia B, X-linked recessive	HDR, insert hFIX-padua exon2-8 in mFIX exon 2	FIX Knockout mice	AAV8-Cas9, AAV8-sgRNA-Donor DNA	160
		HDR, insert into mFIX into murine ROSA26 safe harbor	R333Q Hemophilia mice	Ad5-Cas9-gRNA, Ad5-EF1α-mFIX, intravenous	218

Table 1.1: In vivo CRISPR-Cas strategies developed for functional rescue in animal models… (Continued)

	Human Disease	Editing strategy	Disease model	Administration	Ref
	Hemophilia B, Y371D mutation in FIX	HDR	Y381D Hemophilia B mice	Naked DNA, spCas9, ssODN-HDR Donor, intravenous	219
	Hemophilia A, X-linked recessive	Insert human B-domain deleted FVIII in intron 13 of liver-specific albumin locus	FVIII Knockout mice	AAV8-saCas9-sgRNA, AAV8-donor DNA	161
	Ornithine transcarbamylase deficiency, X-linked recessive	HDR, correct point mutation	spf/ash heterozygous OTC neonate mice	AAV8-TBG-saCas9, AAV8-sgRNA-donor DNA, intravenous	163
		HDR, insert codon-optimized human OTC into intron 4 of mOTC locus		AAV8-TBG-saCas9, AAV8-sgRNA-TBG-hOTCco, intravenous injection	162
	Hypercholesteolemia	BE3, disruption of mouse W159 *Pcsk9*	WT C57BL/6 mice	Ad-BE3-sgRNA, retro-orbital	220
		BE3, disruption of human W159 *PCSK10* or NHEJ-mediated disruption of *PCSK10*	Humanized *PCSK9 knock-in* mice	Ad5-Cas9-sgRNA or Ad5-BE3-sgRNA, Intravenous	221
	Atherogenic dyslipidemia, homozygous familial hypercholesterolemia	BE3, disruption of Q135 *Angptl3*	hyperlipidemic *Ldlr*-knockout mice	Ad-BE3-sgRNA	220
	Phenylketonuria, autosomal recessive	BE3	Homozygous point mutation in *Pah*-F263S, hyperphenylalaninemia mice	AAV8.N-int-BE3 & AAV8.C-int-BE3, intravenous	222

Table 1.1: In vivo CRISPR-Cas strategies developed for functional rescue in animal models… (Continued)

	Human Disease	Editing strategy	Disease model	Administration	Ref
	Hereditary tyrosinemia type I, autosomal recessive, *FAH* loss of function mutation	HDR	*Fah* homozygous mutation in exon 8 mice	spCas9 mRNA by LNP & AAV8-sgRNA-donor template, intravenous	73
		HDR with nCas9	Homozygous 10-bp deletion in exon 2 of Fah in rat HTI model	Ad-nCas9 & Adv-donor template, intravenous	223
		Base editing, ABE6.3, tRNA nCas9-adenosine deaminase	*Fah* homozygous mutation in exon 8 mice	Plasmid by hydrodynamic injection or mRNA/sgRNA by LNP, Intravenous	224
	Hereditary tyrosinemia type I, autosomal recessive, compound heterozygous mutations in *FAH*	Allelic exchange by targeting intron 7, NHEJ	compound heterozygous exon 5 insertion/ exon 8 mutation in *Fah*, HTI moue model	AAV9-SpCas9 & ScAAV8-sgRNA, intravenous	225
Muscle	Duchenne muscular dystrophy, mutation hotspot in exon 50	Reframing or Exon skipping	ΔEx50 C57/BL6J mice	AAV9-CK8-spCas9 & AAV9-sgRNA-51, Intraperitoneal	226
			ΔEx50 dogs	AAV9-CK8-spCas9 & AAV9-sgRNA-51, Intravenous	120
	Duchenne muscular dystrophy, mutation hotspot in exon 44	Reframing or Exon skipping	ΔEx44 C57BL/6 mice	AAV9-CK8-spCas9 & AAV9-sgRNA, Intraperitoneal	227
				AAV9-CK8-spCas9 & scAAV-sgRNA	121
	Duchenne muscular dystrophy, mutation hotspot region (exon 45-55)	Excise exon 52 & 53	mdx^{4cw} mice, nonsense mutation in exon 53	AAV6-CK8-spCas9 & AAV6-sgRNA or AAV6-CK8-saCas9-sgRNA, retro-orbital	115
	Duchenne muscular dystrophy, nonsense mutation in exon 23	Excise exon 23	Mdx mice, pt mutation in exon 23	AAV9-CMV-saCas9 & AAV9-sgRNA	114
				AAV8-CMV-saCas9 & AAV8-sgRNA, Intravenous	119,228

Table 1.1: In vivo CRISPR-Cas strategies developed for functional rescue in animal models… (Continued)

	Human Disease	Editing strategy	Disease model	Administration	Ref
				AAV9-CMV-saCas9 & AAV9-sgRNA, Intravenous	229
				AAV9-miniCMV-spCas9 & AAV9-sgRNA, Intraperitoneal	230
		Excise exon 21-23		Ad-Cas9-sgRNA	231
		HDR		CRISPR Gold nanoparticle Cas9 RNP & donor DNA, intramuscular	75
		NHEJ, indels to result in reframing	DMD KO C57/BL6J mice, ΔEx23	AAV9-Spc5-12-CjCas9-sgRNA, Intramuscular	232
	Wolff-Parkinson-White syndrome, autosomal dominant inherited disease, H530R mutation in *PRKAG2*	Disrupt mutant allele	Heart specific H530R transgenic mice, Heterogenous H530R PRKAG2 knock-in mice	AAV9-Cas9, scAAV9-sgRNA, intravenous or intraventricular	8
Ear	Dominant progressive hearing loss (DFN36), *TMC1*	Knockdown mutant *Tmc1*	Beethoven mice, heterozygous missense mutation in *Tmc1*	Cas9-sgRNA RNP, cationic lipid complex, inner ear	233
				Anc80L65-CMV-SaCas9-sgRNA, inner ear	234
Multiorgan	Mucopolysaccharidosis type I, autosomal recessive	HDR, insert *Idua* into murine *ROSA26* safe harbor	Idua-knock-out MPS I mice	PrecisionX CRISPR/Cas9 SmartNuclease ™, IDUA donor vector, Cationic liposomes, Intravenous	176

1.2.5 Challenges

1.2.5.1 Immunogenicity of the payload and vehicle

One of the most critical barriers to successful therapeutic translation is immune interaction. Most significantly, there are concerns that the adaptive immune response could inhibit the efficacy

of CRISPR-based gene therapies. This could occur through the action of neutralizing antibodies against either the Cas effector protein or the delivery vehicle, or through cytotoxic T-cell responses leading to specific killing of treated cells. Antibodies towards AAV-CRISPR-Cas9 have been shown to neutralize and negate editing efficacy in mice with just one prior exposure to the therapeutics[235]. This is particularly pertinent for clinical application because multiple studies have now identified pre-existing immunity in the human population towards the commonly used Cas9 orthologs from *S. pyogenes* and *S. aureus*[236–239], understandably since both are species that frequently colonize the human microbiome and may cause disease. Identification of circulating antibodies reacting to Cas9 could pose a large hurdle towards RNPs, since these purified Cas9 protein complexes could be recognized and neutralized within the bloodstream. Other Cas-encoding DNA and RNA modalities might generally be less inhibited by this obstacle, since the Cas9 protein will be expressed intracellularly and remain less accessible to antibodies from the circulation.

A more substantial risk in this case, however, is the potential for cytotoxic T-lymphocytes to recognize Cas-expressing cells through the Cas-derived peptides presented at the cell surface via the MHC class I pathway. Given the appropriate costimulatory milieu generated by innate immune signaling in the CTL, the target cell, and the surrounding tissue, this can lead to killing of target cells expressing the Cas proteins. Two *in vivo* mouse studies have highlighted the potential for anti-Cas9 immunity to inhibit efficacy of CRISPR-based drugs. In one, a single dose of AAV-delivered Cas9 was shown to be sufficient to generate anti-Cas9 T-cells in mice and to induce their killing potential[235]. Another study showed that pre-stimulated anti-Cas9 cytotoxic T-cells completely negated the therapeutic effect of Cas9 gene editing within 12 weeks[240]. Although no human studies have shown an inhibitory effect of the immune system as of yet, Cas9-reactive T-

cells have been shown by multiple groups to exist in significant fractions of the population[236,238]. The interaction of the immune system with CRIPSR-based therapies has only just begun to be investigated, but preliminary data suggests it could become an obstacle to *in vivo* CRISPR-Cas therapies and may signal that transient immunosuppression or transient transgene expression or Cas engineering would be necessary to maintain patient safety and efficacy[241].

1.2.5.2 Readout challenges

Designing animal studies to provide an early readout for efficacy and potency is challenging, as shown in the preclinical study of EDIT 101. Editas Medicine used both a disease mouse model and NHP to measure desired physiological response. Well-established tissues in the eye allow for specific isolation, such as the cultivation of the neural retina of adult cadavers into retinal explant cultures for demonstration of editing efficacy in mature human photoreceptors [219]. Neural retinas from the knock-in mouse model of LCA10 were isolated to assess dosing to achieve functional rescue for clinical outcome. However, as photoreceptors in mice are predominantly rods, quantitative assessment of the editing rates in foveal cones were validated in a primate model. Surrogate guide RNAs had to be designed for NHPs due its genetic divergence between humans. Total productive editing of the therapy was subsequently extrapolated from the editing frequency of the tissue sample. This extrapolation and in some instances relying on phenotypic versus genotypic outcomes becomes particularly critical as obtaining human tissues (e.g., retina) may not always be feasible. On the other hand, diseases entailing systemic correction such as DMD as opposed to localized editing in the eye will require multi-site sampling to quantify total editing. One of the drawbacks when using surrogate guides would be how representative the editing results will be in determining the correct dose for therapeutic efficacy. Greater confidence can be drawn by comparing with an additional model as shown in this study. For instance, EDIT 101 in

humanized mouse model when compared with surrogate guides in NHP showed similar editing rates in response to increasing dose.

In contrast, the manufacturing of autologous products will vary due to the material received from the donor, but the final product can be well-characterized before administration. In the study by Stadtmauer *et al*., unique CRISPR-Cas9 engineered T-cells were quantified for the frequency of edits using a PCR based assay and assessed for cytotoxic effects against tumor cells with an *in vitro* potency assay [5]. Unlike *in vivo* targeting, the editing efficiency required to effect clinical benefit was not established for the targets. Whole blood samples and biopsy of the bone marrow and tumor site in patients provided a direct post infusion readout of the therapeutic product.

An additional readout challenge is the careful characterization of the off-target genomic and transcriptomic effects of CRISPR-based therapies. In the last few years, several sensitive methodologies for assaying off-target effects have been developed.[251-257] Even so, the assaying of clinically-relevant off-target effects remains difficult because depending on the genomic context and cell type, some off-targets may be completely benign, while others could have serious consequences. This is a well-recognized issue in the field and many efforts to manage this are underway, particularly engineering of the CRISPR payload at both the protein and gRNA level and management of the window of active exposure to the functional RNP complex. For more in-depth reviews on this aspect of CRISPR-based therapies see.[258-261]

1.2.5.3 Regulatory concerns and n-of-1 trials

CRISPR-based therapeutics face a host of unique regulatory challenges when compared with other therapeutic methods such as small-molecule drugs or antibodies that typically act at the protein level. There are two layers of information abstraction when moving from DNA to RNA to protein, and often this information may have important implications for therapeutic development.

One such example is the fact that even a single genetic disease caused by knockout of a singular gene may have multiple underlying mutations in different patients. Often a single mutation, such as the delta508F for cystic fibrosis, the deletion of chromosomal region 15q11-13 on the maternal allele in Angelman syndrome, and the 1278insTATC mutation in the *HEXA* gene in Tay-Sachs disease, accounts for a majority of cases, but in each of these diseases, tens to hundreds of other mutations affecting the same genes have also been shown to cause the disease. Correction of these mutations by CRISPR-Cas genome editing would likely require different gRNA constructs, or even different Cas effector modalities, even with the same delivery mechanism. These different gRNAs would naturally have different safety and efficacy profiles as a result of the varying functional activities, off-target effects, etc. of distinct gRNAs. Under currently standard regulatory frameworks, this would seem to require many individual therapies based on the same platform to treat the same disease to undergo their own arduous and expensive approval studies.

Oligonucleotide-based drugs, particularly antisense oligos (ASOs), though ripe with therapeutic potential, face a similar regulatory problem. To address this, companies such as Ionis Therapeutics have pioneered n-of-1 trials with personalized ASOs for patient specific mutations causing a variety of diseases including ALS, Batten disease, and cystic fibrosis[242–244]. Despite the success of several of these trials, it is an arduous process for which the FDA is not streamlined. These trials to date have relied on right-to-try clauses, expanded access INDs, and perhaps most critically, heroic fundraising efforts on the part of the patients. Despite personalization being a conspicuous goal of therapeutic science for years, there remains no clear regulatory path for these kinds of personalized medicines to be tested at scale. However, due to the trailblazing of ASO therapies and the promise of CRISPR-based drugs, the FDA and other global regulatory agencies are acutely aware of this problem and may well deploy new guidelines to mitigate it[245,246].

Clearly, safety and efficacy must be demonstrated, even for n-of-1 trials and personalized therapies[247]. However, the process by which that can be accomplished must be simplified and streamlined for drugs that have an established platform. Many aspects of safety testing can be ported between similar drugs which may use the same delivery and mechanism, but with different underlying nucleic acid targets. Additionally, inexpensive, fast, and robust efficacy testing platforms must be developed, which fortunately is an area of significant research interest. Although the landscape of current drug development models is not designed for the dynamic, personalized potential of the future of therapeutic breakthroughs including CRISPR-based drugs, deft regulatory work and new guidelines can address this problem and pave the way to unlock that potential.

1.3 LOOKING FORWARD

In this work, I primarily focus on the immunity issue, as it is both of critical importance and potentially solvable through technological innovations. To this end I take an approach informed by previous de-immunization efforts, evolutionary and structural attributes of the proteins involved, technical advancements such as long-read single molecule sequencing, and merging of computational and experimental design and screening methodologies to enable meaningful strides towards safer and more efficacious therapies with regard to immune interactions. Beginning with a characterization of exactly what immune barriers are faced by AAV-delivered CRISPR-Cas, I then branch out into possible solutions to these issues both by leveraging existing immune orthogonal orthologs of the therapeutic proteins, and by engineering new variants of these proteins with superior immunological attributes. Finally, I discuss how this work fits into the broader field of CRISPR-Cas gene therapy, and how the advancements developed here can be applied in other contexts to facilitate both scientific discovery and the resultant clinical applications.

1.4 ACKNOWLEDGEMENTS

Chapter 1, in part, is a reprint of the material as it appears in The CRISPR Journal 3(4): 253–275 (2020). Tay, Lavina Sierra; Palmer, Nathan; Panwala, Rebecca; Chew, Wei Leong; Mali, Prashant. The dissertation author was the primary researcher and author of this paper. This work was generously supported by NIH grants (R01HG009285, RO1CA222826, RO1GM123313) and by the Agency for Science, Technology, and Research under its Industrial Alignment Fund (Pre-Positioning) (H17/01/a0/012).

CHAPTER 2 – THE ADAPTIVE IMMUNE RESPONSE TO AAV-CRISPR-CAS

Protein therapeutics, including protein-based gene therapy, have several advantages over small-molecule drugs. They generally serve complex, specific functions, and have minimal off-target interference with normal biological processes. However, one of the fundamental challenges to any protein-based therapeutic is the interaction with the adaptive immune system. Neutralization by circulating antibodies through B-cell activation and clearance of treated cells by CD8+ cytotoxic T-lymphocytes (CTLs) create a substantial barrier to effective protein therapies[248–251]. Although for some applications the delay in the adaptive immune response to novel proteins may allow sufficient time for the initial dose to work, subsequent doses face faster and stronger secondary immune responses due to the presence of memory T- and B- cells. In addition, gene transfer studies have shown that host immune responses against the delivery vector and/or therapeutic transgene can eliminate treated cells, thus limiting the efficacy of the treatment[252–258].

Furthermore, protein therapies often require repeated treatments due to degradation of the protein, turnover of treated cells, or, in the case of gene therapy, reduced expression of the transgene[259,260]. This provides an even greater challenge as repeated exposure to the same antigen can elicit a more robust secondary immune response[261], which may completely inhibit subsequent dosage or even sensitize the immune system to antigens remaining from the initial exposure.

An additional important consideration is the immunogenicity of the delivery vehicle or administration route for the Cas9 and associated guide RNA (gRNA). In this regard, adeno-associated viruses (AAVs) have emerged as a highly preferred vehicle for gene delivery, as they are associated with low immunogenicity and toxicity[255,256], which promotes transgene expression[262,263] and treatment efficacy. Despite the relatively low immunogenicity of AAV vectors, antibodies against both the capsid and transgene may still be elicited[257,264–270].

Additionally, the prevalence of neutralizing antibodies (NAB) against AAVs in the human population[271] and cross-reactivity between serotypes[272] remains a hurdle for efficacious AAV therapy. Although AAVs were initially considered non-immunogenic due to their poor transduction of antigen-presenting cells (APCs)[273], it is now known that they can transduce dendritic cells (DCs)[274] and trigger innate immune responses through Toll-like receptor (TLR) signaling pathways[275]. The ability to transduce DCs is dependent on AAV serotype and genome, and may be predictive of overall immunogenicity[276]. A previous study exploring the utility of the AAV-Cas9 system observed a humoral immune response to both AAV and Cas9, as well as an expansion of myeloid and T-cells in response to Cas9[270], highlighting the need to confront this issue when further developing gene therapies.

To assess the degree to which the immune response poses a barrier to AAV-delivered CRISPR-Cas, and to determine what features of the immune response constitute the critical levers governing the success of the therapy, we characterized an experimental model in which we treat mice with AAV-delivered CRISPR-Cas targeting the PCKS9 gene, primarily expressed in hepatocytes. Inhibition of this gene is known to be protective against atherosclerotic disease via the reduction of circulating low density lipoprotein levels. Additionally, the therapeutic effect is easily measured on a continuous basis through the measurement of PCSK9 in sera. In addition to demonstrating efficacy of AAV-delivered CRISPR-Cas in knocking out PCSK9, measured at both the protein and DNA level, we assayed treated mice for induction of protein-targeting antibodies and of T-cell-mediated cytotoxicity against AAV and Cas9 proteins.

2.1 RESULTS

2.1.1 Immune response to AAV and Cas9

One of the major obstacles for sequential gene therapy treatments is the presence of neutralizing antibodies against the delivery vehicle and transgene cargo induced by the first administration of the therapy **(Figure 2.1a)**. To determine the humoral immune response kinetics to AAV-CRISPR therapeutics **(Figure 2.1b)**, focusing as an exemplar on the AAV8 capsid and the Cas9 transgene, we first injected C57BL/6J mice retro-orbitally with 10^{12} vg of AAV8-SaCas9 targeting proprotein convertase subtilisin/kexin type 9 (PCSK9), a promising gene target that when disrupted can reduce LDL levels and protect against cardiovascular disease. Consistent with a previous study[277], mice had reduced PCSK9 serum levels as early as one week post-injection due to successful SaCas9 mediated gene-editing, which was sustained for the entire duration of the experiment (4 weeks) **(Figure 2.1c).** We noted that a subset of the mice developed IgG1 antibodies against the SaCas9 protein **(Figure 2.1d)**. Additionally, mice developed humoral immunity to the AAV8 capsid within one-week post-injection **(Figure 2.1e)**.

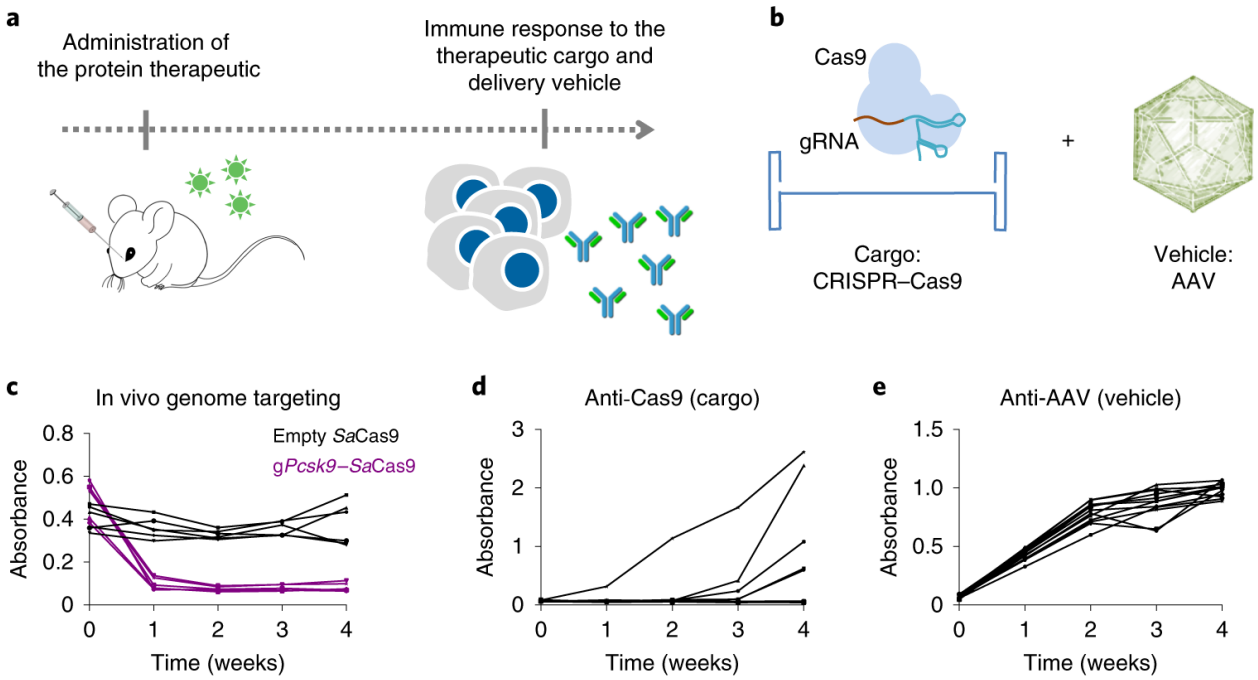


Figure 2.1: Treatment efficacy and humoral response. (a) Proteins have substantial therapeutic potential, but a major drawback is the immune response to both the therapeutic protein and its delivery vehicle. (b) As a case study, we studied the CRISPR–Cas9 systems and corresponding delivery vehicles based on AAVs. (c) Mice were injected retro-orbitally with 1×10^{12} VG per mouse of AAV8-SaCas9 targeting the Pcsk9 gene or a non-targeting control (empty vector). A decrease in PCSK9 serum levels—owing to successful gene targeting—was observed in mice that received the Pcsk9-targeting AAV-SaCas9 virus (n = 6 mice for each group). Each line represents an individual mouse. (d) Immune responses of mice treated with AAV8-SaCas9 to the cargo (Cas9) were detected using ELISAs for the SaCas9 protein (n = 12 mice). Each line represents an individual mouse. (e) Immune responses of mice treated with AAV8-SaCas9 to the delivery vehicle (AAV) were detected using ELISAs for the AAV8 virus capsid (n = 12 mice). Each line represents an individual mouse.

In addition to the antibody-mediated response, we also observed cell-mediated responses to AAV8 and SaCas9 in treated mice using an IFN-γ ELISPOT assay. In this assay, antigenic components are presented to T-cells isolated from treated mice by activated antigen-presenting cells. T-cells that can recognize the antigen will receive signals to proliferate and to secrete cytokines such as IFN-γ, which can be measured using an ELISA-style colorimetric assay, such that an activated T-cell will create a visual spot on the plate. The number of spots roughly corresponds to the number of T-cells able to respond to the antigen being tested. We tested a set

of MHCII-restricted epitopes predicted to be strongly immunogenic from both AAV8 and SaCas9. A subset of these epitopes showed significant CD4+ T-cell responses indicating that cell-mediated immunity against both the delivery vehicle and transgene payload is induced upon receipt of a gene therapy dose. **(Figure 2.2)**.

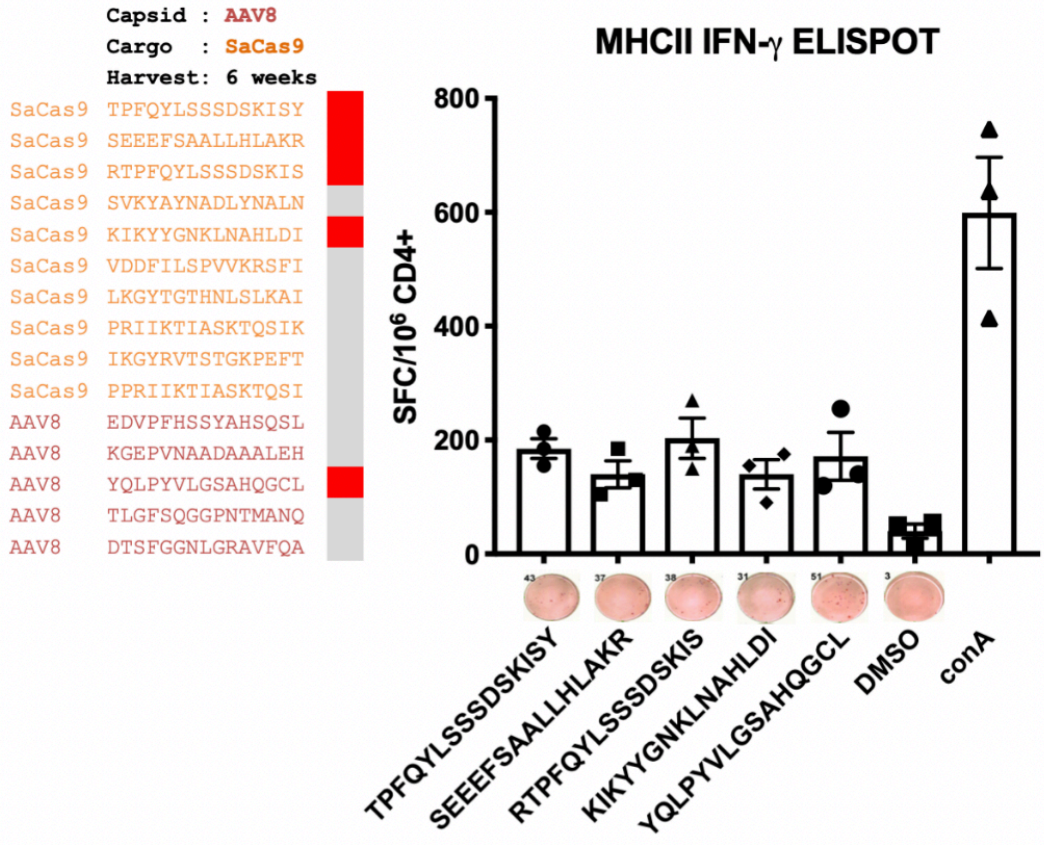


Figure 2.2: Experimental validation of a MHCII peptide predictions via IFN-γ ELISPOT. Mice were injected retro-orbitally with $1x10^{12}$ vg/mouse of AAV8-SaCas9 targeting the Pcsk9 gene and were sacrificed after 6 weeks. Purified CD4+ T-cells from splenocytes were seeded at $2x10^5$ cells per well (results are technical triplicates of n=6 pooled mice and are shown as mean ± s.e.m.). $1x10^5$ lipopolysaccharide-activated antigen presenting cells (APCs) from control mice were added to each well. Cells were incubated with highly immunogenic MHC-II predicted peptides for 20h. Spots were developed with biotinylated anti-IFN-γ. A one-way ANOVA with post hoc Dunnett's test was performed to determine statistical differences with DMSO for all peptides (left panel) (TPFQYLSSSDSKISY, p=0.004; SEEEFSAALLHLAKR, p=0.0339; RTPFQYLSSSDSKIS, p=0.0001; KIKYYGNKLNAHLDI, p=0.0339; YQLPYVLGSAHQGCL, p=0.0015). Data for the significant peptides are plotted (right panel).

Although these experiments demonstrated that AAV-CRISPR-Cas specific T-cells are induced to recognize antigenic components of the therapy, we set out to measure what the physiological effects those T-cells may have on cells successfully transduced by AAV and expressing the Cas9 payload. This was especially important to address given that we were not able to detect a strong CD8+ T-cell response to our predicted MHCI-restricted epitopes via IFN-γ ELISPOT, and it represents the most appropriate experiment to assess what effect cell-mediated immunity may have on therapeutic efficacy. To this end, we immunized mice against Cas9 by injecting whole Cas9 protein in combination with Complete Freund's Adjuvant (CFA). After immunization we challenged these mice by injecting fluorescently labeled splenocytes which had been pulsed with Cas9 epitopes. We observed 39% specific clearance of Cas9-pulsed cells 3.5 weeks after immunization, but no clearance 1 week after immunization, demonstrating that anti-Cas9 T-cells can specifically recognize and kill cells presenting Cas9 epitopes *in vivo* (**Figure 2.3**).

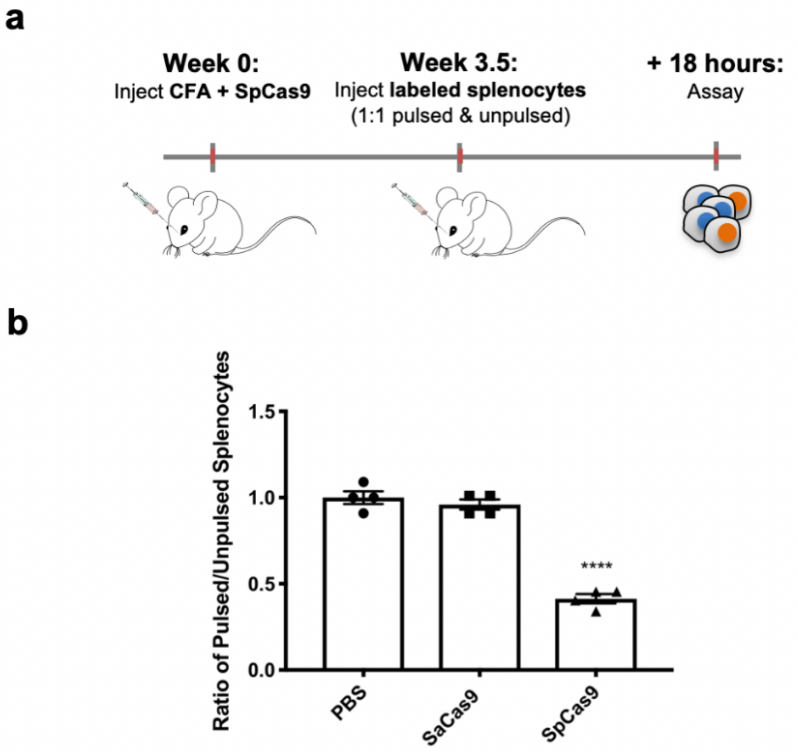


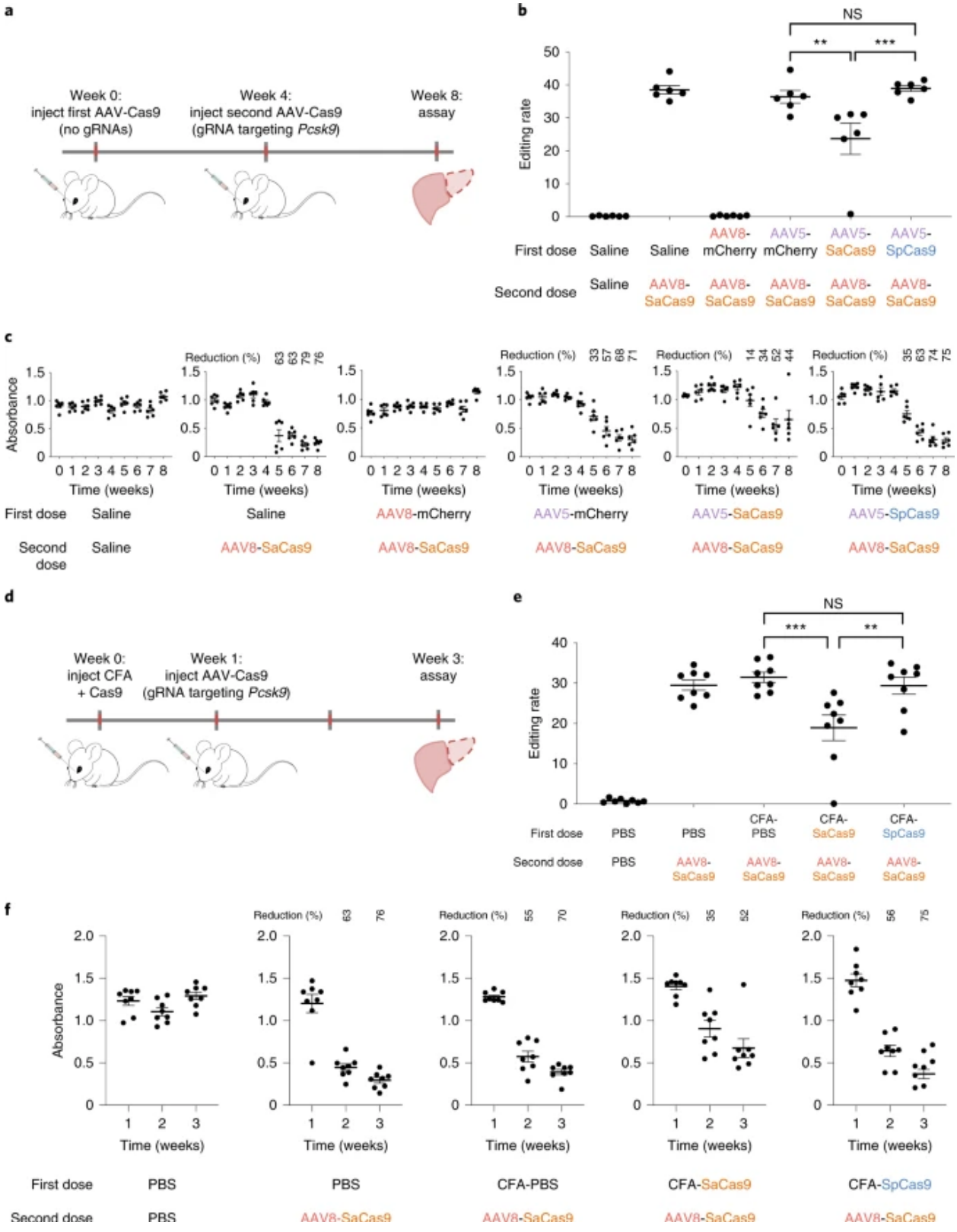
Figure 2.3: Cas9-specific splenocyte clearance in vivo. (a) Mice were immunized with CFA + SpCas9 protein. Splenocytes from naive mice were prepared by labeling with the fluorescent dye CTV or CFSE and pulsing with either a pool of immunogenic Cas9 epitopes (SpCas9 or SaCas9) or DMSO. A 1:1 mixture of these cells was injected retro-orbitally into immunized mice at a total of 6 x 10 7 cells per mouse. After 18 hours, splenocytes from these mice were analyzed by flow cytometry to assay for specific clearance of Cas9 epitope pulsed cells. (b) Ratio of pulsed/unpulsed splenocyte clearance at 3.5 weeks post-immunization. SpCas9-pulsed splenocytes are specifically cleared. Results are shown as mean ± s.e.m. Each data point represents an individual mouse with n=4 mice per group. A one-way ANOVA with post hoc Dunnett's test was performed, $p<0.0001$.

2.1.2 Immune barriers to effective gene editing

Having demonstrated that AAV-Cas9s elicit a significant immune response in the mouse model, we next performed a two-step dosing experiment to test whether these immune responses inhibit the efficacy of multi-dose gene editing. For this experiment, we used another mouse strain, BALB/c, in order to verify and characterize the immune response in two independent strains. The

first round of dosing contained no gRNA and served to immunize the mice against the second dose, which contained an active AAV8-SaCas9 with gRNA targeting PCSK9, allowing us to directly measure genome editing efficiency by sequencing, as well as serum PCSK9 levels as a phenotypic readout for therapeutic efficacy **(Figure 2.4a)**. Additionally, we measured IgG responses to all AAV and Cas9 used in the experiment. As expected from previous preclinical work on AAV therapies, prior exposure and humoral immunity to AAV8 (AAV8-mCherry) abolished the effectiveness of subsequent gene editing when using AAV8 as the delivery vector (AAV8-SaCas9). Importantly, this effect was not seen with previous exposure to AAV5 (AAV5-mCherry), and subsequent dosing with AAV8-SaCas9 resulted in strong genome editing and PCSK9 knockdown similar to the effects of AAV8-SaCas9 dosing in naïve mice **(Figures 2.4b, 2.4c)**.

Figure 2.4: Redosing in immunized mice. (a) Mice were initially immunized with saline, AAV8-mCherry, AAV5-mCherry, AAV5-*Sa*Cas9 or AAV5-*Sp*Cas9 with no gRNA. After 4 weeks, the mice were given a second dose of saline or AAV8-*Sa*Cas9 with a gRNA targeting *Pcsk9*. Serum was collected before the first injection and again during each subsequent week for 8 weeks. Mice were exposed to antigens by retro-orbital injections at 1×10^{12} VG per mouse. (b) Genome editing rates—quantified by sequencing—are entirely abolished when mice are immunized against AAV8 and moderately inhibited when immunized against *Sa*Cas9. No effect is seen when mice are immunized against AAV5 or *Sp*Cas9. Data are mean ± s.e.m. A one-way analysis of variance (ANOVA) with post hoc Tukey's test was performed to determine statistical differences. $^{**}P = 0.0033$; $^{***}P = 0.0004$; NS, not significant. Each data point represents an individual mouse ($n = 6$ for all panels). (c) Final PCSK9 serum levels (week 8), the phenotypic result of gene editing, decrease sharply after an active second dose of AAV8-*Sa*Cas9 with gRNA. This effect is abolished when mice are first immunized against AAV8, but not when mice are first immunized against AAV5. Previous immunization with AAV5-*Sa*Cas9 reduces but does not eliminate editing, whereas previous dosing with AAV5-SpCas9 has no effect on editing. Data of the full time-course for each week are shown. Data are mean ± s.e.m. Each data point represents an individual mouse ($n = 6$ for all panels). (d) Mice were immunized with both CFA and 5 μg Cas9 1 week before injections with active AAV-SaCas9. (e) At week 3, mice immunized with *Sa*Cas9 show a reduced editing rate compared with mice injected with both CFA and PBS. No change in editing rate is seen when immunized with *Sp*Cas9. Data are mean ± s.e.m. A one-way ANOVA with post hoc Dunnett's test was performed to determine statistical differences. $^{***}P = 0.0002$; $^{**}P = 0.0015$; NS, not significant. Each data point represents an individual mouse ($n = 8$). (f) The reduction in serum PCSK9 is partially inhibited when mice are immunized with both CFA and *Sa*Cas9, but not with CFA and PBS, or CFA and *Sp*Cas9. Data are mean ± s.e.m. Each data point represents an individual mouse ($n = 8$).


a
Week 0:
inject first AAV-Cas9
(no gRNAs)
Week 4:
inject second AAV-Cas9
(gRNA targeting Pcsk9)
Week 8:
assay
b
NS
**
***
Editing rate
First dose
Saline
Saline
AAV8-mCherry
AAV5-mCherry
AAV5-SaCas9
AAV5-SpCas9
Second dose
Saline
AAV8-SaCas9
c
Absorbance
Reduction (%)
Time (weeks)
First dose
Saline
Saline
AAV8-mCherry
AAV5-mCherry
AAV5-SaCas9
AAV5-SpCas9
Second dose
Saline
AAV8-SaCas9
d
Week 0:
inject CFA
+ Cas9
Week 1:
inject AAV-Cas9
(gRNA targeting Pcsk9)
Week 3:
assay
e
NS
***
**
Editing rate
First dose
PBS
PBS
CFA-PBS
CFA-SaCas9
CFA-SpCas9
Second dose
PBS
AAV8-SaCas9
f
Absorbance
Reduction (%)
Time (weeks)
First dose
PBS
PBS
CFA-PBS
CFA-SaCas9
CFA-SpCas9
Second dose
PBS
AAV8-SaCas9

Although we do not necessarily expect this observed orthogonality between AAV8 and AAV5 to carry over into the human setting, here it allowed us to specifically test the effects of the immune response to the Cas9 payload with minimal interference from the AAV delivery vector. Mice first immunized against SaCas9 using AAV5 showed a 35% reduced level of editing, a 38% reduction in PCSK9 decrease, and a wider variation between mice. This may reflect a weak immune response to SaCas9 in our mouse model, and/or a domination of private (individual) T-cell responses to SaCas9. IgG ELISAs revealed that only a subset of mice immunized with AAV5-SaCas9 developed an antibody response. We correlated the level of serum antibodies induced upon SaCas9 immunization with the efficiency of PCSK9 editing after the second dose, finding that mice with a stronger antibody response showed lower editing efficiency.

To replicate these results in a different context, and to verify that immunity to Cas9 specifically can create a barrier to effective gene therapy, we conducted a slightly modified immunization experiment. Here we used a Cas9 protein vaccine combined in emulsion CFA as the initial dose, thereby immunizing a Cas9-specific primary response independently of AAV (**Figure 2.4d**). Subsequent dosing with AAV8-SaCas9 targeting PCSK9 recapitulated the results of AAV-based immunization, showing that prior exposure to the SaCas9 protein, but not SpCas9, significantly reduced the effectiveness of SaCas9-based gene editing by 42% and PCSK9 protein reduction by 51% (**Figure 2.4e, 2.4f**). Taken together, anti-AAV and anti-Cas9 immunity represents a significant obstacle to therapeutic efficacy.

2.2 DISCUSSION

The use of protein therapeutics requires ways to evade the host's immune response. Cas9, as an example, has prokaryotic origins and can evoke a long-lived T-cell response[270,278], which may lead to clearance of transduced cells. In addition, circulating antibodies can neutralize the

AAV vector and prevent efficient transduction upon repeated doses. Immunosuppressive drugs could mitigate some of these aspects, but not without significant side-effects, and are not applicable to patients in poor health[279–282]. Similar to what has been done in cancer antibody therapeutics[283], the SpCas9 protein could be de-immunized by swapping high-immunogenicity domains. This is a promising approach; however, it will be complex and laborious as we anticipate tens of mutations to achieve stealth, which might often result in a reduction in activity and an overall less effective therapy.

Another consideration is that various applications of the CRISPR system will have significantly different immune consequences. For example, contrast genome editing applications, in which only transient expression of Cas9 is needed, to cases of gene regulation (CRISPRi or CRISPRa), in which continuous Cas9 expression is required. While ongoing expression applications will have to continuously contend with T-cell surveillance, genome editing with transient expression may not be compromised by a primary T-cell response if Cas9 expression is lost before CTL activation and clonal expansion. Building on this advantage, we note that promising new techniques may achieve stable gene regulation via transient i.e. hit-and-run approaches using epigenome editing[284]. Despite this, efficacious single-dose therapies may require high titers, especially in cases such as Duchenne muscular dystrophy where systemic delivery is needed. Such high doses may lead to toxicity issues, as demonstrated in a recent study of high-dose AAV9 delivery in rhesus macaques[285]. Multiple lower, non-toxic doses delivered sequentially have the potential to achieve high transduction efficiency but must contend with the stronger and faster secondary adaptive immune response mediated by memory T- and B-cells.

Although we characterize the adaptive immune response to both the AAV VP1 and Cas9 proteins, it is not expected that these will induce the same type nor kinetics of response due to the

differing nature of the antigens. The mice receive VP1 protein in the form of a viral capsid, as contrasted with Cas9 in the form of DNA. The delivery of AAV capsids is expected to produce a strong antibody response through the canonical MHC class II pathway. It may also induce a CTL response through MHC class I presentation via transduction of APCs or cross-presentation of endocytosed viral proteins.

Alternatively, the Cas9 transgene is expressed as protein only once inside a transduced cell, and therefore could be expected to induce both an antibody and CTL response through two separate but non-mutually exclusive mechanisms. One potential avenue is that a subset of AAVs transduce APCs, an event that has been previously observed to occur[274]. After expression, Cas9 may be presented on class I MHC molecules through the canonical pathway, or presented on class II MHC molecules after being encapsulated in autophagosomes (a substantial fraction of MHCII-bound peptides is derived from internal proteins through autophagy)[286]. Another potential mechanism involves transduced non-APCs expressing Cas9 and subsequently undergoing apoptosis or necrosis. APCs then scavenge these dying cells, presenting the Cas9 proteins found within on class II MHC molecules through the canonical pathway, or on class I MHC molecules through cross-presentation, a process important for anti-viral immune responses.

We also note some limitations to our work. Mainly, we have used as our model two inbred mice strains, C57BL/6J and BALB/c, which have limited MHC diversity[287], and might not recapitulate other human immunological features, such as differences in antigen processing and presentation. Our use of highly conservative models of potential human immunity suggests that any immune barriers to gene editing we observe here could be significantly magnified in the human setting. This seems to be corroborated by recent studies of pre-existing immunity to Sp- and SaCas9 in humans[16,278,288] that showed measurable effector and regulatory T-cell responses.

One promising approach to avoid immune mediated inhibition of gene therapy is to harness the ability of Treg cells to promote a more tolerogenic immune response to therapeutic proteins[278]. Additionally, B-cell epitopes can also be predicted and incorporated into immune orthogonality analysis. However, since B-cell epitopes may be both linear and conformational, these are more difficult to predict. Advances and further validation of these *in silico* models will allow for better predictions in the future[289–293]. In our study, initial immunization doses were not delivered with a gRNA, therefore Cas9 produced inside the cell or delivered as a protein will be in the apo-Cas9 conformation. This could result in different B-cell epitopes compared to the gRNA-bound Cas9 complex, as the 3D conformations are substantially different. Note that this should not affect MHC-presented peptides however, and thus not affect T-cell responses.

Here we demonstrated in multiple contexts that 1) AAV-CRISPR-Cas application can induce both a humoral and cell-mediated response to both the delivery vehicle and transgene payload, 2) this response, given enough time, activates T-cells sufficient to specifically clear cells expressing Cas9 epitopes, and 3) these effects combined can lead to complete negation of therapeutic effect in the case of immunity to AAV, and significant reduction of efficacy in the case of immunity to Cas9. Special care will need to be given to addressing these potential pitfalls in the course of developing any gene therapeutic that makes use of AAV delivery or CRISPR-Cas gene editing.

2.3 METHODS

2.3.1 Computational Methods

2.3.1.1 MHC Binding Predictions

Immunologically informed comparison was done in a similar fashion, but using only those k-mers predicted to bind to at least one of 81 HLA-1 alleles using netMHC 4.0[294] for class I (alleles can be found at http://www.cbs.dtu.dk/services/NetMHC/MHC_allele_names.txt), and at least one of 5,620 possible MHC II molecules based on 936 HLA-2 alleles using netMHCIIpan 3.1[295] for class II (alleles can be found at http://www.cbs.dtu.dk/services/NetMHCIIpan-3.1/alleles_name.list). We compared the use of netMHC to alternative immune epitope prediction platforms such as the Immune Epitope Database (iedb.org)[296] and found very strong agreement across software. Ultimately, we chose netMHC because of the larger number of HLA alleles it supports. Sequences were defined as binding if the predicted affinity ranked in the top 2% of a test library of 400,000 random peptides as suggested in the software guidelines.

2.3.1.2 Statistics

PCSK9 ELISA data from immunization experiments (Figures 3, S7), were normalized per mouse to the average of the first 4 weeks of the experiment (during which time no active dose was given), and then analyzed using a two-way repeated measures ANOVA to account for both time and group membership as independent variables. Post hoc Tukey tests were used to compare across groups at each timepoint as shown in Figure 3C.

2.3.2 Experimental Methods

2.3.2.1 Vector design and construction

Split-SpCas9 AAV vectors were constructed by sequential assembly of corresponding gene blocks (IDT) into a custom synthesized rAAV2 vector backbone[297,298]. The first AAV contains a gRNA driven by a human RNA polymerase III promoter, U6, and a N-terminal Cas9 (NCas9)

fused to an N-intein driven by a CMV promoter, as well as a polyadenylation (polyA) signal (Truong et al., Moreno et al.) The second AAV cassette contains a CMV driven C-terminal Cas9 (CCas9) fused to a C-intein as well as a polyA signal. gRNA sequences were inserted into NCas9 plasmids by cloning oligonucleotides (IDT) encoding spacers into AgeI cloning sites via Gibson assembly. pX601-AAV-CMV::NLS-SaCas9-NLS-3xHA-bGHpA;U6::BsaI-sgRNA was a gift from Feng Zhang (Addgene plasmid # 61591)

2.3.2.2 AAV Production

AAV2/8, AAV2/2, AAV2/5, AAV2/DJ virus particles were produced using HEK293T cells via the triple transfection method and purified via an iodixanol gradient[299]. Confluency at transfection was between 80% and 90%. Media was replaced with pre-warmed media 2 hours before transfection. Each virus was produced in 5 x 15 cm plates, where each plate was transfected with 7.5 µg of pXR-capsid (pXR-8, pXR-2, pXR-5, pXR-DJ), 7.5 of µg recombinant transfer vector, and 22.5 µg of pAd5 helper vector using PEI (1µg/µl linear PEI in 1x DPBS pH 4.5, using HCl) at a PEI:DNA mass ratio of 4:1. The mixture was incubated for 10 minutes at RT and then applied dropwise onto the media. The virus was harvested after 72 hours and purified using an iodixanol density gradient ultracentrifugation method. The virus was then dialyzed with 1x PBS (pH 7.2) supplemented with 50 mM NaCl and 0.0001% of Pluronic F68 (Thermo Fisher) using 100kDA filters (Millipore), to a final volume of ~1 ml and quantified by qPCR using primers specific to the ITR region, against a standard (ATCC VR-1616).

AAV-ITR-F: 5'-CGGCCTCAGTGAGCGA-3' and

AAV-ITR-R: 5'-GGAACCCCTAGTGATGGAGTT-3'.

2.3.2.3 Animal studies

All animal procedures were performed in accordance with protocols approved by the Institutional Animal Care and Use Committee (IACUC) of the University of California, San Diego. All mice were acquired from Jackson labs. AAV injections were done in adult male C57BL/6J or BALB/c mice (10 weeks) through retro-orbital injections using $1x10^{12}$ vg/mouse.

2.3.2.4 CFA immunizations

CFA immunizations were prepared by mixing CFA (Sigma-Aldrich) with 5 μg Cas9 protein or PBS at a 1:1 ratio using two syringes connected by an elbow joint to create an even emulsion. 200 μL of CFA emulsion was injected subcutaneously into the flanks of adult mice.

2.3.2.5 ELISA

PCSK9: Levels of serum PCSK9 were measured using the Mouse Proprotein Convertase 9/PCSK9 Quantikine ELISA kit (R&D Systems) according to manufacturer's guidelines. Briefly, serum samples were diluted 1:200 in Calibrator diluent and allowed to bind for 2 h onto microplate wells that were precoated with the capture antibody. Samples were then sequentially incubated with PCSK9 conjugate followed by the PCSK9 substrate solution with extensive intermittent washes between each step. The amount of PCSK9 in serum was estimated colorimetrically using a standard microplate reader (BioRad iMark).

Cas9 and AAV: Recombinant SpCas9 protein (PNA Bio, cat. no. CP01), or SaCas9 protein (ABM good, cat no. K144), was diluted in 1x coating buffer (Bethyl), and 0.5 μg was used to coat each well of 96-well Nunc MaxiSorp Plates (ab210903) overnight at 4 °C. For AAV experiments, 10^9 vg of AAV-2, -5, -8 or -DJ in 1x coating buffer was used to coat each well of 96-well Nuc MaxiSorp Plates. Plates were washed three times for 5 min with 350 μl of 1x Wash Buffer (Bethyl) and blocked with 300 μl of 1x BSA Blocking Solution (Bethyl) for 2 hours at RT. The wash

procedure was repeated. Serum samples were added at 1:40 dilution, and plates were incubated for 5 hours at 4 °C with shaking. Wells were washed three times for 5 min, and 100 µl of HRP-labeled goat anti-mouse IgG1 (Bethyl; diluted 1:100,000 in 1% BSA Blocking Solution) was added to each well. After incubating for 1hr at RT, wells were washed four times for 5 min, and 100 µl of TMB Substrate (Bethyl) was added to each well. Optical density (OD) at 450 nm was measured using a plate reader (BioRad iMark).

2.3.2.6 NGS quantification of editing

Genomic DNA was extracted from samples of mouse livers using a DNA extraction kit (Qiagen). A 200 bp region containing the target cut site of the PCSK9 gene was amplified by PCR using 0.5 µg DNA (~100,000 diploid genomes) as the template. Libraries were prepared using NEBNext Illumina library preparation kit and sequenced on an Illumina HiSeq. Each sample was sequenced to a target depth of 100,000 reads. Adaptors were trimmed from resulting fastqs using AdapterRemoval[300] and NHEJ-repaired cleavage events resulting in a mutation were quantified using CRISPResso[301].

2.3.2.7 Splenocyte Clearance Assay

Splenocyte clearance assays were performed similarly to previous work[302]. Briefly, spleens from adult C57BL/6J mice were harvested and treated to remove erythrocytes and dead cells. These cells were then diluted to 10^7 cells/ml and split into two pools, one of which was pulsed for 40 min with a pool of the 30 most immunogenic T-cell epitopes in SpCas9 (according to our predictions) at 1 µg/ml each and labeled with the CellTrace Violet fluorescent dye (ThermoFisher). The other pool was pulsed with a matching amount of DMSO, and labeled with the green fluorescent dye CFSE (ThermoFisher). A 1:1 mixture of these cells was then injected into naïve

or CFA-immunized mice at week 1 or 3.5 retro-orbitally at 3-6 x 10^7 cells per mouse. Spleens from these mice were harvested 16-20 hours later, treated to remove erythrocytes, and analyzed by flow cytometry to assess the degree of specific clearance of the CTV+ cells which were pulsed with Cas9 peptides.

2.3.2.8 Epitope prediction and peptide synthesis

The MHC-binding peptides for our mouse model were predicted using the netMHC-4.0 and netMHCIIpan-3.1 online software with the alleles H-2-Db and H-2-Kb for class I and H-2-IAb for class II. For MHCII, the top 10 peptides for Sp- and SaCas9 and top 5 peptides for AAV-8 and AAV-DJ by percentile binding were selected for synthesis by Synthetic Biomolecules as crude materials. For MHCI, we selected the top 20 peptides for Sp- and SaCas9 and the top 10 for AAV-8 and AAV-DJ. All peptides were dissolved in DMSO with a concentration of 40 mg ml^{-1} and stored at −20 °C.

2.3.2.9 IFN-γ ELISPOT assay

CD8+ T cells were isolated from splenocytes using magnetic bead positive selection (Miltenyi Biotec) 6 weeks after virus infection. A total of 2×10^5 CD8+ T cells were stimulated with 1×10^5 LPS-blasts loaded with 10 μg of individual peptide in 96-well flat-bottom plates (Immobilon-P, Millipore) that were coated with anti-IFN-γ mAb (clone AN18, Mabtech) in triplicate. Concanavalin A (ConA) was used as positive control. After 20 h of incubation, biotinylated anti-mouse IFN-γ mAb (R4-6A2; Mabtech), followed by ABC peroxidase (Vector Laboratories) and then 3-amino-9-ethylcarbazole (Sigma-Aldrich) were added into the wells. Responses are expressed as number of IFN-γ SFCs per 1×10^6 CD8+ T cells.

2.4 ACKNOWLEDGEMENTS

Chapter 2, in part, is a reprint of the material as it appears in Nature Biomedical Engineering 3, 806–816. (2019). Moreno, Ana; Palmer, Nathan; Alemán, Fernando; Chen, Genghao; Pla, Andrew; Jiang, Ning; Chew, Wei Leong; Law, Mansun; Mali, Prashant. The dissertation author was the primary researcher and author of this paper. We thank members of the Mali laboratory for advice and help with experiments and the Salk GT3 viral core for help with the production of AAVs. This research was supported by UCSD Institutional Funds, the Burroughs Wellcome Fund (1013926), the March of Dimes Foundation (5-FY15-450), the Kimmel Foundation (SKF-16-150), and NIH grants (R01HG009285, RO1CA222826, RO1GM123313, R01AI079031 and R01AI106005). A.M.M. acknowledges a graduate fellowship from CONACYT and UCMEXUS. W.L.C. acknowledges the IAF-PP grant (H17/01/a0/012).

CHAPTER 3 – IMMUNE ORTHOGONAL ORTHOLOGS ENABLE REDOSING OF AAV-CRISPR-CAS

The immune response to protein therapeutics outside of the AAV-CRISPR-Cas system has long been recognized as an important challenge that must be handled thoughtfully before moving a potential drug anywhere close to the clinic. A common approach to circumventing these issues has been to utilize human proteins, or to humanize proteins by substitution of non-human components[303,304]. However, this approach is limited to a small set of therapeutic proteins naturally occurring in humans or closely related species. In addition, although the humanization of proteins can result in a significantly less immunogenic product, they still carry immunological risk[304].

Another way to circumvent an immune response to protein therapeutics is the removal of immunogenic T cell epitopes[305–307]. Once these epitopes are identified, substitution of key amino acids may reduce the protein's immunogenicity since modification of critical anchor residues can abrogate binding to MHC molecules and prevent antigen presentation. However, this can prove difficult due to the massive diversity at HLA loci. As epitope engineering must account for the substrate specificity of each different HLA allele, therapeutics would likely require unique modification for each patient. While epitope deletion/mutation has been successfully applied to several proteins[307,308], this can only preserve protein function when limited to small numbers of HLA alleles unrepresentative of the full diversity. Structural modifications such as PEGylation have also been known to reduce immunogenicity by interfering with antigen-processing mechanisms. However, there is evidence that PEG-specific antibodies are elicited in patients treated with PEGylated therapeutic enzymes[309–312].

In order to facilitate efficacious repeat protein therapies, we propose here the use of orthologous proteins whose function is constrained by natural selection, but whose structure is

subject to diversification by genetic drift. An ortholog, given sufficient sequence divergence, will not cross-react with the immune response generated by exposure to the others, allowing repeat doses to avoid neutralization by existing antibodies and treated cells to avoid clearance by activated CTLs.

As a case study for exploring this approach we focused on the CRISPR-Cas9 system, perhaps the most anticipated therapeutic for gene editing[23,24,313–321]. Comparative genomics has demonstrated that Cas9 proteins are widely distributed across bacterial species and have diversified over an extensive evolutionary history[322–326]. Although there may be pre-existing immunity to Cas9s from pathogenic or commensal species[278,288,327], we hypothesized this diversity could provide a mechanism to circumvent inducing immunological memory by utilizing orthologous Cas9 proteins for each treatment.

In the case of the widely used delivery vehicle of AAVs, the path to extensive immune orthogonal orthologs is significantly murkier. Among well-characterized AAVs known to transduce human cells, only a handful of different serotypes have been discovered. Although AAVs come from a rich lineage of parvoviruses, many of its sister species rely on non-human hosts and have evolved cellular entry mechanisms not well suited for transduction of human cells. Substantial AAV diversification, engineering, and efforts to discover of novel strains are underway, but so far clinical use has been restricted to only a few highly similar AAV serotypes.

To evaluate the immune orthogonality of AAV-delivered CRISPR-Cas systems, we analyzed 284 DNA targeting and 84 RNA targeting CRISPR effectors, and 167 AAV VP1 orthologs. By comparing total sequence similarity as well as predicted binding strengths to class I and class II MHC molecules, we constructed graphs of immune cross-reactivity and computed cliques of proteins that are orthogonal in immunogenicity profiles. Although MHC epitopes do not

predict antibody epitopes, the induction of the more powerful memory response is primarily dependent on reactivation of memory B-cells with help from memory T-cells through the presentation of antigens on class II MHC molecules[328,329]. Next, we experimentally confirmed our immunological predictions by assaying treated mice for induction of protein-targeting antibodies and of T-cell-mediated cytotoxicity against AAV and Cas9 proteins. Finally, we demonstrated in multiple contexts that consecutive dosing with immune orthogonal orthologs circumvents the inhibition of effective gene editing caused by immune recognition of the AAV vector and Cas9 transgene.

3.1 RESULTS

3.1.1 Identifying immune-orthogonal proteins

Natural selection produces diverse structural variants with conserved function in the form of orthologous genes. We assayed the relevance of this diversity for immunological cross-reactivity of 284 DNA targeting and 84 RNA targeting CRISPR effectors **(Table 3.1)** and 167 AAV orthologs **(Table 3.2)** by first comparing their overall amino acid sequence similarities, and second, using a more specific constraint of how their respective amino acid sequences are predicted to bind MHC class I and II molecules **(Figure 3.1f)**. From these analyses we obtained first an estimate of the comprehensive immune overlap among Cas and AAV orthologs based purely at the sequence level, and second a more stringent estimate of predicted immune overlap based on predicted MHC binding **(Figure 3.1g, 3.1h)**. By sequence-level clustering and clique finding methods, we defined many sets of Cas9 orthologs containing up to 9 members with no 6-mer overlap **(Figure 3.2)**. Notably, based on MHC-binding predictions, we find among the set of DNA-targeting Cas proteins (240 Cas9s and 44 Cpf1s) that 79% of pairs are predicted to have non cross-

reacting immune responses, i.e. they are predicted to be orthogonal in immune space **(Figure 3.1g)**. On the contrary, among AAV capsid (VP1 protein) orthologs we did not find full orthogonality up to the 14-mer level, even when restricting predictions with MHC-binding strengths **(Figure 3.1h)**, likely reflecting the strong sequence conservation and shorter evolutionary history of AAVs[330]. This analysis suggests, consistent with previous observations[331,332], that exposure to one AAV serotype can induce broad immunity to all AAVs, which presents a significant challenge to AAV delivery platforms, as some serotypes are prevalent in human populations. Despite the most divergent AAV serotype (AAV5) showing the fewest shared immunogenic peptides, there remain tracts of sequences fully conserved within the VP1 orthologs. As expected, predicted immune cross-reaction negatively correlates with phylogenetic distance **(Figure 3.3)**, though there is significant variation not captured by that regression, suggesting that MHC-binding predictions can refine the choice of sequential orthologs beyond phylogenetic distance alone.

Table 3.1: CRISPR effectors analyzed for immune orthogonality.

Accession ID	Name	Gene family	Nucleic acid
AAK33936.1	AAK33936.1 conserved hypothetical protein [Streptococcus pyogenes M1 GAS]	Cas9	DNA
ADC31648.1	ADC31648.1 Csn1 family CRISPR-associated protein [Mycoplasma gallisepticum str. F]	Cas9	DNA
ADI19058.1	ADI19058.1 uncharacterized protein conserved in bacteria [uncultured delta proteobacterium HF0070_07E19]	Cas9	DNA
ADX75954.1	ADX75954.1 CRISPR-associated protein, Csn1 family [Staphylococcus pseudintermedius ED99]	Cas9	DNA
AEX66236.1	AEX66236.1 CRISPR-associated endonuclease [Corynebacterium diphtheriae C7 (beta)]	Cas9	DNA
APG80630.1	APG80630.1 CRISPR-associated endonuclease Cas9 [Candidatus Parvarchaeum acidiphilum ARMAN-4]	Cas9	DNA
BAK69486.1	BAK69486.1 putative CRISPR associated protein [Campylobacter lari]	Cas9	DNA
CBK78998.1	CBK78998.1 CRISPR-associated endonuclease, Csn1 family [Coprococcus catus GD/7]	Cas9	DNA
CCA84553.1	CCA84553.1 conserved hypothetical protein [Ralstonia syzygii R24]	Cas9	DNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
EEZ71796.1	EEZ71796.1 CRISPR-associated protein, Csn1 family [Neisseria cinerea ATCC 14685]	Cas9	DNA
EFT93846.1	EFT93846.1 CRISPR-associated protein, Csn1 family [Enterococcus faecalis TX0012]	Cas9	DNA
EHN59352.1	EHN59352.1 CRISPR-associated protein [Oenococcus kitaharae DSM 17330]	Cas9	DNA
EIE39736.1	EIE39736.1 Csn1 family CRISPR-associated protein [Mycoplasma canis PG 14]	Cas9	DNA
Ga0054994_10813	Ga0054994_10813 Geobacillus stearothermophilus Cas9	Cas9	DNA
J7RUA5.1	J7RUA5.1 CRISPR-associated endonuclease Cas9 [Staphylococcus aureus]	Cas9	DNA
NP_246064.1	NP_246064.1 hypothetical protein PM1127 [Pasteurella multocida subsp. multocida str. Pm70]	Cas9	DNA
NP_472073.1	NP_472073.1 hypothetical protein lin2744 [Listeria innocua Clip11262]	Cas9	DNA
NP_664481.1	NP_664481.1 hypothetical protein SpyM3_0677 [Streptococcus pyogenes MGAS315]	Cas9	DNA
NP_721764.1	NP_721764.1 hypothetical protein SMU_1405c [Streptococcus mutans UA159]	Cas9	DNA
NP_735360.1	NP_735360.1 hypothetical protein gbs0911 [Streptococcus agalactiae NEM316]	Cas9	DNA
NP_907605.1	NP_907605.1 hypothetical protein WS1445 [Wolinella succinogenes DSM 1740]	Cas9	DNA
NP_907747.1	NP_907747.1 hypothetical protein WS1613 [Wolinella succinogenes DSM 1740]	Cas9	DNA
NP_938445.1	NP_938445.1 hypothetical protein DIP0036 [Corynebacterium diphtheriae NCTC 13129]	Cas9	DNA
NP_970941.1	NP_970941.1 CRISPR-associated Cas5e [Treponema denticola ATCC 35405]	Cas9	DNA
OJI07263.1	OJI07263.1 hypothetical protein BK997_03320 [Candidatus Micrarchaeum acidiphilum ARMAN-1]	Cas9	DNA
tr|I0AP30|I0AP30_IGNAJ	tr|I0AP30|I0AP30_IGNAJ CRISPR-associated endonuclease Cas9 OS=Ignavibacterium album (strain DSM 19864 / JCM 16511 / NBRC 101810 / Mat9-16) OX=945713 GN=cas9	Cas9	DNA
WP_013852048.1	WP_013852048.1 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus pasteurianus]	Cas9	DNA
WP_036475267.1	WP_036475267.1 type II CRISPR RNA-guided endonuclease Cas9 [Neisseria lactamica]	Cas9	DNA
YP_001122287.1	YP_001122287.1 CRISPR-associated large protein [Francisella tularensis subsp. tularensis WY96-3418]	Cas9	DNA
YP_001239928.1	YP_001239928.1 hypothetical protein BBta_3952 [Bradyrhizobium sp. BTAi1]	Cas9	DNA
YP_001296401.1	YP_001296401.1 Probable CRISPR-associated (Cas) protein. csn1 family, subtype Nmeni [Flavobacterium psychrophilum JIP02/86]	Cas9	DNA
YP_001343689.1	YP_001343689.1 CRISPR-associated endonuclease Csn1 family protein [Actinobacillus succinogenes 130Z]	Cas9	DNA
YP_001398821.1	YP_001398821.1 CRISPR-associated Cas5e family protein [Campylobacter jejuni subsp. doylei 269.97]	Cas9	DNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
YP_001411379.1	YP_001411379.1 CRISPR-associated endonuclease Csn1 family protein [Parvibaculum lavamentivorans DS-1]	Cas9	DNA
YP_001450662.1	YP_001450662.1 CRISPR-associated endonuclease Csn1 family protein [Streptococcus gordonii str. Challis substr. CH1]	Cas9	DNA
YP_001483000.1	YP_001483000.1 hypothetical protein C8J_1425 [Campylobacter jejuni subsp. jejuni 81116]	Cas9	DNA
YP_001531750.1	YP_001531750.1 CRISPR-associated protein [Dinoroseobacter shibae DFL 12 = DSM 16493]	Cas9	DNA
YP_001598557.1	YP_001598557.1 hypothetical protein NMCC_0397 [Neisseria meningitidis 053442]	Cas9	DNA
YP_001602368.1	YP_001602368.1 hypothetical protein GDI_2123 [Gluconacetobacter diazotrophicus PA1 5]	Cas9	DNA
YP_001691366.1	YP_001691366.1 hypothetical protein FMG_0058 [Finegoldia magna ATCC 29328]	Cas9	DNA
YP_001875142.1	YP_001875142.1 CRISPR-associated endonuclease Csn1 family protein [Elusimicrobium minutum Pei191]	Cas9	DNA
YP_001878601.1	YP_001878601.1 hypothetical protein Amuc_2010 [Akkermansia muciniphila ATCC BAA-835]	Cas9	DNA
YP_001955845.1	YP_001955845.1 restriction endonuclease [Bifidobacterium longum DJO10A]	Cas9	DNA
YP_001956000.1	YP_001956000.1 Csn1-like CRISPR-associated protein [uncultured Termite group 1 bacterium phylotype Rs-D17]	Cas9	DNA
YP_001956166.1	YP_001956166.1 CRISPR-associated protein Csn1 [uncultured Termite group 1 bacterium phylotype Rs-D17]	Cas9	DNA
YP_002123679.1	YP_002123679.1 CRISPR-Associated Protein Csn1 [Streptococcus equi subsp. zooepidemicus MGCS10565]	Cas9	DNA
YP_002274753.1	YP_002274753.1 Csn1 family CRISPR-associated protein [Gluconacetobacter diazotrophicus PA1 5]	Cas9	DNA
YP_002285828.1	YP_002285828.1 hypothetical protein Spy49_0823 [Streptococcus pyogenes NZ131]	Cas9	DNA
YP_002342100.1	YP_002342100.1 hypothetical protein NMA0631 [Neisseria meningitidis Z2491]	Cas9	DNA
YP_002344900.1	YP_002344900.1 CRISPR-associated protein [Campylobacter jejuni subsp. jejuni NCTC 11168 = ATCC 700819]	Cas9	DNA
YP_002507391.1	YP_002507391.1 CRISPR-associated protein, Csn1 family [Clostridium cellulolyticum H10]	Cas9	DNA
YP_002551549.1	YP_002551549.1 crispr-associated protein, csn1 family [Acidovorax ebreus TPSY]	Cas9	DNA
YP_002937591.1	YP_002937591.1 CRISPR-system related protein [Eubacterium rectale ATCC 33656]	Cas9	DNA
YP_002996940.1	YP_002996940.1 hypothetical protein SDEG_1231 [Streptococcus dysgalactiae subsp. equisimilis GGS_124]	Cas9	DNA
YP_003082577.1	YP_003082577.1 putative CRISPR-associated protein [Neisseria meningitidis alpha14]	Cas9	DNA
YP_003171950.1	YP_003171950.1 CRISPR-associated protein Csn1 [Lactobacillus rhamnosus GG]	Cas9	DNA
YP_003306403.1	YP_003306403.1 CRISPR-associated protein, Csn1 family [Streptobacillus moniliformis DSM 12112]	Cas9	DNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
YP_003360699.1	YP_003360699.1 CRISPR-associated endonuclease Csn1 [Bifidobacterium dentium Bd1]	Cas9	DNA
YP_003448082.1	YP_003448082.1 CRISPR-associated protein, Csn1 family [Azospirillum sp. B510]	Cas9	DNA
YP_003484612.1	YP_003484612.1 hypothetical protein SmuNN2025_0694 [Streptococcus mutans NN2025]	Cas9	DNA
YP_003516037.1	YP_003516037.1 CRISPR associated protein [Helicobacter mustelae 12198]	Cas9	DNA
YP_003552871.1	YP_003552871.1 CRISPR-associated protein, Csn1 family [Candidatus Puniceispirillum marinum IMCC1322]	Cas9	DNA
YP_003574635.1	YP_003574635.1 Csn1 family CRISPR-associated protein [Prevotella ruminicola 23]	Cas9	DNA
YP_003586920.1	YP_003586920.1 hypothetical protein ZPR_4422 [Zunongwangia profunda SM-A87]	Cas9	DNA
YP_003801227.1	YP_003801227.1 CRISPR-associated protein, Csn1 family [Olsenella uli DSM 7084]	Cas9	DNA
YP_003812627.1	YP_003812627.1 hypothetical protein HDN1F_34120 [gamma proteobacterium HdN1]	Cas9	DNA
YP_003937986.1	YP_003937986.1 CRISPR associated protein [Bifidobacterium bifidum S17]	Cas9	DNA
YP_003968716.1	YP_003968716.1 CRISPR-associated protein, Csn1 family (plasmid) [Ilyobacter polytropus DSM 2926]	Cas9	DNA
YP_004049328.1	YP_004049328.1 hypothetical protein NLA_17660 [Neisseria lactamica 020-06]	Cas9	DNA
YP_004168469.1	YP_004168469.1 CRISPR-associated protein, csn1 family [Nitratifractor salsuginis DSM 16511]	Cas9	DNA
YP_004232757.1	YP_004232757.1 CRISPR-associated protein, Csn1 family [Acidovorax avenae subsp. avenae ATCC 19860]	Cas9	DNA
YP_004239119.1	YP_004239119.1 CRISPR-associated protein, Csn1 family [Weeksella virosa DSM 16922]	Cas9	DNA
YP_004248194.1	YP_004248194.1 CRISPR-associated protein, Csn1 family [Sphaerochaeta globosa str. Buddy]	Cas9	DNA
YP_004288385.1	YP_004288385.1 CRISPR-associated protein [Streptococcus gallolyticus subsp. gallolyticus ATCC BAA-2069]	Cas9	DNA
YP_004295898.1	YP_004295898.1 CRISPR-associated protein, Csn1 family [Nitrosomonas sp. AL212]	Cas9	DNA
YP_004345959.1	YP_004345959.1 CRISPR-associated protein, Csn1 family [Fluviicola taffensis DSM 16823]	Cas9	DNA
YP_004373648.1	YP_004373648.1 CRISPR-associated protein, Csn1 family [Coriobacterium glomerans PW2]	Cas9	DNA
YP_004386148.1	YP_004386148.1 CRISPR-associated protein, Csn1 family [Alicycliphilus denitrificans K601]	Cas9	DNA
YP_004399187.1	YP_004399187.1 CRISPR-associated protein, Csn1 family [Lactobacillus buchneri NRRL B-30929]	Cas9	DNA
YP_004401929.1	YP_004401929.1 CRISPR-associated protein, Csn1 family [Streptococcus suis ST3]	Cas9	DNA
YP_004711987.1	YP_004711987.1 hypothetical protein EGYY_25780 [Eggerthella sp. YY7918]	Cas9	DNA
YP_004740688.1	YP_004740688.1 hypothetical protein Ccan_14650 [Capnocytophaga canimorsus Cc5]	Cas9	DNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
YP_004823543.1	YP_004823543.1 hypothetical protein PARA_18570 [Haemophilus parainfluenzae T3T1]	Cas9	DNA
YP_004841202.1	YP_004841202.1 hypothetical protein LSA_08670 [Lactobacillus sanfranciscensis TMW 1.1304]	Cas9	DNA
YP_004843922.1	YP_004843922.1 putative CRISPR-associated (Cas) protein [Flavobacterium branchiophilum FL-15]	Cas9	DNA
YP_004940969.1	YP_004940969.1 putative CRISPR-associated (Cas) protein [Flavobacterium columnare ATCC 49512]	Cas9	DNA
YP_005054169.1	YP_005054169.1 CRISPR-associated protein, Csn1 family [Filifactor alocis ATCC 35896]	Cas9	DNA
YP_005095062.1	YP_005095062.1 CRISPR-associated protein, Csn1 family [Streptococcus macedonicus ACA-DC 198]	Cas9	DNA
YP_005112965.1	YP_005112965.1 hypothetical protein BF638R_3991 [Bacteroides fragilis 638R]	Cas9	DNA
YP_005124303.1	YP_005124303.1 CRISPR-associated endonuclease [Corynebacterium diphtheriae 241]	Cas9	DNA
YP_005143942.1	YP_005143942.1 CRISPR-associated endonuclease [Corynebacterium diphtheriae VA01]	Cas9	DNA
YP_005156749.1	YP_005156749.1 CRISPR-associated endonuclease [Corynebacterium diphtheriae 31A]	Cas9	DNA
YP_005159132.1	YP_005159132.1 CRISPR-associated endonuclease [Corynebacterium diphtheriae BH8]	Cas9	DNA
YP_005163828.1	YP_005163828.1 CRISPR-associated endonuclease [Corynebacterium diphtheriae HC02]	Cas9	DNA
YP_005204039.1	YP_005204039.1 CRISPR-associated protein, SAG0894 family [Streptococcus infantarius subsp. infantarius CJ18]	Cas9	DNA
YP_005388840.1	YP_005388840.1 CRISPR-associated protein Csn1 [Streptococcus pyogenes MGAS15252]	Cas9	DNA
YP_005394506.1	YP_005394506.1 putative BCR [Riemerella anatipestifer ATCC 11845 = DSM 15868]	Cas9	DNA
YP_005586691.1	YP_005586691.1 endonuclease [Bifidobacterium longum subsp. longum KACC 91563]	Cas9	DNA
YP_005658406.1	YP_005658406.1 hypothetical protein CJM1_1467 [Campylobacter jejuni subsp. jejuni M1]	Cas9	DNA
YP_005707461.1	YP_005707461.1 csn1 family CRISPR-associated protein [Enterococcus faecalis OG1RF]	Cas9	DNA
YP_005820658.1	YP_005820658.1 CRISPR-associated protein, Csn1 family [Fibrobacter succinogenes subsp. succinogenes S85]	Cas9	DNA
YP_005824175.1	YP_005824175.1 membrane protein [Francisella cf. tularensis subsp. novicida 3523]	Cas9	DNA
YP_005848005.1	YP_005848005.1 hypothetical protein IALB_3034 [Ignavibacterium album JCM 16511]	Cas9	DNA
YP_005851953.1	YP_005851953.1 hypothetical protein [Lactobacillus delbrueckii subsp. bulgaricus 2038]	Cas9	DNA
YP_005851955.1	YP_005851955.1 hypothetical protein [Lactobacillus delbrueckii subsp. bulgaricus 2038]	Cas9	DNA
YP_005862383.1	YP_005862383.1 CRISPR-associated protein [Lactobacillus johnsonii DPC 6026]	Cas9	DNA
YP_005863223.1	YP_005863223.1 hypothetical protein HN6_00086 [Lactobacillus salivarius CECT 5713]	Cas9	DNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
YP_005879219.1	YP_005879219.1 hypothetical protein NMV_1993 [Neisseria meningitidis 8013]	Cas9	DNA
YP_005880236.1	YP_005880236.1 Csn1 family CRISPR-associated protein [Mycoplasma gallisepticum str. R(high)]	Cas9	DNA
YP_005891308.1	YP_005891308.1 hypothetical protein NMAA_0315 [Neisseria meningitidis WUE 2594]	Cas9	DNA
YP_005963707.1	YP_005963707.1 CRISPR-associated protein [Listeria monocytogenes 10403S]	Cas9	DNA
YP_005966656.1	YP_005966656.1 CRISPR-associated protein [Listeria monocytogenes J0161]	Cas9	DNA
YP_006002222.1	YP_006002222.1 CRISPR-associated endonuclease, Csn1 family [Streptococcus thermophilus ND03]	Cas9	DNA
YP_006002994.1	YP_006002994.1 CRISPR-associated endonuclease, Csn1 family [Streptococcus thermophilus ND03]	Cas9	DNA
YP_006013330.1	YP_006013330.1 hypothetical protein SDE12394_06440 [Streptococcus dysgalactiae subsp. equisimilis ATCC 12394]	Cas9	DNA
YP_006034250.1	YP_006034250.1 CRISPR-associated protein [Streptococcus gallolyticus subsp. gallolyticus ATCC 43143]	Cas9	DNA
YP_006034260.1	YP_006034260.1 CRISPR-associated protein [Streptococcus gallolyticus subsp. gallolyticus ATCC 43143]	Cas9	DNA
YP_006040517.1	YP_006040517.1 CRISPR-associated protein, Csn1 family [Streptococcus thermophilus JIM 8232]	Cas9	DNA
YP_006069875.1	YP_006069875.1 CRISPR-associated endonuclease, Csn1 family [Streptococcus salivarius JIM8777]	Cas9	DNA
YP_006080899.1	YP_006080899.1 CRISPR-associated protein, Csn1 family [Streptococcus suis D9]	Cas9	DNA
YP_006235671.1	YP_006235671.1 CRISPR-associated protein Csn1 [Helicobacter cinaedi PAGU611]	Cas9	DNA
YP_006250978.1	YP_006250978.1 CRISPR-associated protein csn1 [Streptococcus mutans LJ23]	Cas9	DNA
YP_006298249.1	YP_006298249.1 CRISPR-associated protein Cas9/Csn1, subtype II/NMEMI [Prevotella intermedia 17]	Cas9	DNA
YP_006339747.1	YP_006339747.1 CRISPR-system-like protein [Streptococcus thermophilus MN-ZLW-002]	Cas9	DNA
YP_006340526.1	YP_006340526.1 Csn1 [Streptococcus thermophilus MN-ZLW-002]	Cas9	DNA
YP_006359179.1	YP_006359179.1 CRISPR-system-like protein [Streptococcus suis ST1]	Cas9	DNA
YP_006374110.1	YP_006374110.1 hypothetical protein TMO_c0518 [Tistrella mobilis KA081020-065]	Cas9	DNA
YP_006408468.1	YP_006408468.1 CRISPR-associated protein Cas9/Csn1, subtype II/NMEMI [Belliella baltica DSM 15883]	Cas9	DNA
YP_006428096.1	YP_006428096.1 hypothetical protein Ornrh_2170 [Ornithobacterium rhinotracheale DSM 15997]	Cas9	DNA
YP_006487067.1	YP_006487067.1 hypothetical protein EHR_06155 [Enterococcus hirae ATCC 9790]	Cas9	DNA
YP_006490549.1	YP_006490549.1 CRISPR-associated protein csn1 [Streptococcus mutans GS-5]	Cas9	DNA
YP_006536854.1	YP_006536854.1 CRISPR associated protein [Enterococcus faecalis D32]	Cas9	DNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
YP_006583854.1	YP_006583854.1 Csn1 family CRISPR-associated protein [Mycoplasma gallisepticum NY01_2001.047-5-1P]	Cas9	DNA
YP_006586107.1	YP_006586107.1 Csn1 family CRISPR-associated protein [Mycoplasma gallisepticum CA06_2006.052-5-2P]	Cas9	DNA
YP_006680178.1	YP_006680178.1 CRISPR-associated protein [Listeria monocytogenes SLCC2540]	Cas9	DNA
YP_006723584.1	YP_006723584.1 hypothetical protein B739_1085 [Riemerella anatipestifer RA-CH-1]	Cas9	DNA
YP_006726535.1	YP_006726535.1 CRISPR-associated protein, Csn1 family [Lactobacillus buchneri CD034]	Cas9	DNA
YP_006745617.1	YP_006745617.1 CRISPR associated protein [Leuconostoc gelidum JB7]	Cas9	DNA
YP_006752400.1	YP_006752400.1 hypothetical protein BN194_23340 [Lactobacillus casei W56]	Cas9	DNA
YP_006816880.1	YP_006816880.1 CRISPR-associated endonuclease Csn1 family protein [Actinobacillus suis H91-0380]	Cas9	DNA
YP_006858515.1	YP_006858515.1 CRISPR-associated protein [Campylobacter jejuni subsp. jejuni PT14]	Cas9	DNA
YP_006859751.1	YP_006859751.1 CRISPR-associated protein [Streptococcus dysgalactiae subsp. equisimilis RE378]	Cas9	DNA
YP_006866082.1	YP_006866082.1 CRISPR-associated protein Cas9/Csn1, subtype II [Psychroflexus torquis ATCC 700755]	Cas9	DNA
YP_006904759.1	YP_006904759.1 hypothetical protein SDSE_1207 [Streptococcus dysgalactiae subsp. equisimilis AC-2713]	Cas9	DNA
YP_006932845.1	YP_006932845.1 CRISPR-associated protein, Csn1 family [Streptococcus pyogenes A20]	Cas9	DNA
YP_006951177.1	YP_006951177.1 CRISPR-associated protein [Streptococcus agalactiae SA20-06]	Cas9	DNA
YP_007258684.1	YP_007258684.1 Csn1 family CRISPR-associated protein [Mycoplasma cynos C142]	Cas9	DNA
YP_007355826.1	YP_007355826.1 hypothetical protein G148_0828 [Riemerella anatipestifer RA-CH-2]	Cas9	DNA
YP_007413015.1	YP_007413015.1 Hypothetical protein zj316_0330 [Lactobacillus plantarum ZJ316]	Cas9	DNA
YP_007547854.1	YP_007547854.1 CRISPR-associated protein, Csn1 [Bibersteinia trehalosi USDA-ARS-USMARC-192]	Cas9	DNA
YP_007601284.1	YP_007601284.1 CRISPR-associated protein [Helicobacter cinaedi ATCC BAA-847]	Cas9	DNA
YP_007615874.1	YP_007615874.1 hypothetical protein G432_07765 [Sphingomonas sp. MM-1]	Cas9	DNA
YP_007644735.1	YP_007644735.1 CRISPR-associated protein, Csn1 family [Bdellovibrio exovorus JSS]	Cas9	DNA
YP_007801951.1	YP_007801951.1 CRISPR-associated protein, Csn1 family [Gordonibacter pamelaeae 7-10-1-b]	Cas9	DNA
YP_007818384.1	YP_007818384.1 CRISPR-associated protein, Csn1 family [Butyrivibrio fibrisolvens 16/4]	Cas9	DNA
YP_007833383.1	YP_007833383.1 CRISPR-associated endonuclease, Csn1 family [Roseburia intestinalis M50/1]	Cas9	DNA
YP_007968536.1	YP_007968536.1 Hypothetical protein GBS222_0765 [Streptococcus agalactiae]	Cas9	DNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
YP_008027038.1	YP_008027038.1 CRISPR-associated protein Cas9 [Spiroplasma syrphidicola EA-1]	Cas9	DNA
YP_008056512.1	YP_008056512.1 CRISPR-associated protein Cas9/Csn1, subtype II/NMEMI [Streptococcus iniae SF1]	Cas9	DNA
YP_008084634.1	YP_008084634.1 CRISPR-associated protein, Csn1 family [Streptococcus agalactiae 09mas018883]	Cas9	DNA
YP_008086783.1	YP_008086783.1 CRISPR-associated protein, Csn1 family [Streptococcus agalactiae ILRI005]	Cas9	DNA
YP_008117657.1	YP_008117657.1 CRISPR-associated protein, SAG0894 family [Streptococcus agalactiae ILRI112]	Cas9	DNA
YP_008200873.1	YP_008200873.1 CRISPR-associated protein [Lactobacillus casei LOCK919]	Cas9	DNA
YP_008203772.1	YP_008203772.1 CRISPR-associated protein [Lactobacillus rhamnosus LOCK900]	Cas9	DNA
YP_008293597.1	YP_008293597.1 CRISPR-associated protein [Campylobacter jejuni 32488]	Cas9	DNA
YP_008295413.1	YP_008295413.1 CRISPR-associated protein Csn1 [Listeria monocytogenes]	Cas9	DNA
YP_008356623.1	YP_008356623.1 hypothetical protein KE3_1337 [Streptococcus lutetiensis 033]	Cas9	DNA
YP_008416253.1	YP_008416253.1 CRISPR-associated protein Csn1 [Geobacillus sp. JF8]	Cas9	DNA
YP_008434322.1	YP_008434322.1 Csn1 family CRISPR-associated protein [Treponema pedis str. T A4]	Cas9	DNA
YP_008447359.1	YP_008447359.1 CRISPR-associated protein [Lactobacillus paracasei subsp. paracasei 8700:2]	Cas9	DNA
YP_008476484.1	YP_008476484.1 CRISPR-associated protein cas9/csn1, subtype II/nmemi [Fusobacterium nucleatum subsp. vincentii 3_1_36A2]	Cas9	DNA
YP_008509340.1	YP_008509340.1 CRISPR-associated protein [Streptococcus anginosus C1051]	Cas9	DNA
YP_008512693.1	YP_008512693.1 CRISPR-associated protein [Streptococcus intermedius B196]	Cas9	DNA
YP_008562877.1	YP_008562877.1 CRISPR-associated protein [Campylobacter jejuni subsp. jejuni 00-2425]	Cas9	DNA
YP_008618041.1	YP_008618041.1 hypothetical protein BRDCF_09465 [Bacteroides sp. CF50]	Cas9	DNA
YP_008624437.1	YP_008624437.1 CRISPR-associated protein, Csn1 family [Campylobacter jejuni 4031]	Cas9	DNA
YP_008689610.1	YP_008689610.1 CRISPR-associated protein Csn1 [Streptococcus sp. I-G2]	Cas9	DNA
YP_008759810.1	YP_008759810.1 Csn1 family CRISPR-associated protein [Staphylococcus pasteuri SP1]	Cas9	DNA
YP_008779396.1	YP_008779396.1 CRISPR-associated protein [Pediococcus pentosaceus SL4]	Cas9	DNA
YP_008823205.1	YP_008823205.1 hypothetical protein EMQU_0319 [Enterococcus mundtii QU 25]	Cas9	DNA
YP_008868573.1	YP_008868573.1 CRISPR-associated protein Cas9 [Spiroplasma apis B31]	Cas9	DNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
YP_008894594.1	YP_008894594.1 CRISPR-associated protein Csn1 [Mycoplasma gallisepticum S6]	Cas9	DNA
YP_015730.1	YP_015730.1 hypothetical protein MMOB0330 [Mycoplasma mobile 163K]	Cas9	DNA
YP_122507.1	YP_122507.1 hypothetical protein lpp0160 [Legionella pneumophila str. Paris]	Cas9	DNA
YP_139177.1	YP_139177.1 hypothetical protein stu0657 [Streptococcus thermophilus LMG 18311]	Cas9	DNA
YP_141068.1	YP_141068.1 hypothetical protein str0657 [Streptococcus thermophilus CNRZ1066]	Cas9	DNA
YP_213533.1	YP_213533.1 conserved hypothetical protein [Bacteroides fragilis NCTC 9343]	Cas9	DNA
YP_278700.1	YP_278700.1 hypothetical protein MS53_0582 [Mycoplasma synoviae 53]	Cas9	DNA
YP_280216.1	YP_280216.1 cytoplasmic protein [Streptococcus pyogenes MGAS6180]	Cas9	DNA
YP_425545.1	YP_425545.1 CRISPR-associated endonuclease Csn1 family protein [Rhodospirillum rubrum ATCC 11170]	Cas9	DNA
YP_534331.1	YP_534331.1 hypothetical protein RPC_4489 [Rhodopseudomonas palustris BisB18]	Cas9	DNA
YP_534999.1	YP_534999.1 hypothetical protein LSL_0095 [Lactobacillus salivarius UCC118]	Cas9	DNA
YP_568168.1	YP_568168.1 CRISPR-associated Cas5e family protein [Rhodopseudomonas palustris BisB5]	Cas9	DNA
YP_571550.1	YP_571550.1 hypothetical protein Nham_4054 (plasmid) [Nitrobacter hamburgensis X14]	Cas9	DNA
YP_596617.1	YP_596617.1 cytoplasmic protein [Streptococcus pyogenes MGAS9429]	Cas9	DNA
YP_598495.1	YP_598495.1 hypothetical cytosolic protein [Streptococcus pyogenes MGAS10270]	Cas9	DNA
YP_600439.1	YP_600439.1 putative cytoplasmic protein [Streptococcus pyogenes MGAS2096]	Cas9	DNA
YP_602415.1	YP_602415.1 hypothetical protein MGAS10750_Spy0921 [Streptococcus pyogenes MGAS10750]	Cas9	DNA
YP_820161.1	YP_820161.1 CRISPR-system-like protein [Streptococcus thermophilus LMD-9]	Cas9	DNA
YP_820832.1	YP_820832.1 CRISPR-system-like protein [Streptococcus thermophilus LMD-9]	Cas9	DNA
YP_873709.1	YP_873709.1 HNH endonuclease [Acidothermus cellulolyticus 11B]	Cas9	DNA
YP_898402.1	YP_898402.1 membrane protein [Francisella tularensis subsp. novicida U112]	Cas9	DNA
YP_996018.1	YP_996018.1 CRISPR-associated endonuclease Csn1 family protein [Verminephrobacter eiseniae EF01-2]	Cas9	DNA
ZP_00143587.1	ZP_00143587.1 hypothetical protein [Fusobacterium nucleatum subsp. vincentii ATCC 49256]	Cas9	DNA
ZP_02077990.1	ZP_02077990.1 hypothetical protein EUBDOL_01797 [Eubacterium dolichum DSM 3991]	Cas9	DNA
ZP_03683851.1	ZP_03683851.1 hypothetical protein CATMIT_02512, partial [Catenibacterium mitsuokai DSM 15897]	Cas9	DNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
ZP_03755025.1	ZP_03755025.1 hypothetical protein ROSEINA2194_03455 [Roseburia inulinivorans DSM 16841]	Cas9	DNA
ZP_03925169.1	ZP_03925169.1 conserved hypothetical protein [Actinomyces coleocanis DSM 15436]	Cas9	DNA
ZP_03989815.1	ZP_03989815.1 crispr-associated protein [Acidaminococcus sp. D21]	Cas9	DNA
ZP_05061364.1	ZP_05061364.1 CRISPR-associated large protein (provisional), putative [gamma proteobacterium HTCC5015]	Cas9	DNA
ZP_06288774.1	ZP_06288774.1 CRISPR-associated protein, Csn1 family [Prevotella timonensis CRIS 5C-B1]	Cas9	DNA
ZP_06887976.1	ZP_06887976.1 CRISPR-associated protein, Csn1 family [Methylosinus trichosporium OB3b]	Cas9	DNA
ZP_07217791.1	ZP_07217791.1 conserved hypothetical protein [Bacteroides sp. 20_3]	Cas9	DNA
ZP_07316256.1	ZP_07316256.1 CRISPR-associated protein, Csn1 family [Veillonella atypica ACS-134-V-Col7a]	Cas9	DNA
ZP_07398877.1	ZP_07398877.1 csn1 family CRISPR-associated protein [Peptoniphilus duerdenii ATCC BAA-1640]	Cas9	DNA
ZP_07455288.1	ZP_07455288.1 csn1 family CRISPR-associated protein [Eubacterium yurii subsp. margaretiae ATCC 43715]	Cas9	DNA
ZP_07738815.1	ZP_07738815.1 CRISPR-associated protein, Csn1 family [Aminomonas paucivorans DSM 12260]	Cas9	DNA
ZP_07880770.1	ZP_07880770.1 conserved hypothetical protein [Actinomyces sp. oral taxon 180 str. F0310]	Cas9	DNA
ZP_07912707.1	ZP_07912707.1 conserved hypothetical protein [Staphylococcus lugdunensis M23590]	Cas9	DNA
ZP_08015909.1	ZP_08015909.1 hypothetical protein HMPREF9464_01128 [Sutterella wadsworthensis 3_1_45B]	Cas9	DNA
ZP_08029929.1	ZP_08029929.1 CRISPR-associated protein, Csn1 family [Solobacterium moorei F0204]	Cas9	DNA
ZP_08157403.1	ZP_08157403.1 CRISPR-associated protein, Csn1 family [Ruminococcus albus 8]	Cas9	DNA
ZP_08324662.1	ZP_08324662.1 CRISPR-associated protein, Csx12 family [Parasutterella excrementihominis YIT 11859]	Cas9	DNA
ZP_08574780.1	ZP_08574780.1 CRISPR-associated protein, Csn1 family [Lactobacillus coryniformis subsp. torquens KCTC 3535]	Cas9	DNA
ZP_08576281.1	ZP_08576281.1 possible CRISPR associated protein [Lactobacillus farciminis KCTC 3681]	Cas9	DNA
ZP_08660870.1	ZP_08660870.1 possible CRISPR associated protein [Fructobacillus fructosus KCTC 3544]	Cas9	DNA
ZP_08837074.1	ZP_08837074.1 hypothetical protein HMPREF0666_03250 [Prevotella sp. C561]	Cas9	DNA
ZP_09312133.1	ZP_09312133.1 hypothetical protein MoviS_00710 [Mycoplasma ovipneumoniae SC01]	Cas9	DNA
ZP_09352959.1	ZP_09352959.1 CRISPR-associated protein cas9/csn1, subtype II/nmemi [Bacillus smithii 7_3_47FAA]	Cas9	DNA
ZP_09642280.1	ZP_09642280.1 CRISPR-associated protein cas9/csn1, subtype II/nmemi [Odoribacter laneus YIT 12061]	Cas9	DNA
ZP_10010146.1	ZP_10010146.1 CRISPR-associated protein Cas9/Csn1, subtype II/NMEMI [Treponema sp. JC4]	Cas9	DNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
ZP_10206685.1	ZP_10206685.1 CRISPR-associated protein, Csn1 family [Planococcus antarcticus DSM 14505]	Cas9	DNA
ZP_10895610.1	ZP_10895610.1 CRISPR-associated protein Cas9/Csn1, subtype II/NMEMI [Porphyromonas sp. oral taxon 279 str. F0450]	Cas9	DNA
ZP_10898214.1	ZP_10898214.1 CRISPR-associated protein, Csn1 family [Rhodovulum sp. PH10]	Cas9	DNA
ZP_10953934.1	ZP_10953934.1 HNH endonuclease [Alicyclobacillus hesperidum URH17-3-68]	Cas9	DNA
ZP_11022414.1	ZP_11022414.1 CRISPR-associated protein cas9/csn1, subtype II/nmemi [Barnesiella intestinihominis YIT 11860]	Cas9	DNA
ZP_11150502.1	ZP_11150502.1 CRISPR-associated protein, Csn1 family [Alcanivorax pacificus W11-5]	Cas9	DNA
ZP_16930555.1	ZP_16930555.1 csn1 family CRISPR-associated protein [Streptococcus sanguinis SK49]	Cas9	DNA
ZP_17295095.1	ZP_17295095.1 CRISPR-associated protein cas9/csn1, subtype II/nmemi [Bergeyella zoohelcum ATCC 43767]	Cas9	DNA
ZP_18919511.1	ZP_18919511.1 hypothetical protein C882_0672 [Caenispirillum salinarum AK4]	Cas9	DNA
YP_007713556.1	gi|478482906|ref|YP_007713556.1|CASV|cpf1|cpf1	Cpf1	DNA
KK211384.1	KK211384.1 Anaerovibrio sp. RM50 genomic scaffold T525DRAFT_scaffold00007.7, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
KLE02041.1	KLE02041.1 hypothetical protein AA20_01655 [Arcobacter butzleri L348]|cas12a|CRISPR|TypeVA	Cpf1	DNA
KXB38146.1	KXB38146.1 hypothetical protein HMPREF1869_00137 [Bacteroidales bacterium KA00251]|cas12a|CRISPR|TypeVA	Cpf1	DNA
AUKC01000013.1	AUKC01000013.1 Butyrivibrio sp. NC3005 G634DRAFT_scaffold00010.10_C, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
AUKD01000009.1	AUKD01000009.1 Butyrivibrio fibrisolvens MD2001 G635DRAFT_scaffold00009.9_C, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
KKR91555.1	KKR91555.1 hypothetical protein UU43_C0004G0003 [Candidatus Falkowbacteria bacterium GW2011_GWA2_41_14]|cas12a|CRISPR|TypeVA	Cpf1	DNA
KKQ38174.1	KKQ38174.1 hypothetical protein US54_C0016G0015 [Candidatus Roizmanbacteria bacterium GW2011_GWA2_37_7]|cas12a|CRISPR|TypeVA	Cpf1	DNA
KIX20262.1	KIX20262.1 hypothetical protein SY27_14115 [Flavobacterium sp. 316]|cas12a|CRISPR|TypeVA	Cpf1	DNA
EDZ90881.1	EDZ90881.1 hypothetical protein FTG_0873 [Francisella tularensis subsp. novicida FTG]|cas12a|CRISPR|TypeVA	Cpf1	DNA
LFLB01000034.1	LFLB01000034.1 Gammaproteobacteria bacterium LS_SOB contig30559, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
EHR33500.1	EHR33500.1 hypothetical protein HMPREF9709_01099 [Helcococcus kunzii ATCC 51366]|cas12a|CRISPR|TypeVA	Cpf1	DNA
EOS46485.1	EOS46485.1 hypothetical protein C809_02517 [Lachnospiraceae bacterium COE1]|cas12a|CRISPR|TypeVA	Cpf1	DNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
JHWS01000001.1	JHWS01000001.1 Lachnospiraceae bacterium NC2008 T528DRAFT_scaffold00001.1_C, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
KE384587.1	KE384587.1 Moraxella caprae DSM 19149 genomic scaffold K941DRAFT_scaffold00051.51, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
KE384190.1	KE384190.1 Oribacterium sp. NK2B42 genomic scaffold G625DRAFT_scaffold00019.19, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
KK211334.1	KK211334.1 Prevotella brevis ATCC 19188 genomic scaffold T433DRAFT_scaffold00014.14, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
KE384028.1	KE384028.1 Proteocatella sphenisci DSM 23131 genomic scaffold G558DRAFT_scaffold00004.4, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
KE384121.1	KE384121.1 Pseudobutyrivibrio ruminis CF1b genomic scaffold G592DRAFT_scaffold00007.7, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
KFO67989.1	KFO67989.1 hypothetical protein ER57_07115 [Smithella sp. SCADC]|cas12a|CRISPR|TypeVA	Cpf1	DNA
AKC95493.1	AKC95493.1 hypothetical protein VC03_02970 [Sneathia amnii]|cas12a|CRISPR|TypeVA	Cpf1	DNA
KL370853.1	KL370853.1 Succinivibrio dextrinosolvens H5 genomic scaffold T508DRAFT_scaffold00001.1, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
GL995220.1	GL995220.1 Succinivibrionaceae bacterium WG1 genomic scaffold scaffold00033, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
KEJ92204.1	KEJ92204.1 hypothetical protein EH55_04135 [Synergistes jonesii]|cas12a|CRISPR|TypeVA	Cpf1	DNA
KUJ74576.1	KUJ74576.1 hypothetical protein AVO42_04040 [Thiomicrospira sp. XS5]|cas12a|CRISPR|TypeVA	Cpf1	DNA
BBPY01000028.1	BBPY01000028.1 Treponema endosymbiont of Eucomonympha sp. DNA, contig: contig000028, strain: E12, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
BBPZ01000036.1	BBPZ01000036.1 Treponema endosymbiont of Eucomonympha sp. DNA, contig: contig000036, strain: E8, whole genome shotgun sequence|cas12a|CRISPR|TypeVA	Cpf1	DNA
KKQ36153.1	KKQ36153.1 hypothetical protein US52_C0007G0008 [candidate division WS6 bacterium GW2011_GWA2_37_6]|cas12a|CRISPR|TypeVA	Cpf1	DNA
WP_003040289.1	WP_003040289.1 type V CRISPRassociated protein Cpf1 [Francisella tularensis]	Cpf1	DNA
WP_004356401.1	WP_004356401.1 type V CRISPRassociated protein Cpf1 [Prevotella disiens]	Cpf1	DNA
WP_012739647.1	WP_012739647.1 type V CRISPRassociated protein Cpf1 [[Eubacterium] eligens]	Cpf1	DNA
WP_013282991.1	WP_013282991.1 type V CRISPRassociated protein Cpf1 [Butyrivibrio proteoclasticus]	Cpf1	DNA
WP_018359861.1	WP_018359861.1 type V CRISPRassociated protein Cpf1 [Porphyromonas macacae]	Cpf1	DNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
WP_020988726.1	WP_020988726.1 type V CRISPRassociated protein Cpf1 [Leptospira inadai]	Cpf1	DNA
WP_021736722.1	WP_021736722.1 type V CRISPRassociated protein Cpf1 [Acidaminococcus sp. BV3L6]	Cpf1	DNA
KDN25524.1	KDN25524.1 hypothetical protein MBO_03467 [Moraxella bovoculi 237]	Cpf1	DNA
WP_024988992.1	WP_024988992.1 type V CRISPRassociated protein Cpf1 [Prevotella albensis]	Cpf1	DNA
AIZ56868.1	AIZ56868.1 hypothetical protein Mpt1_c09950 [Candidatus Methanoplasma termitum]	Cpf1	DNA
WP_035635841.1	WP_035635841.1 type V CRISPRassociated protein Cpf1 [Lachnospiraceae bacterium ND2006]	Cpf1	DNA
WP_036890108.1	WP_036890108.1 type V CRISPRassociated protein Cpf1 [Porphyromonas crevioricanis]	Cpf1	DNA
WP_044910712.1	WP_044910712.1 type V CRISPRassociated protein Cpf1 [Lachnospiraceae bacterium MC2017]	Cpf1	DNA
WP_044919442.1	WP_044919442.1 type V CRISPRassociated protein Cpf1 [Lachnospiraceae bacterium MA2020]	Cpf1	DNA
KKP36646.1	KKP36646.1 hypothetical protein UR27_C0015G0004 [Candidatus Peregrinibacteria bacterium GW2011_GWA2_33_10]	Cpf1	DNA
KKT48220.1	KKT48220.1 hypothetical protein UW39_C0001G0044 [Parcubacteria group bacterium GW2011_GWC2_44_17]	Cpf1	DNA
JQLU01000005.1	JQLU01000005.1 Carnobacterium gallinarum DSM 4847 strain MT44 BR43DRAFT_scf7180000000012_quiver.5_C, whole genome shotgun sequence|cas13a|CRISPR|Type6A	Cas13	RNA
EDR99618.1	EDR99618.1 hypothetical protein EUBSIR_02687 [[Eubacterium] siraeum DSM 15702]|cas13a|CRISPR|Type6A	Cas13	RNA
JQKK01000015.1	JQKK01000015.1 Lachnospiraceae bacterium MA2020 T348DRAFT_scaffold00014.14_C, whole genome shotgun sequence|cas13a|CRISPR|Type6A	Cas13	RNA
AUJT01000030.1	AUJT01000030.1 Lachnospiraceae bacterium NK4A144 G619DRAFT_scaffold00027.27_C, whole genome shotgun sequence|cas13a|CRISPR|Type6A	Cas13	RNA
ATWC01000054.1	ATWC01000054.1 Lachnospiraceae bacterium NK4A179 G621DRAFT_scaffold00054.54_C, whole genome shotgun sequence|cas13a|CRISPR|Type6A	Cas13	RNA
ACV39665.1	ACV39665.1 hypothetical protein Lebu_1799 [Leptotrichia buccalis C1013b]|cas13a|CRISPR|Type6A	Cas13	RNA
ERL25782.1	ERL25782.1 hypothetical protein HMPREF9108_01633 [Leptotrichia sp. oral taxon 225 str. F0581]|cas13a|CRISPR|Type6A	Cas13	RNA
KB890278.1	KB890278.1 Leptotrichia shahii DSM 19757 genomic scaffold B031DRAFT_scaffold_9.10, whole genome shotgun sequence|cas13a|CRISPR|Type6A	Cas13	RNA
ERK53440.1	ERK53440.1 hypothetical protein HMPREF9015_00520 [Leptotrichia wadei F0279]|cas13a|CRISPR|Type6A	Cas13	RNA
ERK48421.1	ERK48421.1 hypothetical protein HMPREF9015_01858 [Leptotrichia wadei F0279]|cas13a|CRISPR|Type6A	Cas13	RNA
ERK47820.1	ERK47820.1 hypothetical protein HMPREF9015_02301 [Leptotrichia wadei F0279]|cas13a|CRISPR|Type6A	Cas13	RNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
KGL43917.1	KGL43917.1 hypothetical protein EP58_05535 [Listeria newyorkensis]|cas13a|CRISPR|Type6A	Cas13	RNA
CBH27300.1	CBH27300.1 hypothetical protein lse_1149 [Listeria seeligeri serovar 1/2b str. SLCC3954]|cas13a|CRISPR|Type6A	Cas13	RNA
EUJ40903.1	EUJ40903.1 hypothetical protein PWEIH_02614 [Listeria weihenstephanensis FSL R90317]|cas13a|CRISPR|Type6A	Cas13	RNA
ADQ78341.1	ADQ78341.1 hypothetical protein Palpr_0179 [Paludibacter propionicigenes WB4]|cas13a|CRISPR|Type6A	Cas13	RNA
ETE53770.1	ETE53770.1 hypothetical protein U715_11520 [Rhodobacter capsulatus Y262]|cas13a|CRISPR|Type6A	Cas13	RNA
CDC65743.1	CDC65743.1 uncharacterized protein BN714_01570 [Ruminococcus sp. CAG:57]|cas13a|CRISPR|Type6A	Cas13	RNA
LARF01000048.1	LARF01000048.1 Ruminococcus sp. N15.MGS57 contig_47, whole genome shotgun sequence|cas13a|CRISPR|Type6A	Cas13	RNA
ACOK01000100.1	ACOK01000100.1 Ruminococcus flavefaciens FD1 Contig105, whole genome shotgun sequence|cas13a|CRISPR|Type6A	Cas13	RNA
JTLD01000029.1	JTLD01000029.1 Alistipes sp. ZOR0009 L990_29, whole genome shotgun sequence|cas13b|CRISPR|Type6B	Cas13	RNA
GAE20957.1	GAE20957.1 hypothetical protein JCM10003_349 [Bacteroides pyogenes JCM 10003]|cas13b|CRISPR|Type6B	Cas13	RNA
EKB54193.1	EKB54193.1 hypothetical protein HMPREF9699_02005 [Bergeyella zoohelcum ATCC 43767]|cas13b|CRISPR|Type6B	Cas13	RNA
AEK23281.1	AEK23281.1 Hypothetical protein Ccan_11650 [Capnocytophaga canimorsus Cc5]|cas13b|CRISPR|Type6B	Cas13	RNA
KN549099.1	KN549099.1 Chryseobacterium sp. YR477 genomic scaffold EW78DRAFT_scaffold00001.1, whole genome shotgun sequence|cas13b|CRISPR|Type6B	Cas13	RNA
KIX21336.1	KIX21336.1 hypothetical protein SY27_06350 [Flavobacterium sp. 316]|cas13b|CRISPR|Type6B	Cas13	RNA
CCB70204.1	CCB70204.1 Hypothetical protein FBFL15_2182 [Flavobacterium branchiophilum FL15]|cas13b|CRISPR|Type6B	Cas13	RNA
AMA48551.1	AMA48551.1 hypothetical protein AWN65_03295 [Flavobacterium columnare]|cas13b|CRISPR|Type6B	Cas13	RNA
AEW86266.1	AEW86266.1 hypothetical protein FCOL_07235 [Flavobacterium columnare ATCC 49512]|cas13b|CRISPR|Type6B	Cas13	RNA
EHO06562.1	EHO06562.1 hypothetical protein HMPREF9712_03108 [Myroides odoratimimus CCUG 10230]|cas13b|CRISPR|Type6B	Cas13	RNA
EHO08761.1	EHO08761.1 hypothetical protein HMPREF9714_02132 [Myroides odoratimimus CCUG 12901]|cas13b|CRISPR|Type6B	Cas13	RNA
EKB06014.1	EKB06014.1 hypothetical protein HMPREF9711_00870 [Myroides odoratimimus CCUG 3837]|cas13b|CRISPR|Type6B	Cas13	RNA
ADQ80738.1	ADQ80738.1 hypothetical protein Palpr_2606 [Paludibacter propionicigenes WB4]|cas13b|CRISPR|Type6B	Cas13	RNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
KGE88582.1	KGE88582.1 hypothetical protein IX84_07840 [Phaeodactylibacter xiamenensis]|cas13b|CRISPR|Type6B	Cas13	RNA
KGL55352.1	KGL55352.1 hypothetical protein HQ50_05870 [Porphyromonas sp. COT052 OH4946]|cas13b|CRISPR|Type6B	Cas13	RNA
ALJ25681.1	ALJ25681.1 hypothetical protein PGF_00012420 [Porphyromonas gingivalis 381]|cas13b|CRISPR|Type6B	Cas13	RNA
ALJ26029.1	ALJ26029.1 hypothetical protein PGF_00016090 [Porphyromonas gingivalis 381]|cas13b|CRISPR|Type6B	Cas13	RNA
ALO30226.1	ALO30226.1 hypothetical protein PGS_00015470 [Porphyromonas gingivalis A7A128]|cas13b|CRISPR|Type6B	Cas13	RNA
ALA94104.1	ALA94104.1 hypothetical protein PGJ_00015140 [Porphyromonas gingivalis AJW4]|cas13b|CRISPR|Type6B	Cas13	RNA
GAP82240.1	GAP82240.1 hypothetical protein PGANDO_1674 [Porphyromonas gingivalis]|cas13b|CRISPR|Type6B	Cas13	RNA
ERJ81987.1	ERJ81987.1 hypothetical protein HMPREF1988_01768 [Porphyromonas gingivalis F0185]|cas13b|CRISPR|Type6B	Cas13	RNA
ERJ65637.1	ERJ65637.1 hypothetical protein HMPREF1553_02065 [Porphyromonas gingivalis F0568]|cas13b|CRISPR|Type6B	Cas13	RNA
ERJ65503.1	ERJ65503.1 hypothetical protein HMPREF1554_01647 [Porphyromonas gingivalis F0569]|cas13b|CRISPR|Type6B	Cas13	RNA
ERJ64231.1	ERJ64231.1 hypothetical protein HMPREF1555_01956 [Porphyromonas gingivalis F0570]|cas13b|CRISPR|Type6B	Cas13	RNA
AIJ36392.1	AIJ36392.1 hypothetical protein EG14_10345 [Porphyromonas gingivalis]|cas13b|CRISPR|Type6B	Cas13	RNA
EOA10535.1	EOA10535.1 hypothetical protein A343_1752 [Porphyromonas gingivalis JCVI SC001]|cas13b|CRISPR|Type6B	Cas13	RNA
BAK25606.1	BAK25606.1 hypothetical protein PGTDC60_1457 [Porphyromonas gingivalis TDC60]|cas13b|CRISPR|Type6B	Cas13	RNA
ERJ88434.1	ERJ88434.1 hypothetical protein HMPREF1990_01280 [Porphyromonas gingivalis W4087]|cas13b|CRISPR|Type6B	Cas13	RNA
ERJ87335.1	ERJ87335.1 hypothetical protein HMPREF1990_01800 [Porphyromonas gingivalis W4087]|cas13b|CRISPR|Type6B	Cas13	RNA
EIW94777.1	EIW94777.1 hypothetical protein HMPREF1322_1926 [Porphyromonas gingivalis W50]|cas13b|CRISPR|Type6B	Cas13	RNA
AAQ65551.1	AAQ65551.1 hypothetical protein PG_0338 [Porphyromonas gingivalis W83]|cas13b|CRISPR|Type6B	Cas13	RNA
KGN88274.1	KGN88274.1 hypothetical protein HR08_00310 [Porphyromonas gulae]|cas13b|CRISPR|Type6B	Cas13	RNA
KKC50278.1	KKC50278.1 hypothetical protein HR10_10685 [Porphyromonas gulae]|cas13b|CRISPR|Type6B	Cas13	RNA
KGN87312.1	KGN87312.1 hypothetical protein HQ46_09365 [Porphyromonas gulae]|cas13b|CRISPR|Type6B	Cas13	RNA
KGN85385.1	KGN85385.1 hypothetical protein HR15_09830 [Porphyromonas gulae]|cas13b|CRISPR|Type6B	Cas13	RNA
KGN76020.1	KGN76020.1 hypothetical protein HQ40_04325 [Porphyromonas gulae]|cas13b|CRISPR|Type6B	Cas13	RNA
KGO05347.1	KGO05347.1 hypothetical protein HR16_00525 [Porphyromonas gulae]|cas13b|CRISPR|Type6B	Cas13	RNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
KGL48767.1	KGL48767.1 hypothetical protein HQ49_06245 [Porphyromonas gulae]|cas13b|CRISPR|Type6B	Cas13	RNA
KB899147.1	KB899147.1 Porphyromonas gulae DSM 15663 genomic scaffold F452DRAFT_scaffold00001.1, whole genome shotgun sequence|cas13b|CRISPR|Type6B	Cas13	RNA
JHUW01000010.1	JHUW01000010.1 Prevotella sp. MA2016 T360DRAFT_scaffold00006.6_C, whole genome shotgun sequence|cas13b|CRISPR|Type6B	Cas13	RNA
EJP27887.1	EJP27887.1 hypothetical protein HMPREF1146_2324 [Prevotella sp. MSX73]|cas13b|CRISPR|Type6B	Cas13	RNA
KIP58075.1	KIP58075.1 hypothetical protein ST42_02830 [Prevotella sp. P476]|cas13b|CRISPR|Type6B	Cas13	RNA
KIP63359.1	KIP63359.1 hypothetical protein ST44_03600 [Prevotella sp. P5119]|cas13b|CRISPR|Type6B	Cas13	RNA
KIP62088.1	KIP62088.1 hypothetical protein ST45_06380 [Prevotella sp. P5125]|cas13b|CRISPR|Type6B	Cas13	RNA
KIP58950.1	KIP58950.1 hypothetical protein ST43_06385 [Prevotella sp. P560]|cas13b|CRISPR|Type6B	Cas13	RNA
BAKF01000019.1	BAKF01000019.1 Prevotella aurantiaca JCM 15754 DNA, contig: JCM15754.contig00019, whole genome shotgun sequence|cas13b|CRISPR|Type6B	Cas13	RNA
EFU31981.1	EFU31981.1 hypothetical protein HMPREF6485_0083 [Prevotella buccae ATCC 33574]|cas13b|CRISPR|Type6B	Cas13	RNA
JVYU01002440.1	JVYU01002440.1 Prevotella denticola strain 1208_PDEN 5213_4533_340438, whole genome shotgun sequence|cas13b|CRISPR|Type6B	Cas13	RNA
BAJY01000004.1	BAJY01000004.1 Prevotella falsenii DSM 22864 = JCM 15124 DNA, contig: JCM15124.contig00004, whole genome shotgun sequence|cas13b|CRISPR|Type6B	Cas13	RNA
AFJ07523.1	AFJ07523.1 hypothetical protein PIN17_0200 [Prevotella intermedia 17]|cas13b|CRISPR|Type6B	Cas13	RNA
KE392225.1	KE392225.1 Prevotella intermedia ATCC 25611 = DSM 20706 genomic scaffold G553DRAFT_scaffold00013.13, whole genome shotgun sequence|cas13b|CRISPR|Type6B	Cas13	RNA
KJJ86756.1	KJJ86756.1 hypothetical protein M573_117042 [Prevotella intermedia ZT]|cas13b|CRISPR|Type6B	Cas13	RNA
EGQ18444.1	EGQ18444.1 hypothetical protein HMPREF9144_1146 [Prevotella pallens ATCC 700821]|cas13b|CRISPR|Type6B	Cas13	RNA
ERJ98772.1	ERJ98772.1 hypothetical protein HMPREF1218_0639 [Prevotella pleuritidis F0068]|cas13b|CRISPR|Type6B	Cas13	RNA
BAJN01000005.1	BAJN01000005.1 Prevotella pleuritidis JCM 14110 DNA, contig: JCM14110.contig00005, whole genome shotgun sequence|cas13b|CRISPR|Type6B	Cas13	RNA
EKY00089.1	EKY00089.1 hypothetical protein HMPREF9151_01387 [Prevotella saccharolytica F0055]|cas13b|CRISPR|Type6B	Cas13	RNA
BAKN01000001.1	BAKN01000001.1 Prevotella saccharolytica JCM 17484 DNA, contig: JCM17484.contig00001, whole genome shotgun sequence|cas13b|CRISPR|Type6B	Cas13	RNA
AKP71851.1	AKP71851.1 hypothetical protein CG09_1718 [Riemerella anatipestifer]|cas13b|CRISPR|Type6B	Cas13	RNA

Table 3.1: CRISPR effectors analyzed for immune orthogonality… (Continued)

Accession ID	Name	Gene family	Nucleic acid
AGC41344.1	AGC41344.1 hypothetical protein G148_2040 [Riemerella anatipestifer RACH2]|cas13b|CRISPR|Type6B	Cas13	RNA
AKQ40303.1	AKQ40303.1 hypothetical protein AS87_08290 [Riemerella anatipestifer Yb2]|cas13b|CRISPR|Type6B	Cas13	RNA
JTLI01000096.1	JTLI01000096.1 Cetobacterium sp. ZOR0034 L992_97, whole genome shotgun sequence|cas13c|CRISPR|Type6C	Cas13	RNA
KDE72144.1	KDE72144.1 hypothetical protein FUSO8_06265 [Fusobacterium necrophorum DJ2]|cas13c|CRISPR|Type6C	Cas13	RNA
EHO19081.1	EHO19081.1 hypothetical protein HMPREF9466_01873 [Fusobacterium necrophorum subsp. funduliforme 1_1_36S]|cas13c|CRISPR|Type6C	Cas13	RNA
EIJ68860.1	EIJ68860.1 hypothetical protein HMPREF1049_0423 [Fusobacterium necrophorum subsp. funduliforme ATCC 51357]|cas13c|CRISPR|Type6C	Cas13	RNA
JHXW01000011.1	JHXW01000011.1 Fusobacterium perfoetens ATCC 29250 T364DRAFT_scaffold00009.9_C, whole genome shotgun sequence|cas13c|CRISPR|Type6C	Cas13	RNA

Table 3.2: AAV VP1 orthologs analyzed for immune orthogonality.

Accession ID	Name
3J1Q_A	Adeno-associated virus DJ, A Retargeted Gene Therapy Vector
AAB95450.1	capsid protein VP1 [Adeno-associated virus 6]
AAB95452.1	capsid protein VP1 [Adeno-associated virus 3B]
AAO88183.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88184.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88185.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88186.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88187.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88188.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88189.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88190.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88191.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88192.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88193.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88194.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88195.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88196.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88197.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88198.1	capsid protein [Non-human primate Adeno-associated virus]

Table 3.2: AAV VP1 orthologs analyzed for immune orthogonality… (Continued)

Accession ID	Name
AAO88199.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88200.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88201.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88202.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88203.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88204.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88205.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88206.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88207.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88208.1	capsid protein [Non-human primate Adeno-associated virus]
AAO88209.1	capsid protein [Non-human primate Adeno-associated virus]
AAR07955.1	capsid protein [Snake adeno-associated virus]
AAS99238.1	capsid protein VP1 [Adeno-associated virus]
AAS99239.1	capsid protein VP1 [Adeno-associated virus]
AAS99240.1	capsid protein VP1 [Adeno-associated virus]
AAS99241.1	capsid protein VP1 [Adeno-associated virus]
AAS99242.1	capsid protein VP1 [Adeno-associated virus]
AAS99243.1	capsid protein VP1 [Adeno-associated virus]
AAS99244.1	capsid protein VP1 [Adeno-associated virus]
AAS99245.1	capsid protein VP1 [Adeno-associated virus]
AAS99246.1	capsid protein VP1 [Adeno-associated virus]
AAS99247.1	capsid protein VP1 [Adeno-associated virus]
AAS99248.1	capsid protein VP1 [Adeno-associated virus]
AAS99249.1	capsid protein VP1 [Adeno-associated virus]
AAS99250.1	capsid protein VP1 [Adeno-associated virus]
AAS99251.1	capsid protein VP1 [Adeno-associated virus]
AAS99252.1	capsid protein VP1 [Adeno-associated virus]
AAS99253.1	capsid protein VP1 [Adeno-associated virus]
AAS99254.1	capsid protein VP1 [Adeno-associated virus]
AAS99255.1	capsid protein VP1 [Adeno-associated virus]
AAS99256.1	capsid protein VP1 [Adeno-associated virus]
AAS99257.1	capsid protein VP1 [Adeno-associated virus]
AAS99258.1	capsid protein VP1 [Adeno-associated virus]
AAS99259.1	capsid protein VP1 [Adeno-associated virus]
AAS99260.1	capsid protein VP1 [Adeno-associated virus]
AAS99261.1	capsid protein VP1 [Adeno-associated virus]

Table 3.2: AAV VP1 orthologs analyzed for immune orthogonality… (Continued)

Accession ID	Name
AAS99262.1	capsid protein VP1 [Adeno-associated virus]
AAS99263.1	capsid protein VP1 [Adeno-associated virus]
AAS99264.1	capsid protein VP1 [Adeno-associated virus 9]
AAS99265.1	capsid protein VP1 [Adeno-associated virus]
AAS99266.1	capsid protein VP1 [Adeno-associated virus]
AAS99267.1	capsid protein VP1 [Adeno-associated virus]
AAS99268.1	capsid protein VP1 [Adeno-associated virus]
AAS99269.1	capsid protein VP1 [Adeno-associated virus]
AAS99270.1	capsid protein VP1 [Adeno-associated virus]
AAS99271.1	capsid protein VP1 [Adeno-associated virus]
AAS99273.1	capsid protein VP1 [Adeno-associated virus]
AAS99274.1	capsid protein VP1 [Adeno-associated virus]
AAS99275.1	capsid protein VP1 [Adeno-associated virus]
AAS99276.1	capsid protein VP1 [Adeno-associated virus]
AAS99277.1	capsid protein VP1 [Adeno-associated virus]
AAS99278.1	capsid protein VP1 [Adeno-associated virus]
AAS99279.1	capsid protein VP1 [Adeno-associated virus]
AAS99280.1	capsid protein VP1 [Adeno-associated virus]
AAS99281.1	capsid protein VP1 [Adeno-associated virus]
AAS99282.1	capsid protein VP1 [Adeno-associated virus]
AAS99283.1	capsid protein VP1 [Adeno-associated virus]
AAS99284.1	capsid protein VP1 [Adeno-associated virus]
AAS99285.1	capsid protein VP1 [Adeno-associated virus]
AAS99286.1	capsid protein VP1 [Adeno-associated virus]
AAS99287.1	capsid protein VP1 [Adeno-associated virus]
AAS99288.1	capsid protein VP1 [Adeno-associated virus]
AAS99289.1	capsid protein VP1 [Adeno-associated virus]
AAS99290.1	capsid protein VP1 [Adeno-associated virus]
AAS99291.1	capsid protein VP1 [Adeno-associated virus]
AAS99292.1	capsid protein VP1 [Adeno-associated virus]
AAS99293.1	capsid protein VP1 [Adeno-associated virus]
AAS99294.1	capsid protein VP1 [Adeno-associated virus]
AAS99295.1	capsid protein VP1 [Adeno-associated virus]
AAS99296.1	capsid protein VP1 [Adeno-associated virus]
AAS99297.1	capsid protein VP1 [Adeno-associated virus]
AAS99298.1	capsid protein VP1 [Adeno-associated virus]

Table 3.2: AAV VP1 orthologs analyzed for immune orthogonality… (Continued)

Accession ID	Name
AAS99299.1	capsid protein VP1 [Adeno-associated virus]
AAS99300.1	capsid protein VP1 [Adeno-associated virus]
AAS99301.1	capsid protein VP1 [Adeno-associated virus]
AAS99302.1	capsid protein VP1 [Adeno-associated virus]
AAS99303.1	capsid protein VP1 [Adeno-associated virus]
AAS99304.1	capsid protein VP1 [Adeno-associated virus]
AAS99305.1	capsid protein VP1 [Adeno-associated virus]
AAS99306.1	capsid protein VP1 [Adeno-associated virus]
AAS99307.1	capsid protein VP1 [Adeno-associated virus]
AAS99308.1	capsid protein VP1 [Adeno-associated virus]
AAS99309.1	capsid protein VP1 [Adeno-associated virus]
AAS99310.1	capsid protein VP1 [Adeno-associated virus]
AAS99311.1	capsid protein VP1 [Adeno-associated virus]
AAS99312.1	capsid protein VP1 [Adeno-associated virus]
AAS99313.1	capsid protein VP1 [Adeno-associated virus]
AAS99314.1	capsid protein VP1 [Adeno-associated virus]
AAT46337.1	capsid protein VP1 [Adeno-associated virus 10]
AAT46339.1	capsid protein VP1 [Adeno-associated virus 11]
AAT48613.1	capsid protein [Avian adeno-associated virus ATCC VR-865]
AAT48615.1	capsid protein [Avian adeno-associated virus strain DA-1]
AAU05358.1	capsid protein VP1 [Adeno-associated virus]
AAU05360.1	capsid protein VP1 [Adeno-associated virus]
AAU05362.1	capsid protein VP1 [Adeno-associated virus]
AAU05364.1	capsid protein VP1 [Adeno-associated virus]
AAU05368.1	capsid protein VP1 [Adeno-associated virus]
AAU05370.1	capsid protein VP1 [Adeno-associated virus]
AAU05371.1	capsid protein VP1 [Adeno-associated virus]
AAU05372.1	capsid protein VP1 [Adeno-associated virus]
AAZ79672.1	VP1 capsid [Mouse adeno-associated virus 1]
AAZ79676.1	VP1 capsid [Rat adeno-associated virus 1]
ABA71699.1	capsid protein [Adeno-associated virus VR-195]
ABA71701.1	capsid protein [Adeno-associated virus VR-355]
ABI16639.1	capsid protein VP1 [Adeno-associated virus 12]
ABZ10812.1	capsid protein VP1 [Adeno-associated virus 13]
ACB55301.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55302.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]

Table 3.2: AAV VP1 orthologs analyzed for immune orthogonality… (Continued)

Accession ID	Name
ACB55303.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55304.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55305.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55306.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55307.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55308.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55309.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55310.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55311.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55312.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55313.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55314.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55315.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55316.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
ACB55317.1	capsid protein VP1, partial (endogenous virus) [Adeno-associated virus]
AGA15924.1	capsid protein VP1 [Adeno-associated virus]
AGA15925.1	capsid protein VP1 [Adeno-associated virus]
AGA15926.1	capsid protein VP1 [Adeno-associated virus]
AGA15927.1	capsid protein VP1 [Adeno-associated virus]
AHK22793.1	capsid protein [Avian adeno-associated virus]
AKU89595.1	capsid protein VP1 [Adeno-associated virus - Anc80L65]
AKU89596.1	capsid protein [Adeno-associated virus]
AKU89597.1	capsid protein [Adeno-associated virus]
AKU89598.1	capsid protein [Adeno-associated virus]
AKU89599.1	capsid protein [Adeno-associated virus]
AKU89600.1	capsid protein [Adeno-associated virus]
AKU89601.1	capsid protein [Adeno-associated virus]
AKU89602.1	capsid protein [Adeno-associated virus]
AKU89603.1	capsid protein [Adeno-associated virus]
AOL02447.1	capsid protein [Adeno-associated virus]
APD78414.1	capsid protein [Simian Adeno-associated virus]
NP_043941.1	capsid protein VP1 [Adeno-associated virus 3]
NP_044927.1	capsid protein VP1 [Adeno-associated virus 4]
NP_049542.1	capsid protein VP1 [Adeno-associated virus 1]
pdb|4IOV|A	capsid protein VP1 [Adeno-associated virus - rh32.33]
YP_068409.1	capsid protein VP1 [Adeno-associated virus 5]

Table 3.2: AAV VP1 orthologs analyzed for immune orthogonality… (Continued)

Accession ID	Name
YP_077178.1	capsid protein VP1 [Adeno-associated virus 7]
YP_077180.1	capsid protein VP1 [Adeno-associated virus 8]
YP_077183.1	capsid protein [Avian adeno-associated virus strain DA-1]
YP_680426.1	capsid protein VP1 [Adeno-associated virus 2]

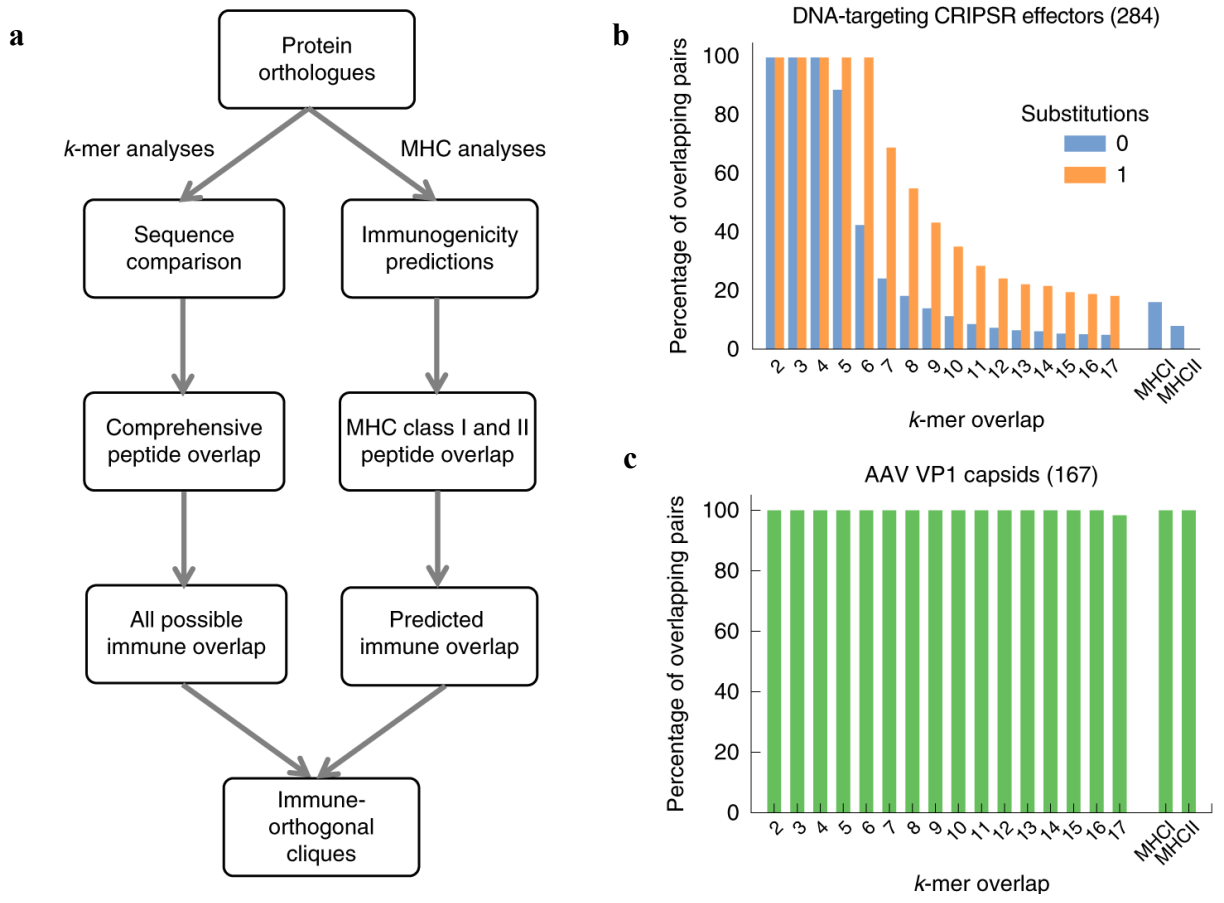


Figure 3.1: Assessing immune orthogonality among CRISPR effectors and AAV capsids. (a) The *in silico* workflow used to find immune-orthogonal protein-homologue cliques. (b) Immunologically uninformed sequence comparison was carried out by checking all *k*-mers in a protein for their presence in another protein sequence with either zero or one mismatch. The *x* axis corresponds to *k*, whereas MHCI and MHCII show overlap only of peptides predicted to bind to MHC class I and class II molecules; 48% of Cas9 pairs show no 6-mer overlap, and 79% of pairs show no overlapping MHC-binding peptides. (c) Immunologically uninformed sequence comparison as in (b) but for AAV VP1 capsid proteins. All AAV pairs contain overlapping MHC-binding peptides.

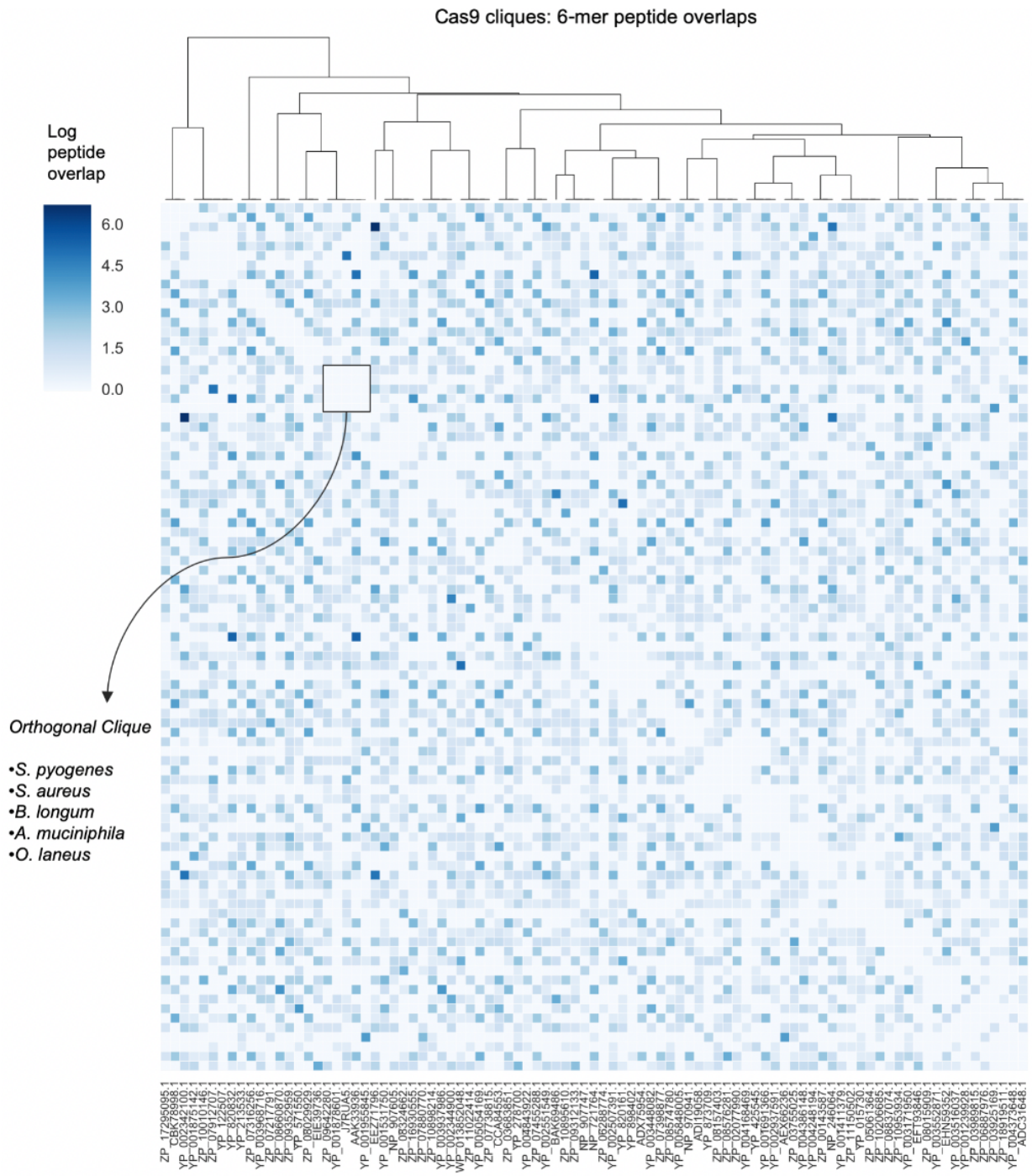


Figure 3.2: Cas9 immune orthogonal cliques. Cliques corresponding to 6-mer overlaps are depicted. An example of an orthogonal clique is highlighted, which includes Cas9s from: S. pyogenes, S. aureus, B. longum, A. muciniphila, and O. laneus.

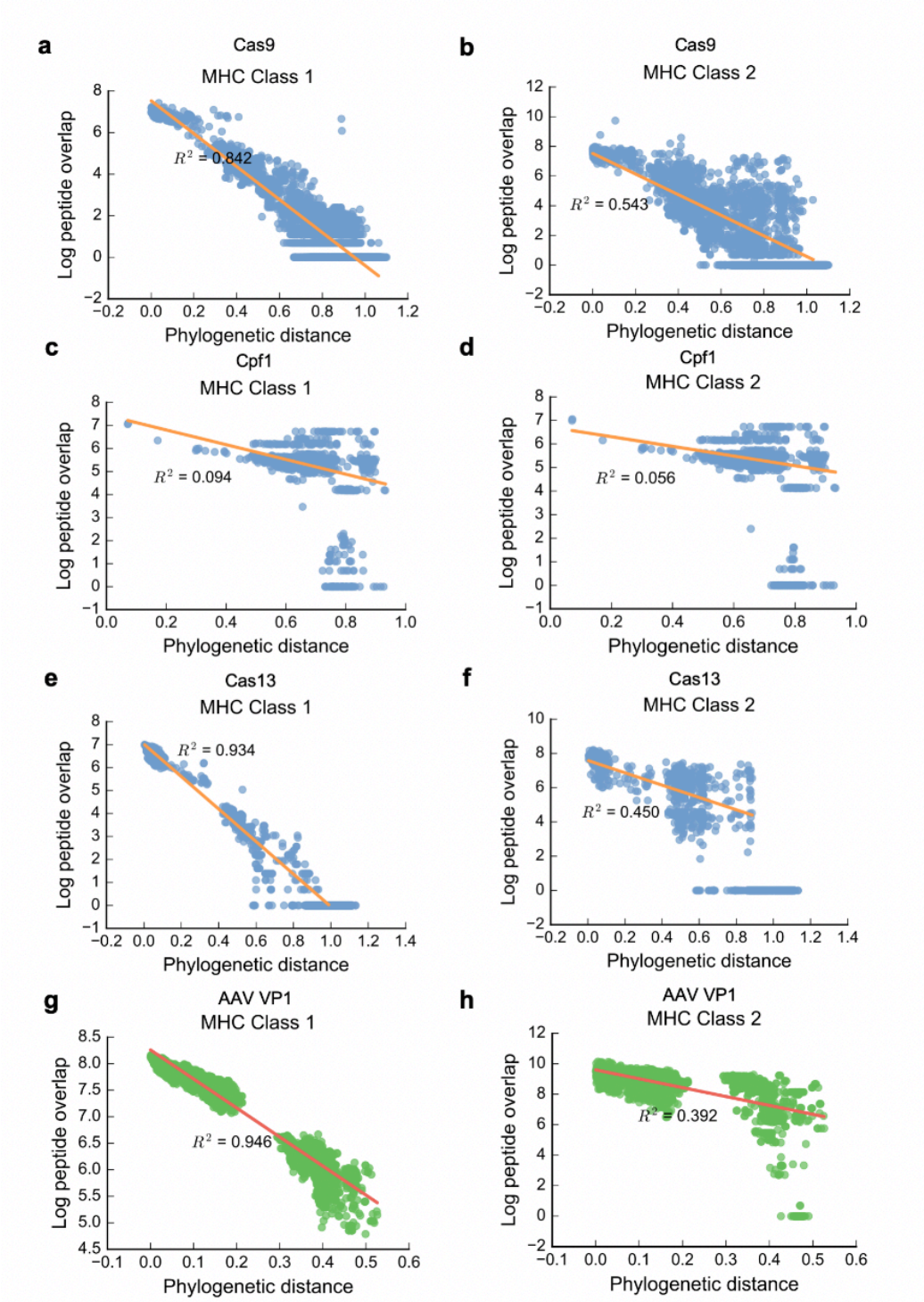
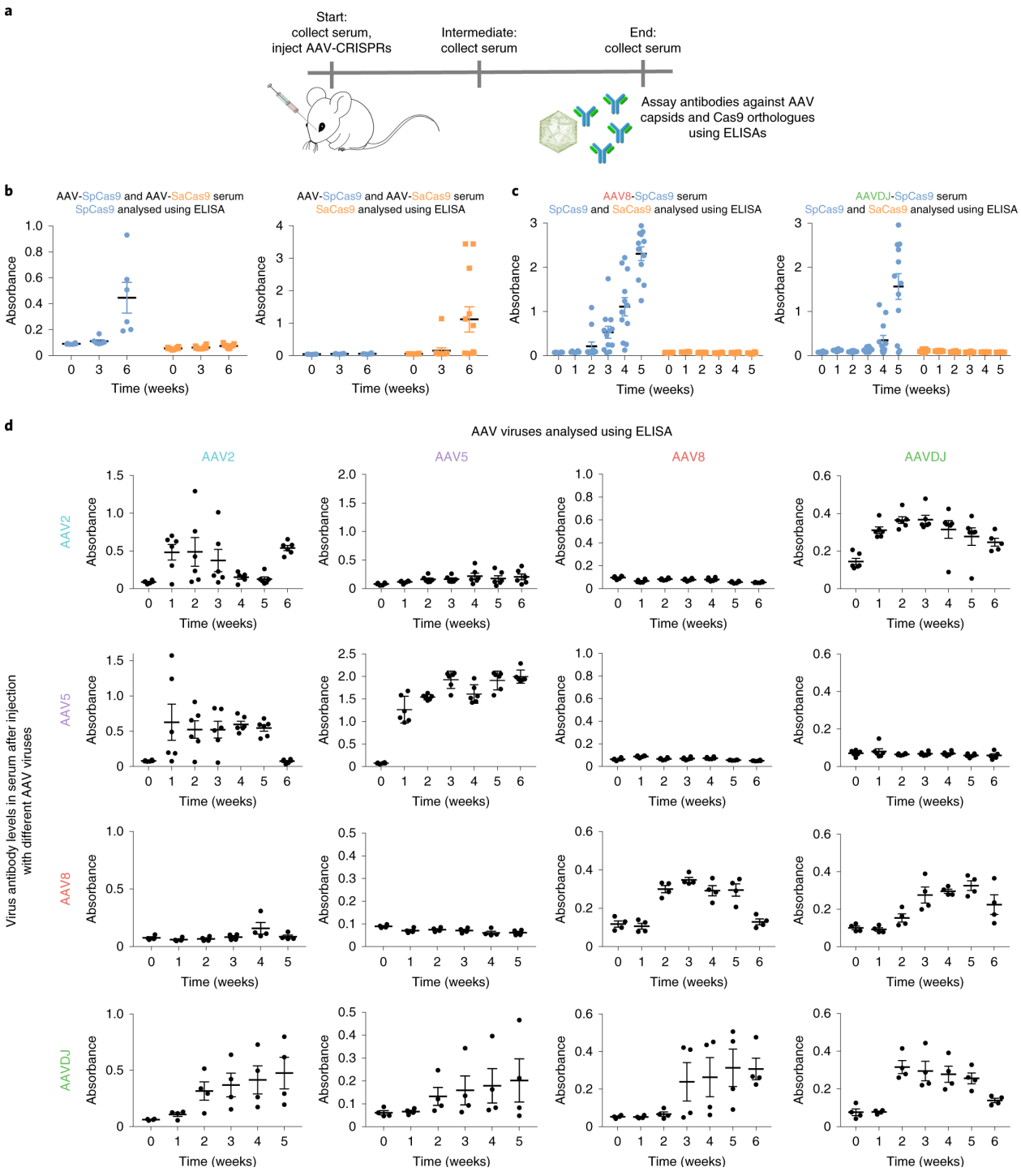


Figure 3.3: **In silico analyses of immunogenicity of Cas9 and AAV orthologues.** Linear regressions exclude pairs of orthologues with no overlap. (a, b) Cas9 MHC class I and class II peptide overlap vs. phylogenetic distance. (c, d) Cpf1 MHC class I and class II peptide overlap vs. phylogenetic distance. (e, f) Cas13 MHC class I and class II peptide overlap vs. phylogenetic distance. (g, h) AAV VP1 MHC class I and class II peptide overlap vs. phylogenetic distance.

3.1.2 Confirming humoral immune orthogonality among Cas9 proteins

To test our immunological predictions and to establish the utility of this approach, we narrowed in on a 5-member clique containing the ubiquitously used *S. pyogenes* Cas9 in addition to the well-characterized *S. aureus* Cas9 **(Figure 3.2)**. To determine whether either of these proteins have cross-reacting antibody responses, we injected mice with 10^{12} vg of either AAV8 or AAVDJ capsids containing SaCas9 or SpCas9 transgenes via retro-orbital injections and harvested serum at days 0 (pre-injection), and periodically over 4-6 weeks **(Figure 3.4a)**. SpCas9-specific antibodies were detected in the plasma of all mice injected with SpCas9 (n=6), and notably none of the mice injected with SaCas9 (n=12) **(Figure 3.4b)**. Half of the mice injected with SaCas9 AAVs (n=12) developed detectable antibodies against SaCas9, whereas none of the mice injected with SpCas9 AAVs (n=6) developed an antibody response against SaCas9. These results were confirmed in an independent study in which SpCas9-specific antibodies, but not SaCas9-specific antibodies, were detected in the plasma of mice injected with AAV-SpCas9 (n=12). These mice were injected retro-orbitally with 10^{12} vg of AAV8-SpCas9 or AAVDJ-SpCas9, and also received an additional intramuscular injection with 10^{11} vg at week 4. **(Figure 3.4c).** Taken together, our data confirms that SpCas9 and SaCas9 have humoral immune orthogonality. As an additional step, we tested another Cas9 ortholog from *Campylobacter jejuni*, useful for AAV-based delivery due to its small size. Mice injected retro-orbitally with 10^{12} vg AAV8-CjCas9 (n=12) showed no significant humoral response to Sp- or SaCas9 after 4 weeks (**Figure 3.5**), confirming immune orthogonality for a set of 3 unique Cas9 orthologs.

Figure 3.4: **Humoral immune orthogonality among Cas9 and AAV proteins**. (a) Mice were exposed to antigens by retro-orbital injections at 1×10^{12} VG per mouse. Serum was collected before injection on day 0, and at multiple points over the course of 4–6 weeks. (b) Levels of anti-*Sp*Cas9 antibodies generated in mice injected with either *Sp*Cas9 ($n = 6$) or *Sa*Cas9 ($n = 12$), and levels of anti-*Sa*Cas9 antibodies generated in mice injected with either *Sp*Cas9 ($n = 6$) or *Sa*Cas9 ($n = 12$). Data are mean ± s.e.m. Each data point represents an individual mouse. (c) Levels of anti-*Sp*Cas9 and anti-*Sa*Cas9 antibodies generated by mice injected retro-orbitally 1×10^{12} VG on day 0 and intramuscularly with 1×10^{11} VG on week 4 with AAV8-*Sp*Cas9 ($n = 12$; left) or AAVDJ-*Sp*Cas9 ($n = 12$; right). Data are mean ± s.e.m. Each data point represents an individual mouse. (d) Levels of anti-AAV8/AAVDJ/AAV2/AAV5 antibodies generated in mice injected with either AAV8 or AAVDJ ($n = 4$ for all panels), or with either AAV2 or AAV5 ($n = 6$ for all panels except for the AAVDJ serum ELISA at the week 6 time point for which $n = 5$). Data are mean ± s.e.m. Each data point represents an individual mouse.


a
Start:
collect serum,
inject AAV-CRISPRs
Intermediate:
collect serum
End:
collect serum
Assay antibodies against AAV
capsids and Cas9 orthologues
using ELISAs
b
AAV-SpCas9 and AAV-SaCas9 serum
SpCas9 analysed using ELISA
AAV-SpCas9 and AAV-SaCas9 serum
SaCas9 analysed using ELISA
c
AAV8-SpCas9 serum
SpCas9 and SaCas9 analysed using ELISA
AAVDJ-SpCas9 serum
SpCas9 and SaCas9 analysed using ELISA
Absorbance
Time (weeks)
d
AAV viruses analysed using ELISA
AAV2
AAV5
AAV8
AAVDJ
Virus antibody levels in serum after injection
with different AAV viruses

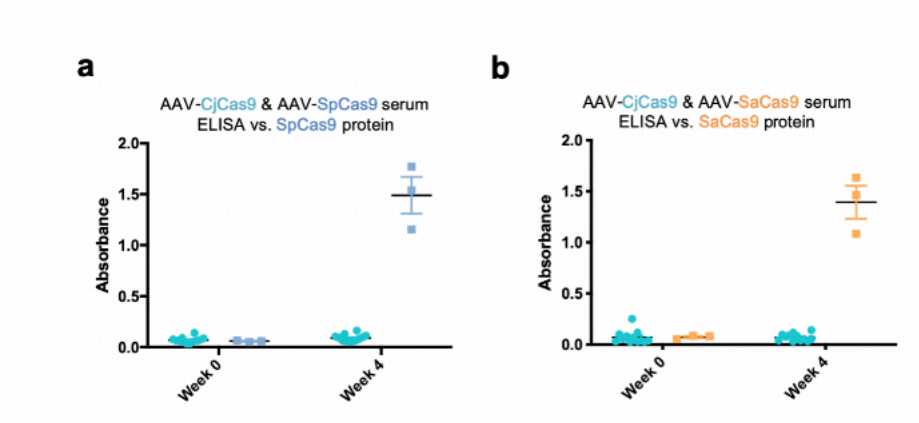


Figure 3.5: Confirming immune orthogonality of C. jejuni Cas9 to Sp- and SaCas9. Anti-SpCas9 antibodies generated in mice injected with SpCas9 (n=3) and CjCas9 (n=12), and anti-SaCas9 antibodies generated in mice injected with SaCas9 (n=3) and CjCas9 (n=12) are shown. Results are shown as mean ± s.e.m. Each data point represents an individual mouse. Mice injected retro-orbitally with 1x10 12 vg/mouse AAV8-CjCas9 targeting PCSK9 showed no significant response to Sp- or SaCas9.

3.1.3 Confirming broad immune cross-reactivity among AAV serotypes

AAVs are becoming a preferred delivery vehicle due to their ability to avoid induction of a strong CD8+ T-cell response, however, the presence of neutralizing antibodies remains a significant barrier to successful application of AAV therapies. Consistent with previous results[331], we found shared immunogenic peptides among all human AAV serotypes **(Figure 3.6)**. We confirmed the lack of orthogonality for two serotypes, AAV8 and AAVDJ, in which we found that antibodies produced in mice injected with AAV8 or AAVDJ react to both AAV8 and AAVDJ antigens **(Figure 3.4d)**. Our analysis suggests that there are no two known AAVs for which exposure to one would guarantee immune naïveté to another across all MHC genotypes. However, immune cross-reaction could be minimized through the use of AAV5[332,333], the most phylogenetically divergent serotype**.** Our predictions identify only a single shared highly immunogenic peptide between AAV5 and the commonly used AAV2 and AAV8 in the mouse model (though several other shared peptides of mild MHC affinity exist). We confirmed this via

ELISAs, where mice injected with AAV2 did not elicit antibodies against AAV5 and AAV8, and mice injected with AAV5 did not elicit antibodies against AAVDJ and AAV8 **(Figure 3.4d)**.

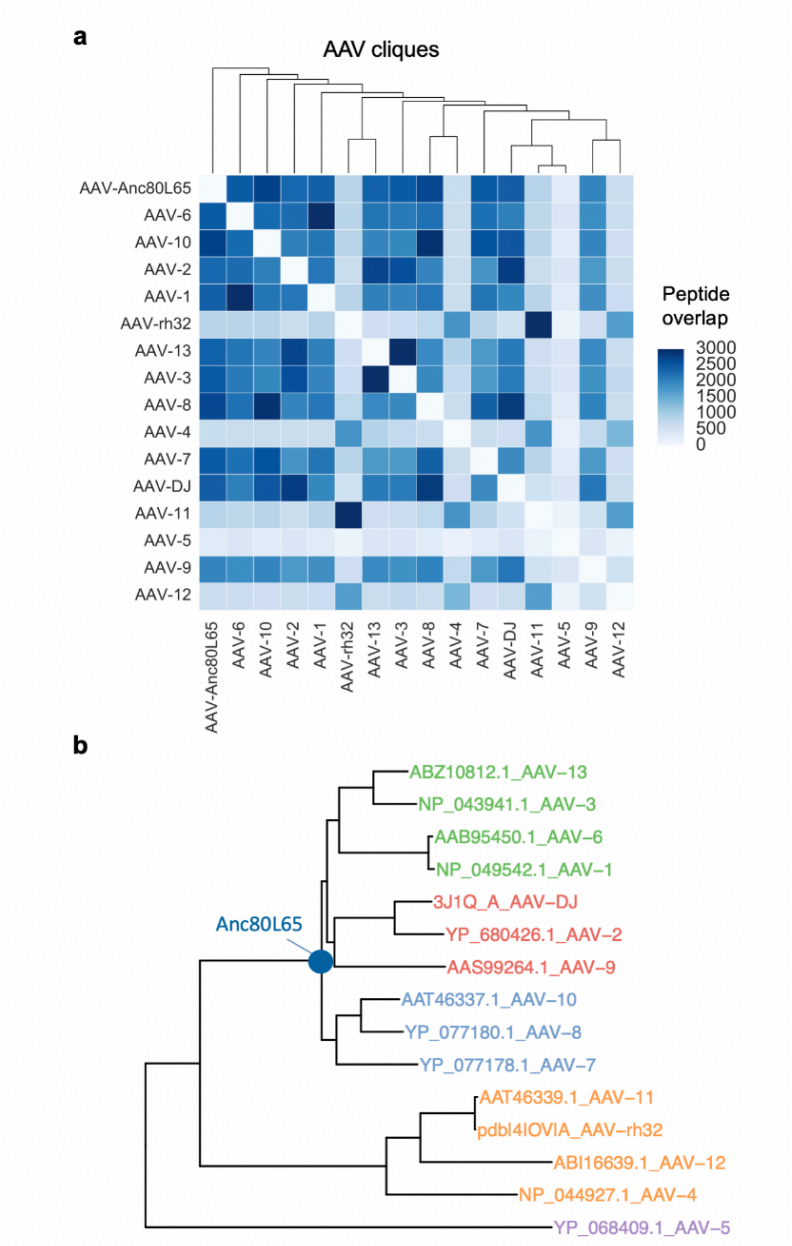
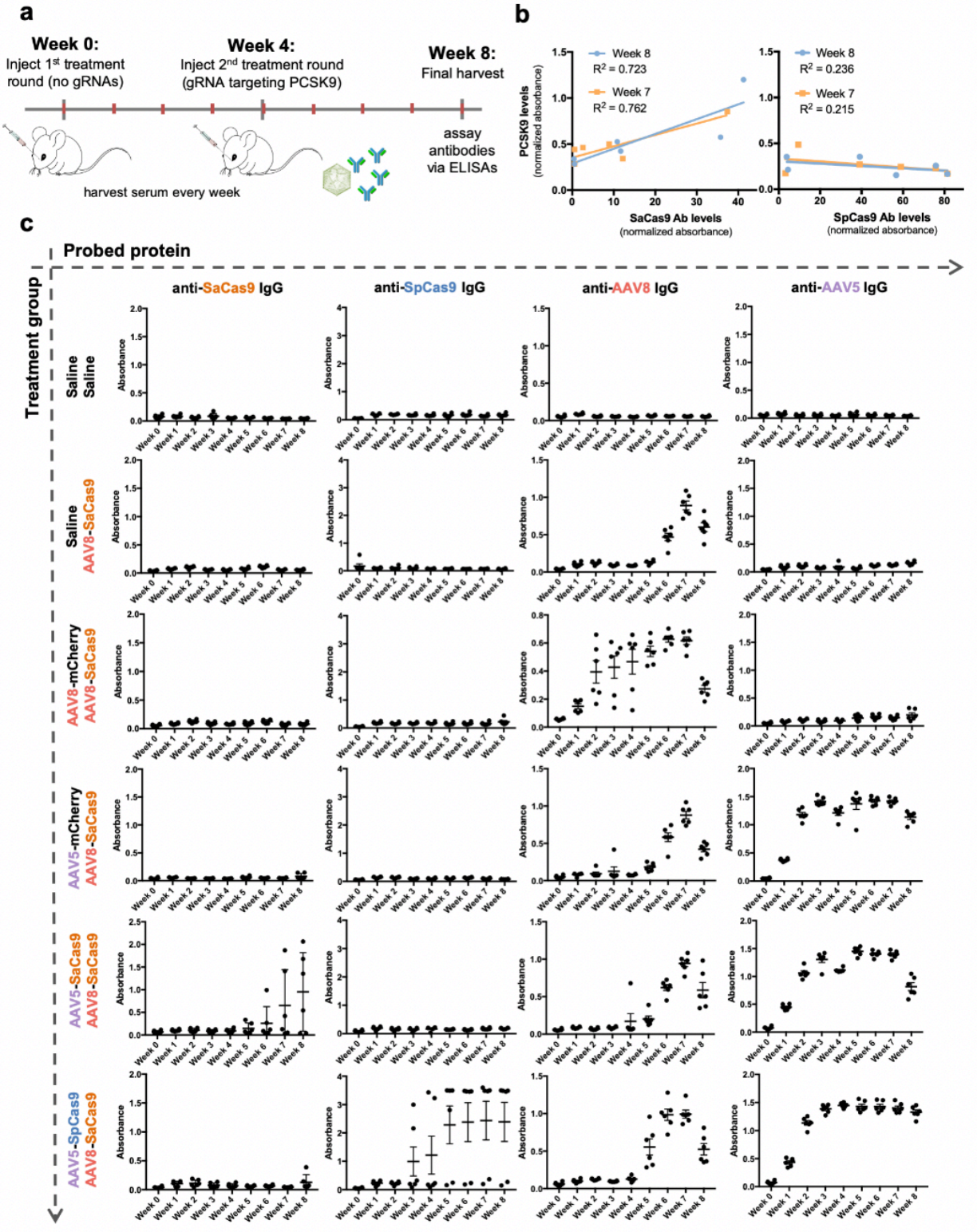


Figure 3.6: Major human AAV serotype groups. (a) AAV immune orthogonal cliques over 81 HLA alleles. AAV5 is the most immune-divergent in comparison to the other serotypes. No orthogonal cliques exist. (b) AAV phylogeny showing major serotype groupings as well as the position of the reconstructed sequence Anc80L65.

3.1.4 Overcoming immune barriers to effective gene editing

Having demonstrated that AAV-Cas9s elicit an immune response in the mouse model, and that the humoral responses to SpCas9 and SaCas9 do not cross-react, we next performed a two-step dosing experiment to test whether these immune responses inhibit the efficacy of multi-dose gene editing, and whether using immune-orthogonal orthologs in sequence can avoid immune-mediated inhibition of gene editing **(Figure 2.4a, 3.7)**. For this experiment, we used another mouse strain, BALB/c, in order to verify and characterize the immune response in two independent strains. The first round of dosing contained no gRNA and served to immunize the mice against the second dose, which contained an active AAV8-SaCas9 with gRNA targeting PCSK9, allowing us to directly measure genome editing efficiency by sequencing, as well as serum PCSK9 levels as a phenotypic readout for therapeutic efficacy. Additionally, we measured IgG responses to all AAV and Cas9 used in the experiment. As expected from previous preclinical work on AAV therapies, prior exposure and humoral immunity to AAV8 (AAV8-mCherry) abolished the effectiveness of subsequent gene editing when using AAV8 as the delivery vector (AAV8-SaCas9). Importantly, this effect was not seen with previous exposure to AAV5 (AAV5-mCherry), and subsequent dosing with AAV8-SaCas9 resulted in strong genome editing and PCSK9 knockdown similar to the effects of AAV8-SaCas9 dosing in naïve mice **(Figures 2.4b, 2.4c)**.

Figure 3.7: Time course of multiple dosing with immune orthogonal orthologues. (a) Mice were initially immunized (via retro-orbital injections of saline or AAV-CRISPRs at 1x10 12 vg/mouse) with saline, AAV8-mCherry, AAV5-mCherry, AAV5-SaCas9, or AAV5-SpCas9 with no gRNA. At 4 weeks, the mice were given a second dose of saline or AAV8-SaCas9 with a gRNA targeting PCSK9. Serum was harvested prior to the first injection, and weekly thereafter. (b) Per mouse SaCas9 and SpCas9 antibody levels were correlated with PCSK9 levels at weeks 7 and 8 to determine if mice mounting stronger immune responses had reduced editing (n=6 for SaCas9 and SpCas9 groups). PCSK9 knockdown significantly correlates with SaCas9, but not SpCas9 antibody levels (F-test for non-zero slope; SaCas9 week 8: p=0.032, SaCas9 week 7: p=0.023; SpCas9 week 8: p=0.329, SpCas9 week 7: p=0.354). (c) Time course of serum samples taken each week. Shown are ELISAs for antibodies specific to SpCas9, SaCas9, AAV8, and AAV5. For all panels, results are shown as mean ± s.e.m. Each data point represents an individual mouse, with n=6 mice per group.


a
Week 0:
Inject 1st treatment round (no gRNAs)
Week 4:
Inject 2nd treatment round (gRNA targeting PCSK9)
Week 8:
Final harvest
assay antibodies via ELISAs
harvest serum every week
b
Week 8
R2 = 0.723
Week 7
R2 = 0.762
PCSK9 levels (normalized absorbance)
SaCas9 Ab levels (normalized absorbance)
Week 8
R2 = 0.236
Week 7
R2 = 0.215
SpCas9 Ab levels (normalized absorbance)
c
Probed protein
Treatment group
anti-SaCas9 IgG
anti-SpCas9 IgG
anti-AAV8 IgG
anti-AAV5 IgG
Saline
Saline
Saline
AAV8-SaCas9
AAV8-mCherry
AAV8-SaCas9
AAV5-mCherry
AAV8-SaCas9
AAV5-SaCas9
AAV8-SaCas9
AAV5-SpCas9
AAV8-SaCas9
Absorbance
Week 0
Week 1
Week 2
Week 3
Week 4
Week 5
Week 6
Week 7
Week 8

Although we do not necessarily expect this observed orthogonality between AAV8 and AAV5 to carry over into the human setting, here it allowed us to specifically test the effects of the immune response to the Cas9 payload with minimal interference from the AAV delivery vector. Mice first immunized against SaCas9 using AAV5 showed a 35% reduced level of editing, a 38% reduction in PCSK9 decrease, and a wider variation between mice. This may reflect a weak immune response to SaCas9 in our mouse model, and/or a domination of private (individual) T-cell responses to SaCas9. IgG ELISAs revealed that only a subset of mice immunized with AAV5-SaCas9 developed an antibody response. We correlated the level of serum antibodies induced upon SaCas9 immunization with the efficiency of PCSK9 editing after the second dose, finding that mice with a stronger antibody response showed lower editing efficiency (**Figure 3.7b**). In contrast, we found that mice first dosed with AAV5-SpCas9 showed robust editing similar to that in naïve mice, suggesting that the predicted and measured immune orthogonality of Sp- and SaCas9 can be harnessed to circumvent immune barriers to gene editing.

To replicate these results in a different context, and to verify that immunity to Cas9 specifically can create a barrier to effective gene therapy, we conducted a slightly modified immunization experiment. Here we used a Cas9 protein vaccine combined in emulsion with Complete Freund's Adjuvant (CFA) as the initial dose, thereby immunizing a Cas9-specific primary response independently of AAV (**Figure 2.4e**). Subsequent dosing with AAV8-SaCas9 targeting PCSK9 recapitulated the results of AAV-based immunization, showing that prior exposure to the SaCas9 protein, but not SpCas9, significantly reduced the effectiveness of SaCas9-based gene editing by 42% and PCSK9 reduction by 51% (**Figure 2.4g, 2.4g**).

Taken together, anti-AAV and anti-Cas9 immunity represents a significant obstacle to therapeutic efficacy, and the use of immune orthogonal AAVs and Cas proteins, by bypassing

immune recall, enables effective gene-editing from repeated administrations of these therapeutic modalities.

3.2 DISCUSSION

To elude the host immune response, we developed here a framework to compare protein orthologs and their predicted binding to MHC I and MHC II by checking a sliding window of all k-mers in a protein for their presence in another, focusing on peptides predicted to bind to at least one MHC allele. Through this analysis, we identified cliques of Cas9 proteins that are immune orthogonal. Based on these predictions, specific T-cell responses from one ortholog would not cross-react with another ortholog of the same clique, preventing the re-activation of CD8+ cytotoxic T-cells, as well as the CD4+ T-cell help necessary to re-activate memory B-cells. We confirmed these results through ELISAs and verified three well-characterized Cas9 proteins (SpCas9, SaCas9, and CjCas9) to be immune orthogonal. Finally, we demonstrated in multiple contexts that consecutive dosing with the same AAV or Cas9 ortholog can face diminished editing efficacy which can be overcome with the use of immune orthogonal orthologs. Therefore, we expect that proteins belonging to the same clique can be used sequentially without eliciting memory T- and B- cell responses.

An important caveat is that each sequential ortholog should also be immune orthogonal to the pre-existing immune repertoire. Very recent work has begun to explore pre-existing immunity to Sp- and SaCas9[278,288,327] in human donors. One potential repository of Cas9 to which humans may not have any pre-existing immunity is in the genomes of extremophiles. However, although humans are not likely to be exposed to these organisms previously, their Cas9s may nevertheless be closely related to commensal or pathogenic species, and therefore immune orthogonality to pre-existing immunity must be rigorously evaluated. To explore this issue, we categorized 240 Cas9s

orthologs based on their species of origin as commensal, pathogenic, environmental, or extremophile, and compared the immune orthogonality between these groups (**Figure S9**). Preliminary analysis suggests that there may be extremophile Cas9s divergent enough as to be orthogonal to pre-existing immunity, even when taking into account both pathogens and commensals. Many more candidates are likely to be discovered as we continue cataloging microbial diversity in a variety of environments using metagenomic approaches. Alongside this process, more diverse Cas9s must be tested and studied to determine if and under what conditions they will be usable in a mammalian setting.

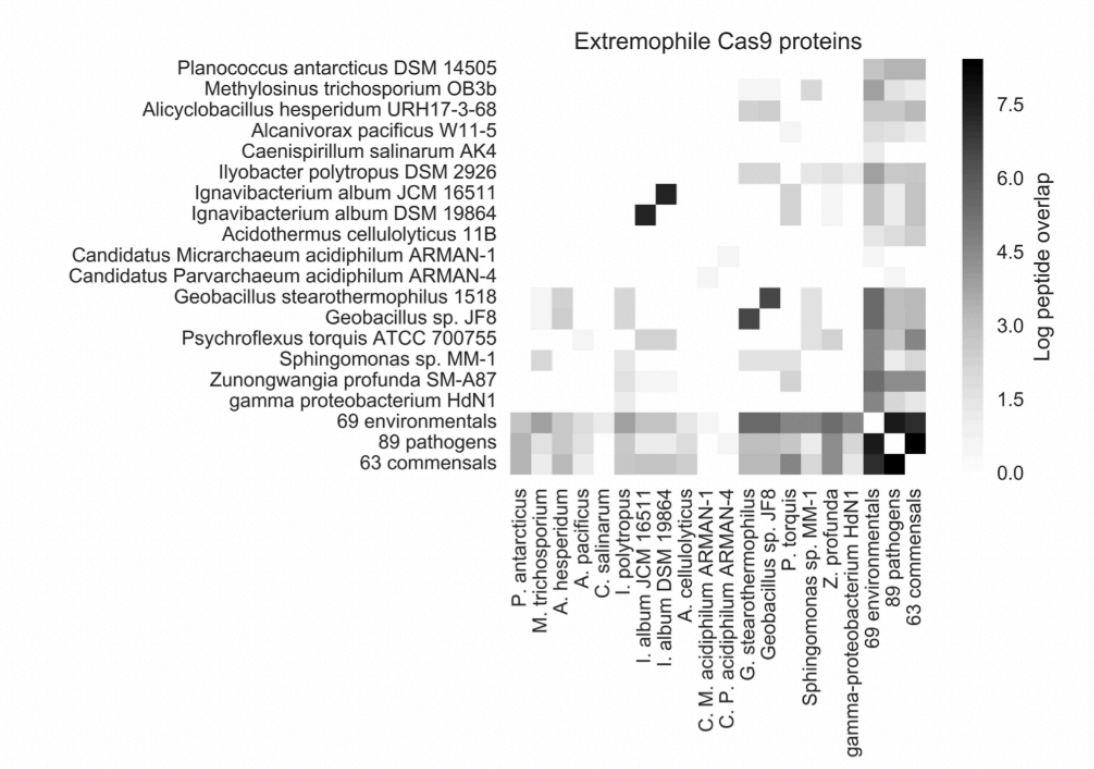


Figure 3.8: Exploring pre-existing CRISPR immunity. Cas9 orthologues were grouped into pathogens, commensals, environmentals, and extremophiles based on their species of origin (see species classification methods). Extremophile orthologues were assayed against pools of all 9-mer peptides originating from Cas9s from pathogenic, commensal, and environmental species. A few orthologues, including some highly divergent sequences from archaeal species, show near complete 9-mer orthogonality to large contingents of possible pre-existing Cas9 immunity.

Due to the importance of AAVs as a delivery agent in gene therapy, we also analyzed AAV serotypes through our MHC I and II comparison framework and have demonstrated that no two AAVs are predicted to be entirely immune orthogonal. However, with a known HLA genotype, it may be possible to define a personalized regimen of immune orthogonal AAVs using currently defined serotypes. For instance, use of AAV5 minimizes immune cross-reactivity in mice and non-human primates, as demonstrated by a recent study in which chimeric-AAV5 immunized mice and non-human primates successfully received a second dose of treatment with AAV1[333]. However, in the human setting we predict that there may be substantially more immune overlap between AAV5 and other AAVs. Additionally, it has been shown that memory B-cells heavily contribute to the antibody response to similar but not identical antigens[334], indicating that partial orthogonality may not be sufficient. Our analysis suggests that creating a pair of globally orthogonal AAV capsids for human application would require >10 mutations in one of the two proteins. This hypothetical orthogonal AAV capsid presents a substantial engineering challenge, as it requires mutating many of the most conserved regions to achieve immune orthogonality.

Previous work has identified that MHC affinity is highly dependent on anchor residues at either end of the binding pocket[335]. Residue diversity is more tolerated in the center of the binding pocket, though it may be these residues that most impact antigen specificity, as it is thought that they are central to interaction with the T-cell receptor (TCR). Comparing the number of orthologous pairs in 9-mer space with the number of predicted orthologous pairs based on class II binding predictions suggests that only approximately 65% of 9-mer peptides serve as appropriate MHC class II binding cores, even across the thousands of HLA-2 combinations we explore here. This under-sampling of peptide space by MHC molecules likely reflects the requirement for hydrophobic anchor residues and leaves some space for protein de-immunization by mutation of

immunogenic peptides to ones which never serve as MHC binding cores. Achieving this while preserving protein function, however, has proven difficult even for few HLA alleles, and remains a major protein engineering challenge. New technologies for directly measuring TCR affinity with MHC-presented antigens[336] will also further clarify the key antigenic peptides contributing to the immune response, and will be useful to inform approaches here.

On a tangential note, recent work has indicated that MHC class I peptides may have significant contribution from spliced host and pathogen-derived peptides created by proteasomal processing[337]. It is unclear how this may affect cross-recognition of proteins we predict to be immune orthogonal. On the one hand, it provides a mechanism whereby very short antigenic sequences spliced to the same host protein may result in cross-recognition of substantially different foreign antigens, however, we expect this to be unlikely due to the massive number of possible spliced peptides between the antigen and entire host proteome.

Overall, we believe our framework provides a potential solution for efficacious gene therapy, not solely for Cas9-mediated genome engineering, but also for other protein therapeutics that might necessitate repetitive treatments. Although using this approach still requires mitigating the primary immune response, particularly antibody neutralization and CTL clearance, we expect that epitope deletion and low-immunogenicity delivery vectors such as AAVs will mitigate this problem, and the potential for repeated dosage will reduce the need for very high first-dose titers and efficiency.

3.3 METHODS

3.3.1 Computational Methods

3.3.1.1 k-mer Analyses

For Cas9, we initially chose 91 orthologs cited in exploratory studies cataloging the diversity of the Cas9 protein[322,326,338–341], including several that are experimentally well-characterized. We subsequently expanded our analysis to a total of 240 Cas9 orthologs and 44 Cpf1/Cas12a orthologs for DNA-targeting CRISPR-associated effector proteins, and 84 RNA-targeting CRISPR-associated effectors including Cas13a, b and c. For AAVs, we analyzed 167 sequences, focusing in on all 13 characterized human serotypes, as well as one isolate from rhesus macaque (rh32), one engineered variant (DJ), and one reconstructed ancestral protein (Anc80L65). We then compared total sequence similarity (immunologically uninformed) as well as predicted binding to class I and class II MHC molecules (immunologically informed) between these proteins. Immunologically uninformed sequence comparison was carried out by checking a sliding window of all contiguous k-mers in a protein for their presence in another protein sequence with either zero or one mismatch.

3.3.1.2 MHC Binding Predictions

Immunologically informed comparison was done in a similar fashion, but using only those k-mers predicted to bind to at least one of 81 HLA-1 alleles using netMHC 4.0[294] for class I (alleles can be found at http://www.cbs.dtu.dk/services/NetMHC/MHC_allele_names.txt), and at least one of 5,620 possible MHC II molecules based on 936 HLA-2 alleles using netMHCIIpan 3.1[295] for class II (alleles can be found at http://www.cbs.dtu.dk/services/NetMHCIIpan-3.1/alleles_name.list). We compared the use of netMHC to alternative immune epitope prediction platforms such as the Immune Epitope Database (iedb.org)[296] and found very strong agreement across software. Ultimately, we chose netMHC because of the larger number of HLA alleles it supports. Sequences were defined as binding if the predicted affinity ranked in the top 2% of a test library of 400,000 random peptides as suggested in the software guidelines. Generation of immune

orthogonal cliques was carried out using the Bron-Kerbosch algorithm. Briefly, a graph was constructed with each ortholog as a vertex, where the edges are defined by the number of shared immunogenic peptides between the connecting vertices. Sets of proteins for which every pair in the set is immune orthogonal constitutes a clique.

3.3.1.3 Phylogenetics and Species Classification

For phylogenetic analyses, protein sequences were aligned using MUSCLE, and distance was calculated using the BLOSUM 62 matrix excluding indels. Phylogeny of AAV serotypes was created using neighbor-joining on major serotype sequences. Categorization of Cas9 orthologs into commensal, pathogenic, environmental, and extremophile species of origin was done by assessing the source of the sample sequence. Sequences isolated from species which had been observed in human-associated samples were classified as pathogenic if they had been reported to cause disease (this would include species which are normally commensal, but opportunistically pathogenic), and commensal otherwise. Sequences from species which are not known to be associated with the human microbiome were classified as environmental unless the species was uniquely isolated from extreme environments including but not limited to geothermal vents, deep anoxic groundwater, fossil fuel deposits, and polar ice.

3.3.2 Experimental Methods

3.3.2.1 Vector design and construction

Split-SpCas9 AAV vectors were constructed by sequential assembly of corresponding gene blocks (IDT) into a custom synthesized rAAV2 vector backbone[297,298]. The first AAV contains a gRNA driven by a human RNA polymerase III promoter, U6, and a N-terminal Cas9 (NCas9)

fused to an N-intein driven by a CMV promoter, as well as a polyadenylation (polyA) signal (Truong et al., Moreno et al.) The second AAV cassette contains a CMV driven C-terminal Cas9 (CCas9) fused to a C-intein as well as a polyA signal. gRNA sequences were inserted into NCas9 plasmids by cloning oligonucleotides (IDT) encoding spacers into AgeI cloning sites via Gibson assembly. pX601-AAV-CMV::NLS-SaCas9-NLS-3xHA-bGHpA;U6::BsaI-sgRNA was a gift from Feng Zhang (Addgene plasmid # 61591)

3.3.2.2 AAV Production

AAV2/8, AAV2/2, AAV2/5, AAV2/DJ virus particles were produced using HEK293T cells via the triple transfection method and purified via an iodixanol gradient[299]. Confluency at transfection was between 80% and 90%. Media was replaced with pre-warmed media 2 hours before transfection. Each virus was produced in 5 x 15 cm plates, where each plate was transfected with 7.5 µg of pXR-capsid (pXR-8, pXR-2, pXR-5, pXR-DJ), 7.5 of µg recombinant transfer vector, and 22.5 µg of pAd5 helper vector using PEI (1µg/µl linear PEI in 1x DPBS pH 4.5, using HCl) at a PEI:DNA mass ratio of 4:1. The mixture was incubated for 10 minutes at RT and then applied dropwise onto the media. The virus was harvested after 72 hours and purified using an iodixanol density gradient ultracentrifugation method. The virus was then dialyzed with 1x PBS (pH 7.2) supplemented with 50 mM NaCl and 0.0001% of Pluronic F68 (Thermo Fisher) using 100kDA filters (Millipore), to a final volume of ~1 ml and quantified by qPCR using primers specific to the ITR region, against a standard (ATCC VR-1616).

AAV-ITR-F: 5'-CGGCCTCAGTGAGCGA-3' and

AAV-ITR-R: 5'-GGAACCCCTAGTGATGGAGTT-3'.

3.3.2.3 Animal studies

All animal procedures were performed in accordance with protocols approved by the Institutional Animal Care and Use Committee (IACUC) of the University of California, San Diego. All mice were acquired from Jackson labs. AAV injections were done in adult male C57BL/6J or BALB/c mice (10 weeks) through retro-orbital injections using $1x10^{12}$ vg/mouse.

3.3.2.4 CFA immunizations

CFA immunizations were prepared by mixing CFA (Sigma-Aldrich) with 5 μg Cas9 protein or PBS at a 1:1 ratio using two syringes connected by an elbow joint to create an even emulsion. 200 μL of CFA emulsion was injected subcutaneously into the flanks of adult mice.

3.3.2.5 ELISA

PCSK9: Levels of serum PCSK9 were measured using the Mouse Proprotein Convertase 9/PCSK9 Quantikine ELISA kit (R&D Systems) according to manufacturer's guidelines. Briefly, serum samples were diluted 1:200 in Calibrator diluent and allowed to bind for 2 h onto microplate wells that were precoated with the capture antibody. Samples were then sequentially incubated with PCSK9 conjugate followed by the PCSK9 substrate solution with extensive intermittent washes between each step. The amount of PCSK9 in serum was estimated colorimetrically using a standard microplate reader (BioRad iMark).

Cas9 and AAV: Recombinant SpCas9 protein (PNA Bio, cat. no. CP01), or SaCas9 protein (ABM good, cat no. K144), was diluted in 1x coating buffer (Bethyl), and 0.5 μg was used to coat each well of 96-well Nunc MaxiSorp Plates (ab210903) overnight at 4 °C. For AAV experiments, 10^9 vg of AAV-2, -5, -8 or -DJ in 1x coating buffer was used to coat each well of 96-well Nuc MaxiSorp Plates. Plates were washed three times for 5 min with 350 μl of 1x Wash Buffer (Bethyl) and blocked with 300 μl of 1x BSA Blocking Solution (Bethyl) for 2 h at RT. The wash procedure

was repeated. Serum samples were added at 1:40 dilution, and plates were incubated for 5 h at 4 °C with shaking. Wells were washed three times for 5 min, and 100 µl of HRP-labeled goat anti-mouse IgG1 (Bethyl; diluted 1:100,000 in 1% BSA Blocking Solution) was added to each well. After incubating for 1hr at RT, wells were washed four times for 5 min, and 100 µl of TMB Substrate (Bethyl) was added to each well. Optical density (OD) at 450 nm was measured using a plate reader (BioRad iMark).

3.3.2.6 NGS quantification of editing

Genomic DNA was extracted from samples of mouse livers using a DNA extraction kit (Qiagen). A 200 bp region containing the target cut site of the PCSK9 gene was amplified by PCR using 0.5 µg DNA (~100,000 diploid genomes) as the template. Libraries were prepared using NEBNext Illumina library preparation kit and sequenced on an Illumina HiSeq. Each sample was sequenced to a target depth of 100,000 reads. Adaptors were trimmed from resulting fastqs using AdapterRemoval[300] and NHEJ-repaired cleavage events resulting in a mutation were quantified using CRISPResso[301].

3.3.2.8 Statistics

PCSK9 ELISA data from immunization experiments (Figures 3, S7), were normalized per mouse to the average of the first 4 weeks of the experiment (during which time no active dose was given), and then analyzed using a two-way repeated measures ANOVA to account for both time and group membership as independent variables. Post hoc Tukey tests were used to compare across groups at each timepoint as shown in Figure 3C.

3.3.2.9 Epitope prediction and peptide synthesis

The MHC-binding peptides for our mouse model were predicted using the netMHC-4.0 and netMHCIIpan-3.1 online software with the alleles H-2-Db and H-2-Kb for class I and H-2-IAb for class II. For MHCII, the top 10 peptides for Sp- and SaCas9 and top 5 peptides for AAV-8 and AAV-DJ by percentile binding were selected for synthesis by Synthetic Biomolecules as crude materials. For MHCI, we selected the top 20 peptides for Sp- and SaCas9 and the top 10 for AAV-8 and AAV-DJ. All peptides were dissolved in DMSO with a concentration of 40 mg ml^{-1} and stored at −20 °C.

3.4 ACKNOWLEDGEMENTS

Chapter 3, in part, is a reprint of the material as it appears in Nature Biomedical Engineering 3, 806–816. (2019). Moreno, Ana; Palmer, Nathan; Alemán, Fernando; Chen, Genghao; Pla, Andrew; Jiang, Ning; Chew, Wei Leong; Law, Mansun; Mali, Prashant. The dissertation author was the primary researcher and author of this paper. We thank members of the Mali laboratory for advice and help with experiments and the Salk GT3 viral core for help with the production of AAVs. This research was supported by UCSD Institutional Funds, the Burroughs Wellcome Fund (1013926), the March of Dimes Foundation (5-FY15-450), the Kimmel Foundation (SKF-16-150), and NIH grants (R01HG009285, RO1CA222826, RO1GM123313, R01AI079031 and R01AI106005). A.M.M. acknowledges a graduate fellowship from CONACYT and UCMEXUS. W.L.C. acknowledges the IAF-PP grant (H17/01/a0/012).

CHAPTER 4 – APPLYING LORAX TO PROGRESSIVELY DEIMMUNIZE CAS9

Immunogenicity is a major concern for protein-based therapeutics, particularly when using proteins derived from non-human species. Induction of the immune response can render treatments ineffective and cause serious, even life-threatening side effects[342,343] One strategy to overcome immune induction by protein therapeutics is to mutate particularly immunogenic epitopes in the target[344,345]. However, this strategy is hindered by the ability of the adaptive immune system to recognize multiple epitopes across large regions of the antigen. While epitope deletion efforts to date have demonstrated reduced immunogenicity, their scope has been limited to only a few major T- and B-cell epitopes[241]. As of yet it is not possible to make these studies comprehensive due to 1) the vast possible epitope space, and 2) protein engineering methods with limited throughput.

Variant library screening has proven to be an effective approach to protein engineering but applying it in this case faces several technical challenges. One problem that arises when attempting to create a de-immunized protein is the need to mutate multiple sites simultaneously across the full length of the protein, as immune epitopes are often widely distributed in the protein sequence. Mutating multiple sites simultaneously requires fully degenerate combinatorial libraries that can very quickly balloon to unmanageable numbers of variants. Narrowing down this space by intelligent selection of library members is necessary to define a tractable mutational landscape to explore and critical for maximizing the chance of functional hits. Another problem is that reading out combinatorial mutations scattered across large (>1kb) regions of the protein is extremely difficult using typical short read sequencing platforms. The best current technique to get around this issue has been to read out only short barcode sequences attached to each variant in order to genotype libraries post-screen. This method has proven effective in numerous cases but is limited by the difficulty of constructing large combinatorial libraries in which each member has a short,

unique barcode whose relation to the overall sequence is known. To date, these issues have generally limited combinatorial library screens to short regions able to be sequenced directly.

To overcome these challenges and to allow for the generation of heavily mutagenized proteins we have developed several methodological innovations which, taken together, comprise a protein engineering platform capable of screening millions of combinatorial variants simultaneously with mutations spread across the full length of arbitrarily large proteins, with computation-guided mutation design to maximize the probability of exploring functional mutation space. In addition to these advantages, our platform can be applied iteratively to tackle particularly challenging engineering tasks by exploring huge swaths of combinatorial mutation space unapproachable using previous techniques. Furthermore, while this methodology is particularly suited to the unique challenges of protein de-immunization, which we demonstrate here as our proof-of-concept use case, it is also applicable to any potential protein engineering goal, so long as there exists an appropriate screening procedure to select for the desired protein functionality. The specific innovations which enable this technology are threefold: library design, library construction, and sequencing readout linking genotype to phenotype.

The unmatched throughput and accuracy of short-read sequencing has revolutionized the technical possibilities in biology. However, it does have its limitations, which come into play when attempting long-range protein engineering. Instead, we apply long-read nanopore sequencing to measure the results of the screens of our combinatorial libraries. This circumvents the limit of short target regions and obviates the need for barcodes altogether by single-molecule sequencing of the entire target gene, enabling library design strategies which can explore any region of the protein in combination with any other region without any complicated cloning procedures required to facilitate barcoding.

To date, the adoption of nanopore sequencing has been limited by its high error rate, around 95% accuracy per DNA base, as compared to established short read techniques which are multiple orders of magnitude more accurate. To address this challenge, we design our libraries such that each variant we engineer will have multiple nucleotide changes for each single target amino acid change, effectively increasing the sensitivity of nanopore based readouts exponentially with increasing numbers of nucleotide changes per library member. The large majority of amino acid substitutions are amenable to a library design paradigm in which each substitution is encoded by two, rather than one, nucleotide changes, due to the degeneracy of the genetic code and the highly permissive third "wobble" position of codons. For example, if the wild-type amino acid leucine is encoded by the codon CTG, typically a substitution to the amino acid proline would be encoded by the single nucleotide change T to C at the second position, resulting in a CCG codon. However, it is also possible to use any of the other three codons encoding proline, CCT, CCC, and CCA, each of which is two nucleotide changes away from the wild-type sequence. These changes are much easier to reliably detect with error-prone long-read nanopore sequencing.

In order to narrow down the vast mutational space associated with combinatorial libraries, we utilize an approach guided by evolution and natural variation. As de-immunizing protein engineering seeks to alter the amino acid sequence of a protein without disrupting functionality, it is extremely useful to narrow down mutations to those less likely to result in non-functional variants. One way to identify these mutants is to leverage the large amounts of sequencing data available to identify low-frequency SNPs that have been observed in natural environments. Such variants are likely to have limited effect on protein function, as highly deleterious alleles would likely be quickly selected out of natural populations and therefore not appear in sequencing data.

Using these likely neutral amino acid substitutions in combinatorial libraries should substantially increase the likelihood of functional hits with enough epitope variation to evade immune induction.

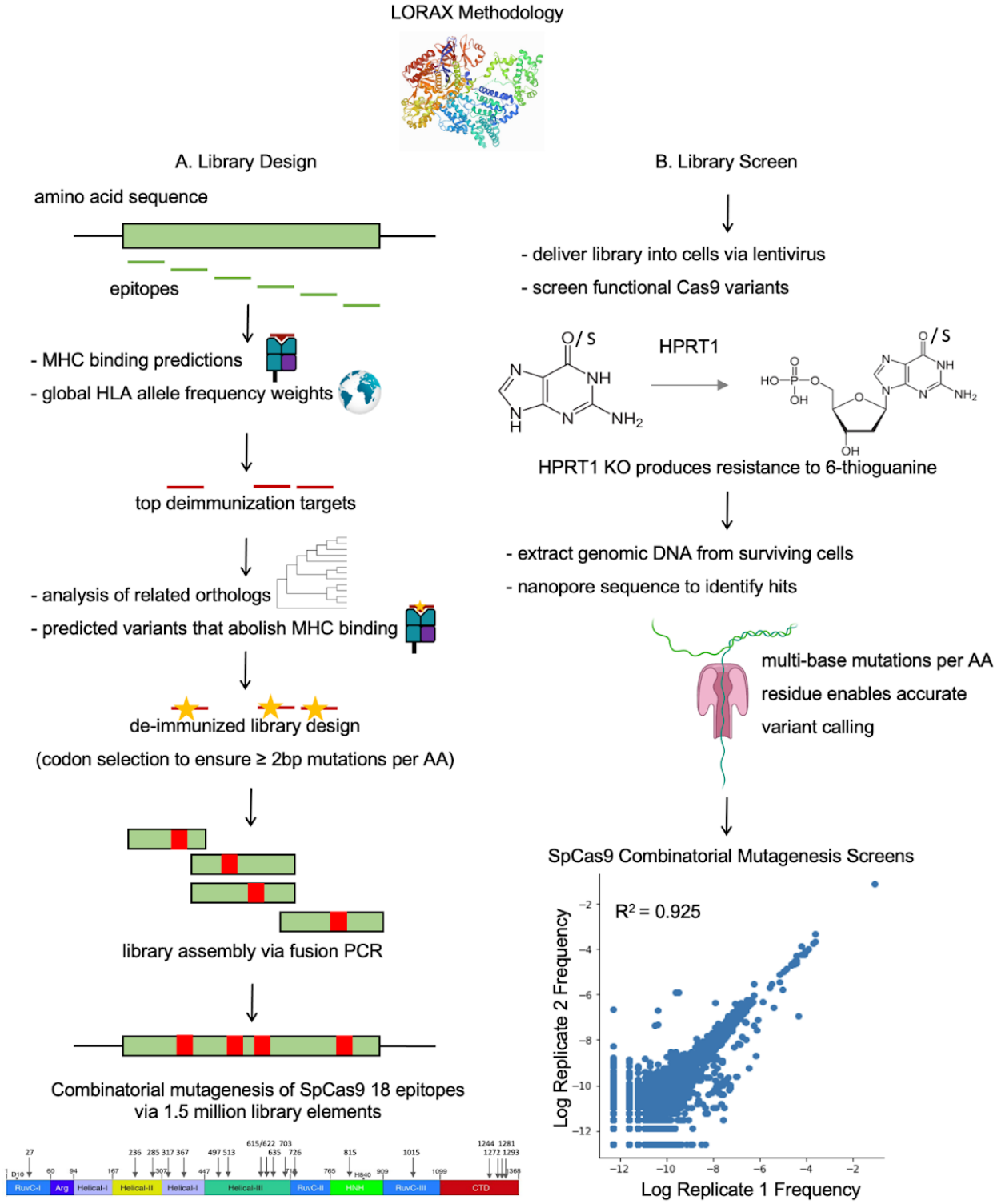
An orthogonal, though complementary approach which we also use here is to screen potential mutations for immunogenicity *in silico* using the netMHC epitope prediction software[346]. To this end, we screen both all possible mutations, as well as only those found in our phylogenetic analysis described above in our top immunogenic epitopes, i.e., our target mutation sites, in order to determine to what degree the mutations we engineer are likely to result in the de-immunization of that particular epitope. This is a critical step as some mutations will have little effect on overall immunogenicity both in terms of MHC binding and B-cell / T-cell receptor interactions.

4.1 RESULTS

4.1.1 LORAX

While variant library screening has proven to be an effective approach to protein engineering[347–354], applying it to deimmunization faces three important technical challenges. One, the need to mutate multiple sites simultaneously across the full length of the protein; two, reading out the associated combinatorial mutations scattered across large (>1kb) regions of the protein via typical short-read sequencing platforms; and three, engineering fully degenerate combinatorial libraries which can very quickly balloon to unmanageable numbers of variants[355,356]. To overcome these challenges, we developed several methodological innovations which, taken together, comprise a novel **lo**ng **ra**nge multiple**x**ed (LORAX) protein engineering platform capable of screening millions of combinatorial variants simultaneously with mutations spread across the full length of arbitrarily large proteins (**Figure 4.1**).

Figure 4.1: LORAX protein engineering methodology to screen progressively deimmunized Cas9 variants. (Left panel: library design) Low-frequency SNPs that have limited effect on Cas9 function were identified and immunogenicity was evaluated *in silico* using the netMHC epitope prediction software to identify candidate mutations. This analysis was performed for many Cas9 orthologs. Mutations were generated such that two bps were changed to account for Nanopore sequencing accuracy. A library was then generated by fusion PCR of blocks containing wildtype and mutations at specific epitopes. Location of epitopes in SpCas9 that were combinatorially mutated and screened is shown. (Right panel: library screen) The screen was performed by transducing HeLa cells with a lentiviral library containing the Cas9 variants and a guide that cuts the HPRT1 gene. HPRT1 knockout produces resistance to 6-TG. After two weeks, DNA is extracted from surviving cells, Cas9 variant sequences are PCR amplified from the genomic DNA and Nanopore sequenced. High accuracy of variant identification is possible due to the use of two bp mutations for each amino acid change. Post-screen library element frequencies across two independent replicates are shown. Replicate correlation was calculated excluding the over-represented wild-type sequence.


LORAX Methodology
A. Library Design
amino acid sequence
epitopes
- MHC binding predictions
- global HLA allele frequency weights
top deimmunization targets
- analysis of related orthologs
- predicted variants that abolish MHC binding
de-immunized library design
(codon selection to ensure ≥ 2bp mutations per AA)
library assembly via fusion PCR
Combinatorial mutagenesis of SpCas9 18 epitopes
via 1.5 million library elements
RuvC-I
Arg
Helical-I
Helical-II
Helical-I
Helical-III
RuvC-II
HNH
RuvC-III
CTD
B. Library Screen
- deliver library into cells via lentivirus
- screen functional Cas9 variants
HPRT1
HPRT1 KO produces resistance to 6-thioguanine
- extract genomic DNA from surviving cells
- nanopore sequence to identify hits
multi-base mutations per AA
residue enables accurate
variant calling
SpCas9 Combinatorial Mutagenesis Screens
R² = 0.925
Log Replicate 2 Frequency
Log Replicate 1 Frequency

4.1.2 Library design

To narrow down the vast mutational space associated with combinatorial libraries, we utilize an approach guided by evolution and natural variation[357,358]. As deimmunizing protein engineering seeks to alter the amino acid sequence of a protein without disrupting functionality, it is extremely useful to narrow down mutations to those less likely to result in non-functional variants. To identify these mutants, we generated large alignments of Cas9 orthologs from publicly available data to identify low-frequency SNPs that have been observed in natural environments. Such variants are likely to have limited effect on protein function, as highly deleterious alleles would tend to be quickly selected out of natural populations (if Cas9 activity is under purifying selection) and therefore not appear in sequencing data[359]. To further subset these candidate mutations, we evaluated for immunogenicity *in silico* using the netMHC epitope prediction software[346], in order to determine to what degree the candidate mutations are likely to result in the deimmunization of the most immunogenic epitopes in which they appear. This is a critical step as many mutations may have little effect on overall immunogenicity[344,360]. Screening for decreased peptide-MHC class I binding filters out amino acid substitutions which are likely immune-neutral, substantially increasing the likelihood of functional hits with enough epitope variation to evade immune induction[344,345].

4.1.3 Readout

Next, to enable readout, we applied long-read nanopore sequencing to measure the results of the screens of our combinatorial libraries. This circumvents the limit of short target regions and obviates the need for barcodes altogether by single-molecule sequencing of the entire target gene, enabling library design strategies which can explore any region of the protein in combination with

any other region without any complicated cloning procedures required to facilitate barcoding[361]. To date, the adoption of nanopore sequencing has been limited by its high error rate, around 95% accuracy per DNA base[362], as compared to established short read techniques which are multiple orders of magnitude more accurate. To address this challenge, we designed our libraries such that each variant we engineered would have multiple nucleotide changes for each single target amino acid change, effectively increasing the sensitivity of nanopore based readouts with increasing numbers of nucleotide changes per library member. The large majority of amino acid substitutions are amenable to a library design paradigm in which each substitution is encoded by two, rather than one, nucleotide changes, due to the degeneracy of the genetic code and the highly permissive third "wobble" position of codons.

4.1.4 Immunogenicity scoring

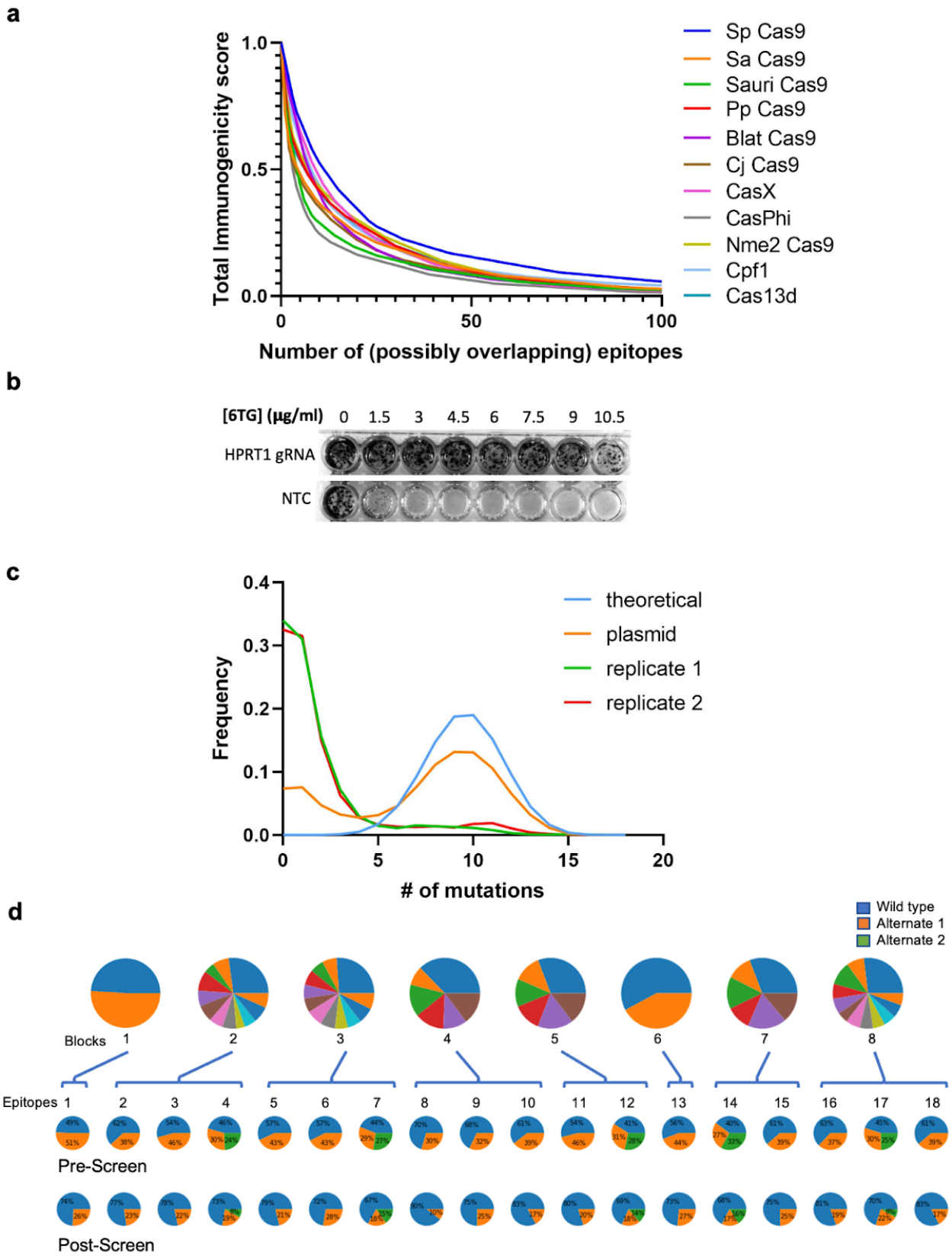
The scale of engineering which would be required to generate an effectively deimmunized Cas9 is not fully understood, as combinatorial deimmunization efforts at the scale of proteins thousands of amino acids long have not yet been possible. Therefore, to roughly estimate these parameters we developed an immunogenicity scoring metric which takes into account all epitopes across a protein and the known diversity of MHC variants in a species weighted by population frequency to generate a single combined score representing the average immunogenicity of a full-length protein as a function of each of its immunogenic epitopes[363]. Formally, this score is calculated as:

$$I_x = \frac{\sum_i^m \sum_j^n w_j(1 - \log(k_{ij} \times \hat{j}))}{y}$$

where I_x = Immunogenicity score of protein x, i = epitopes, j = HLA alleles, j = allele specific standardization coefficient, w_j = HLA allele weights, k_{ij} = predicted binding affinity of epitope i to

allele *j*, and *y* = protein specific scaling factor. We then predicted the overall effect of mutating the top epitopes in several Cas9 orthologs (**Figure 4.2a**). As might be expected, this analysis suggests that single-epitope strategies are woefully inadequate to deimmunize a whole protein for multiple HLA types, and that there are diminishing returns as more and more epitopes are deimmunized. Our analysis suggests that it may require on the order of tens of deimmunized epitopes to make a significant impact on overall, population-wide protein immunogenicity. The scale of engineering demanded by these immunological facts has previously been intractable, but by applying LORAX we conjectured one could now make substantial steps, several mutations at a time, through the mutational landscape of the Cas9 protein.

Figure 4.2: LORAX screen design and results. (a) Immunogenicity scores for Cas orthologs, demonstrating reduced immunogenicity (averaged across HLA types) as the number of mutated epitopes increases. (b) Presence of HPRT1 converts 6TG into a toxic nucleotide analog. HeLa cells transduced with wildtype Cas9 and either a HPRT1 targeting or nontargeting (NTC) guide. Only cells where the HPRT1 gene is disrupted can live in various concentrations of 6TG. 6 μg/mL 6TG was used for the screen as this concentration was sufficient for complete killing of NTC-bearing cells. (c) Variant Cas9 sequences were amplified from the plasmid library or genomic DNA post-screen. Long-read nanopore sequencing was performed and the mutational density distribution for the predicted library, the constructed Cas9 variant library, and the two replicates post-screen are plotted. (d) Cas9 block composition and pre- and post-screen allele frequencies at each of the 18 mutational sites. Each block and site shows enrichment of the wild-type allele, but all sites retain a substantial fraction of mutant alleles.


a
Total Immunogenicity score
1.0
0.5
0.0
0
50
100
Number of (possibly overlapping) epitopes
Sp Cas9
Sa Cas9
Sauri Cas9
Pp Cas9
Blat Cas9
Cj Cas9
CasX
CasPhi
Nme2 Cas9
Cpf1
Cas13d
b
[6TG] (µg/ml) 0 1.5 3 4.5 6 7.5 9 10.5
HPRT1 gRNA
NTC
c
Frequency
0.4
0.3
0.2
0.1
0.0
0
5
10
15
20
# of mutations
theoretical
plasmid
replicate 1
replicate 2
d
Wild type
Alternate 1
Alternate 2
Blocks 1 2 3 4 5 6 7 8
Epitopes 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Pre-Screen
Post-Screen

4.1.5 Library construction

Specifically, applying the procedure above, we designed a library of Cas9 variants based on the SpCas9 backbone containing 23 different mutations across 17 immunogenic epitopes (**Figure 4.1**). Combining these in all possible combinations yields a library of 1,492,992 unique elements. With this design, we then constructed the library in a stepwise process. First, the full-length gene was broken up into short blocks of no more than 1000 bp, which overlap by 30 bp on each end. Each block is designed such that it contains no more than 4 target epitopes to mutagenize. With few epitopes per block and few variant mutations per epitope, it becomes feasible to chemically synthesize each combination of mutations for each block. Each of these combinations was then synthesized and mixed at equal ratios to make a degenerate block mix. This was repeated for each of the blocks necessary to complete the full-length protein sequence. Oxford Nanopore (ONT) MinION sequencing confirmed the majority of the pre-screened library consists of Cas9 sequences with significant numbers of mutations, with most falling into a broad peak between 6 and 14 mutations per sequence, each of which knocking out a key immunogenic epitope (**Figure 4.2c**).

4.1.6 Screening

To identify functional variants still capable of editing DNA, we next designed and carried out a positive selection screen targeting the hypoxanthine phosphoribosyltransferase 1 (HPRT1) gene[364]. In the context of the screen, HPRT1 converts 6-thioguanine (6TG), an analogue of the DNA base guanine, into 6-thioguanine nucleotides that are cytotoxic to cells via incorporation into the DNA during S-phase[365]. Thus, only cells containing functional Cas9 variants capable of disrupting the HPRT1 gene can survive in 6TG-containing cell culture media. To first identify the

optimal 6TG concentration, HeLa cells were transduced with lentivirus particles containing wild-type Cas9 and either a HPRT1-targeting guide RNA (gRNA) or a non-targeting guide. After selection with puromycin, cells were treated with 6TG concentrations ranging from 0-14 μg/mL for one week. Cells were stained with crystal violet at the end of the experiment and imaged. 6 μg/mL was selected as all cells containing non-targeting guide had died while cells containing the HPRT1 guide remained viable (**Figure 4.2b**).

To perform the screen, replicate populations of HeLa cells were transduced with lentiviral particles containing the variant SpCas9 library along with the HPRT1-targeting gRNA at 0.3 MOI and at greater than 75-fold coverage of the library elements. Cells were selected using puromycin after two days and 6TG was added once cells reached 75% confluency. After two weeks, genomic DNA was extracted from remaining cells and full-length Cas9 amplicons were nanopore sequenced on the MinION platform.

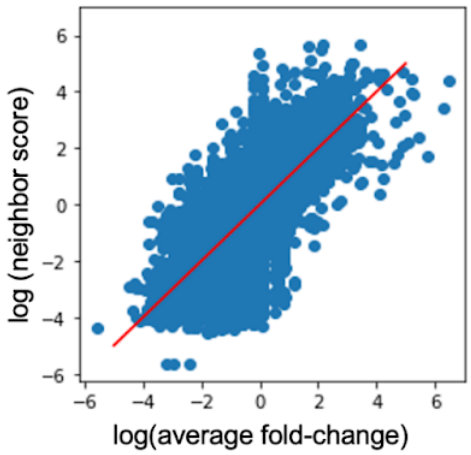
Sequencing revealed that the library was significantly shifted in the mutation density distribution, suggesting that the majority of the library with large (>4) numbers of mutations resulted in non-functional proteins which were unable to survive the screen. Meanwhile, wild-type, single, and double mutants were generally enriched as these proteins proved more likely to retain functionality and pass through the screen (**Figure 4.2c**). Additionally, the two independent replicates of the screen showed strong correlation (R^2 = 0.925) providing further evidence of robustness (**Figure 4.1**). We also analyzed the change in overall frequency of mutations in the pre- and post-screen libraries to see if a pattern of mutation effects could be inferred. Although the wild-type allele was enriched at every site in the post-screen sequences, nearly every site retained a significant fraction of mutated alleles, suggesting that the mutations, at least individually, are fairly well-tolerated and do not disrupt Cas9 functionality (**Figure 4.2d**).

4.1.7 Hit selection

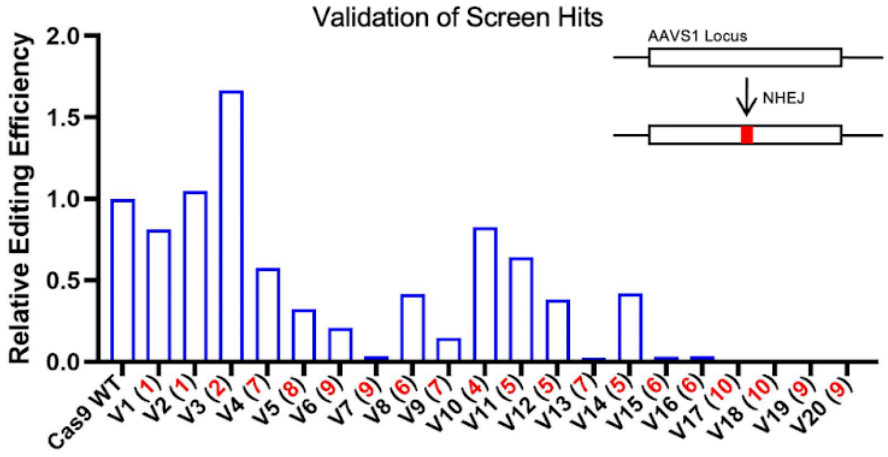
In order to select hits for downstream validation and analysis, we devised a method for differentiating high-support hits likely to be real from noise-driven false positive hits. To do this we hypothesized that the fitness landscape of the screen mutants is likely to be smooth, i.e. variants that contain similar mutations are more likely to have similar fitnesses in terms of editing efficiency compared to randomly selected pairs[366]. We confirmed this by computing a predicted screen score for each variant based on a weighted regression of its nearest neighbors in the screen. This metric correlates well with the actual screen scores and approaches the screen scores even more closely as read coverage increases. This provides good evidence that the fitness landscape is indeed somewhat smooth (**Figure 4.3a**). Next, we reasoned that because the fitness landscape is smooth, real hits should reside in broad fitness peaks which include many neighbors that also show high screen scores, whereas hits that are less supported by near neighbors are more likely to be spurious as they represent non-smooth fitness peaks. Formalizing this logic, we performed a network analysis to differentiate noise-driven hits from bona fide hits by looking at the degree of connectivity with other hits (**Figure 4.4a**).

Figure 4.3: Validations of LORAX screen identified Cas9 variants. (a) Correlation between the fold change of a Cas9 variant and its predicted fold-change based on a k-nearest neighbors regression. Neighboring variants are those that share similar mutational patterns. The strong correlation suggests a smooth fitness landscape in which variants with similar mutation patterns will be more similar in fitness, on average, than those with divergent mutation patterns. (b) Cas9 wildtype or variants V1-20 and sgRNA targeting the AAVS1 locus were introduced into HEK293T cells. NHEJ mediated editing at the AAVS1 locus was quantified via NGS for Cas9 WT and variants V1-20 is plotted. The number in parentheses represents the number of mutations in the variant. Variant genotypes are listed in the lower panel.

a



b



Variant Genotypes

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
V1	-	-	-	-	-	-	-	L615G	-	-	-	-	-	-	-	-	-	-
V2	-	-	-	-	-	-	-	-	L622Q	-	-	-	-	-	-	-	-	-
V3	-	-	-	-	-	-	L513T	-	L622Q	-	-	-	-	-	-	-	-	-
V4	-	-	Y285Q	-	-	-	L513T	-	L622Q	-	-	L726G	L815D	-	L1244G	-	L1281A	-
V5	-	-	Y285Q	-	S367C	-	L513T	-	L622Q	-	-	L726G	L815D	-	L1244G	-	L1281A	-
V6	P27L	-	Y285Q	-	S367C	-	L513T	-	L622Q	-	-	L726G	L815D	-	L1244G	-	L1281A	-
V7	-	-	Y285Q	S317H	S367C	-	L513T	-	L622Q	-	-	L726G	L815D	-	L1244G	-	L1281A	-
V8	-	L236C	-	-	S367C	F497T	-	L615G	-	-	-	-	-	Y1015G	-	-	-	Y1293Q
V9	-	L236C	-	-	S367C	F497T	-	L615G	-	-	-	-	-	Y1015G	-	I1272Q	-	Y1293Q
V10	-	-	-	S317H	-	F497T	L513G	-	L622Q	-	-	-	-	-	-	-	-	-
V11	-	-	-	-	-	F497T	-	-	L622Q	-	-	L726G	-	-	-	-	L1281A	Y1293Q
V12	-	-	-	-	-	-	-	-	L622Q	L635D	F703A	-	-	Y1015G	L1244G	-	-	-
V13	-	-	-	-	-	-	-	-	L622Q	L635D	F703A	-	L815D	-	L1244G	I1272Q	L1281E	-
V14	-	-	-	-	S367C	-	-	-	L622Q	-	-	L726P	-	Y1015G	-	I1272Q	-	-
V15	-	L236C	-	S317H	-	-	-	-	-	-	-	-	L815D	-	-	I1272Q	L1281E	Y1293Q
V16	-	-	-	S317C	-	-	L513T	-	-	L635D	F703A	-	-	Y1015G	-	-	-	Y1293Q
V17	-	-	-	S317H	S367C	-	-	-	-	-	F703A	L726G	L815D	Y1015G	L1244G	I1272Q	L1281A	Y1293Q
V18	P27L	-	-	S317H	S367C	F497T	-	L615G	-	-	F703A	-	-	Y1015K	-	I1272Q	L1281A	Y1293Q
V19	-	-	Y285Q	S317C	-	F497T	-	-	-	L635D	F703A	L726P	-	Y1015G	L1244G	-	L1281A	-
V20	-	-	-	S317C	-	-	-	-	L622Q	L635D	F703A	L726P	L815D	Y1015K	-	-	L1281A	Y1293Q

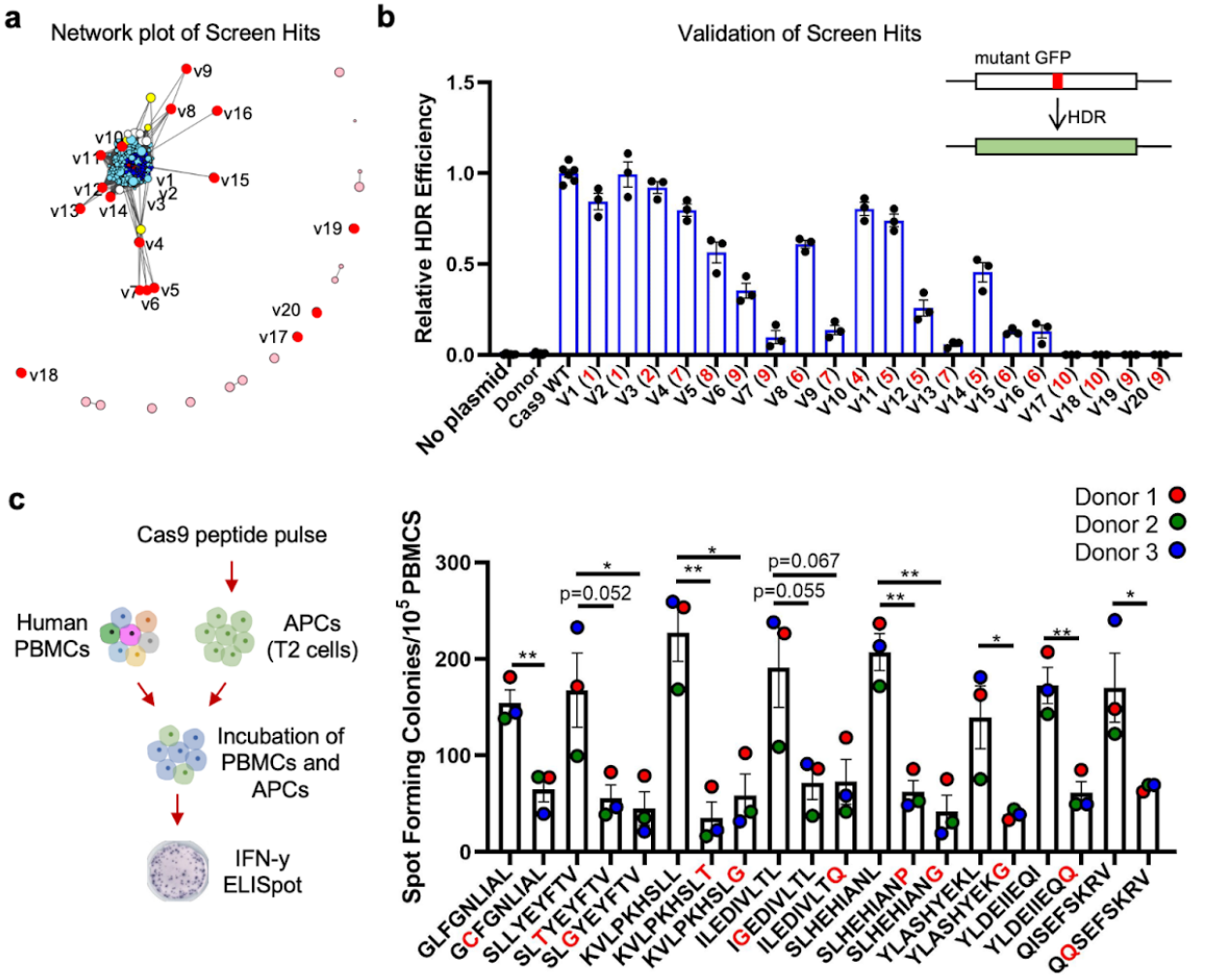


Figure 4.4: Validation of LORAX screen identified Cas9 variants for genome editing and de-immunization. (a) Network reconstruction connecting Cas9 variants with similar mutational patterns. Node colors indicate the number of deimmunized epitopes (dark blue < 3, light blue = 3, white = 4, yellow = 5, pink > 5). Circles in red represent tested variants and labeled with their respective names. (b) HEK293T bearing a GFP coding sequence disrupted by the insertion of a stop codon and a 68-bp genomic fragment of the AAVS1 locus were used as a reporter line. Wildtype (WT) or Cas9 variants, a sgRNA targeting the AAVS1 locus, and a donor plasmid capable of restoring GFP function via homology directed repair (HDR) were transfected into these cells and flow cytometry was performed on day 3. Relative quantification of GFP expression restoration by HDR is plotted. The number in parentheses represents the number of mutations in the variant. Values represented as mean +/- SEM (n=3). (c) T2 cells were pulsed with wildtype and variant peptides, cultured with PBMCs, and an ELISpot assay was performed to assess PBMC IFN-γ secretion to wildtype and variant peptides. Number of spot forming colonies for each peptide is plotted (n=3, mean +/- SEM, *p<0.05, **p<0.01, unpaired t-test two-tailed). Red letters in the peptide sequences represent the mutated amino acid.

4.1.8 Validations

4.1.8.1 Editing efficiency

Applying these analyses to the screen output led us to select and construct 20 variants (V1-20) (**Figure 4.3b**) for validation and characterization. We applied two independent methods to quantify editing of the deimmunized Cas9 variants. First, we performed a gene-rescue experiment using low frequency homology directed repair (HDR) to repair a genetically encoded broken green fluorescent protein (GFP) gene[23] (**Figure 4.4b**). And second, we quantified NHEJ mediated editing by genomic DNA extraction and Illumina next generation sequencing (NGS) using the CRISPResso2 package (**Figure 4.3b**)[367]. Variants highly connected to neighbors were capable of editing, whereas those not connected were non-functional, validating the network-based approach we used to select hits as enriching for truly functional sequences. Among the screen hits was the L614G mutation first identified by Ferdosi and colleagues[241] as a functional Cas9 variant with a critical immunodominant epitope de-immunized (V1). This concordance with previous work provided further confidence in our screening method. Interestingly, we discovered another deimmunizing mutation within the same epitope, L622Q (V2), which similarly retains Cas9 functionality, but appears to be more epistatically permissive, as many of our multi-mutation hits combine this mutation with other deimmunized epitopes.

From these multi-mutation hits we chose V4, which demonstrated high editing capability while still bearing simultaneous mutations across seven distinct epitopes, as well as family members V3, a variant bearing two mutations, and V5, a variant bearing the seven changes from V4 plus one additional mutation. We then further evaluated the efficacy of these mutants side-by-side with WT SpCas9 across a panel of genes and cell types, and assessed V4 activity across both

targeted genome editing and epigenome regulation experiments[298] (**Figure 4.5a-c**). Together, these results confirmed that leveraging our unique combinatorial library design and screening strategy, we were able to produce Cas9 variants with multiple top immunogenic epitopes simultaneously mutated (**Figure 4.6a**) while still retaining significant genome targeting functionality.

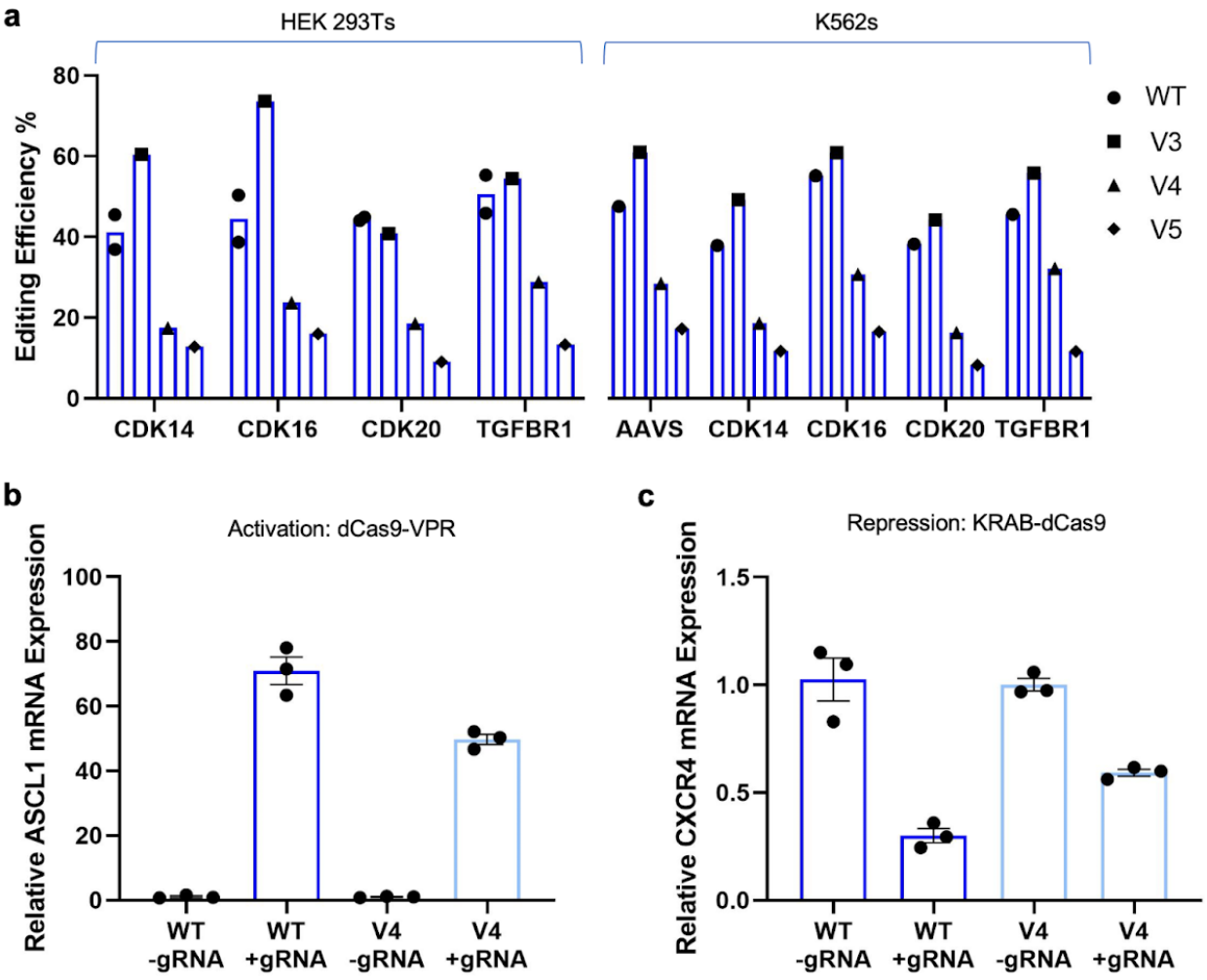


Figure 4.5: Characterization of Cas9 variants V3 (2 mutations), V4 (7 mutations), and V5 (8 mutations) across genome and epigenome targeting assays. (a) Cas9 wild-type or variants V3, V4, or V5, along with sgRNAs targeting the respective genes, were introduced into HEK293T and K562 cells. Editing efficiency of variants across 4 loci in HEK293Ts and 5 loci in K562s is plotted. (b) ASCL1 mRNA expression in cells transfected with dCas9 WT-VPR or dCas9 V4-VPR and sgRNA or no sgRNA is shown. Values represented as mean +/- SEM (n=3). (c) CXCR4 mRNA expression in cells transfected with dCas9 WT-KRAB or dCas9 V4-KRAB and sgRNA or no sgRNA is shown. Values represented as mean +/- SEM (n=3).

4.1.8.2 Deimmunization

To confirm that mutation of these epitopes indeed elicited deimmunization, we assessed T-cell response to wildtype and variant peptides by measuring IFN-γ secretion in the ELISpot assay[238,241]. We chose to use peripheral blood mononuclear cells (PBMCs) from three separate donors that carried the HLA-A*0201 allele as peptides were presented to cells using the TAP-deficient cell line T2 (HLA-A*0201 positive)[368]. Correspondingly, we synthesized peptides for epitopes 2, 7, 8, 9, 12, 15, and 16 as our predictions suggested these epitopes would induce a reduction in immune response for the HLA-A*0201 allele (**Figure 4.6a**). Importantly, since SpCas9v4 carries four of these mutations, this assay would also provide confirmation of deimmunization for this variant. We found that mutant peptides for all epitopes tested indeed resulted in fewer spot forming colonies for all three donors compared to wild type peptides (**Figure 4.4c, Figure 4.6b**), thereby confirming our predictions.

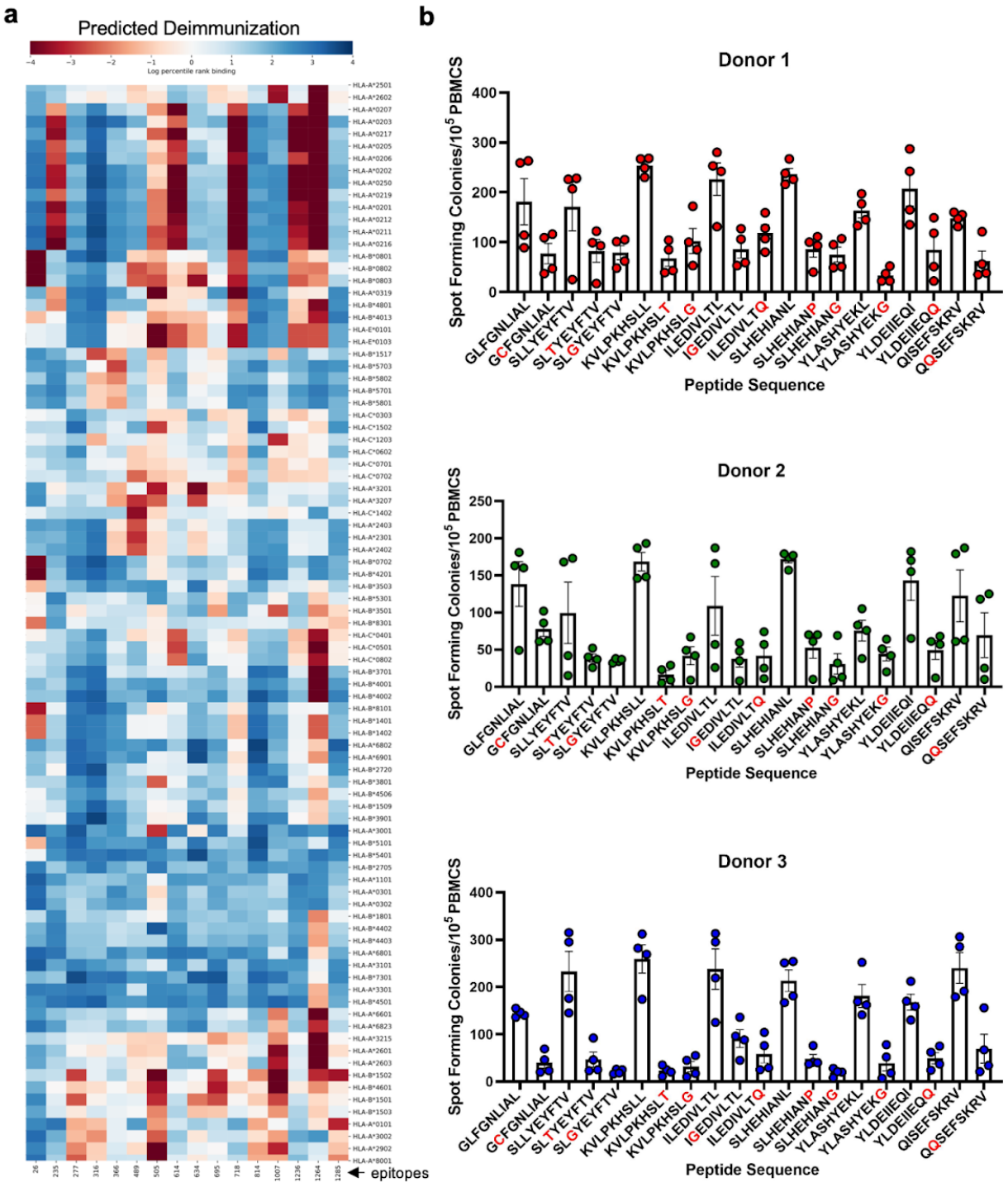


Figure 4.6: Predicted and experimentally confirmed deimmunization across Cas9 epitopes. (a) Predicted mutation-specific reduction in immunogenicity based on the epitope mutated and the HLA typing is depicted for each mutation included in the screen. (b) Technical replicates of spot forming colonies in the ELISpot assay are plotted for each donor (n=4).

4.1.8.3 Gene activation/repression

Based on this, we next evaluated delivery of SpCas9WT and SpCas9v4 and CRISPRoff versions of the same as icRNAs. CRISPRoff represents one of the newest additions to the CRISPR toolbox with the exciting capability to permanently silence gene expression upon transient expression[369]. We conjectured that wtCas9 and CRISPRoff would represent exciting applications of icRNAs for hit-and-run genome and epigenome targeting, as the prolonged persistence could potentially boost targeting, while the use of partially deimmunized Cas9 proteins would enable greater safety in therapeutic contexts. Specifically, icRNA for WT SpCas9 or SpCas9v4, along with sgRNA targeting the AAVS1 locus, or icRNA for CRISPRoff versions along with sgRNA targeting the B2M gene were transfected into HEK293T[370]. Excitingly, we observed both robust genome and epigenome targeting via the icRNA delivery format (**Figure 4.7a, b**).

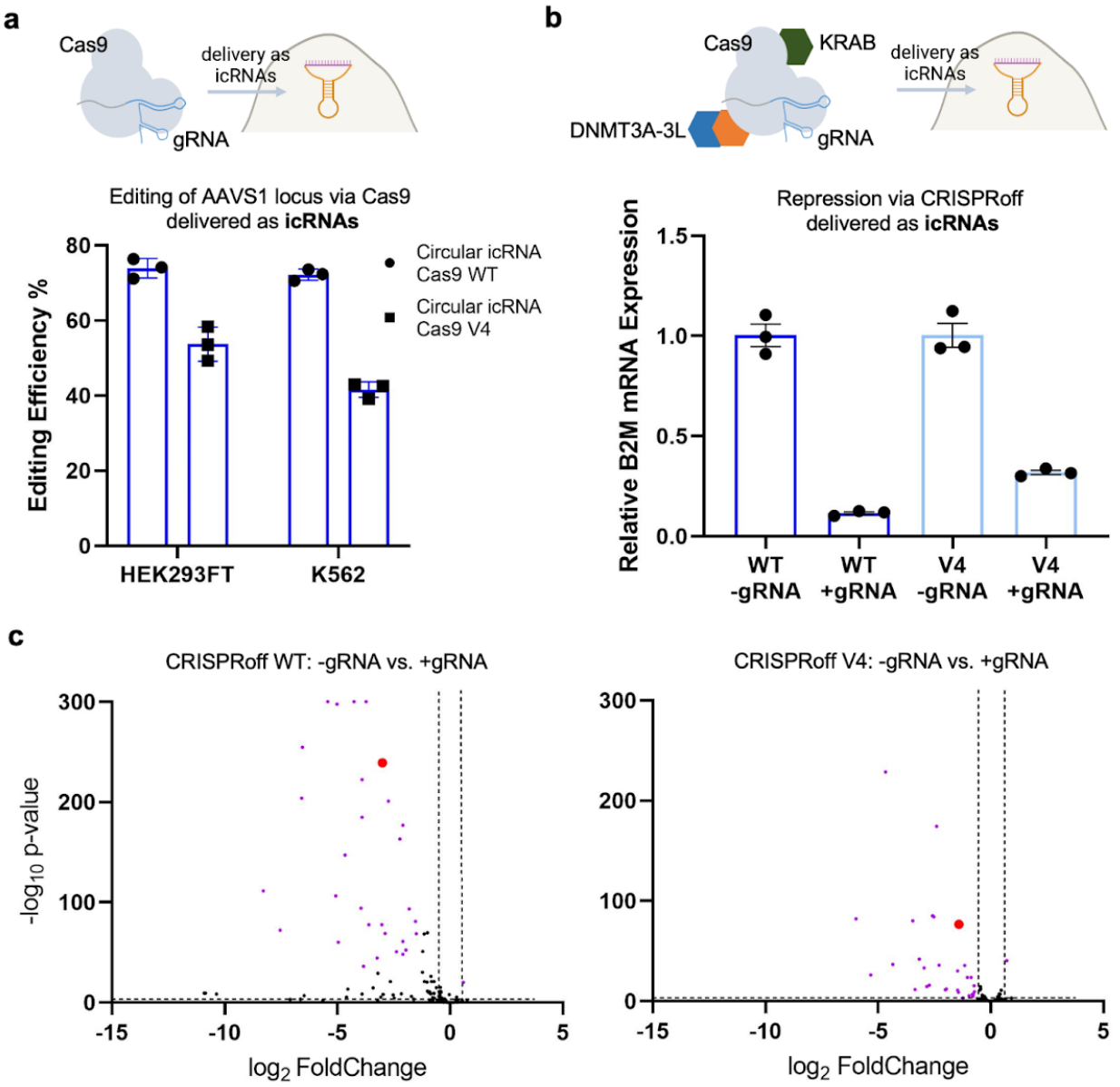


Figure 4.7: Delivery of de-immunized Cas9 variants as icRNAs for genome and epigenome targeting. (a) Circular icRNA for Cas9 wildtype or variant V4, along with a sgRNA targeting the AAVS1 locus, were introduced into HEK293T and K562 cells. Editing efficiency at the AAVS1 locus in the two cell lines are plotted. Values represented as mean +/- SEM (n=3). (b) Circular icRNA for CRISPRoff wildtype or variant V4, along with a sgRNA targeting the B2M gene, were introduced into HEK293T cells. B2M gene repression of CRISPRoff constructs in the presence or absence of sgRNA is plotted. Values represented as mean +/- SEM (n=3). (c) Volcano plot demonstrating differentially expressed genes for CRISPRoff wildtype with or without the B2M guide and variant V4 with or without guide. Dotted lines indicate the cutoff for significance ($\log_2$(fold change) greater than 0.5 or less than -0.5 and $-\log_{10}$ p-value greater than 3). B2M downregulation is confirmed by the red dots. Differentially expressed genes found in both wildtype and V4 are labeled with purple dots.

Lastly, to assess the specificity of SpCas9v4, we performed RNA sequencing on WT and SpCas9v4 CRISPRoff samples with and without the B2M guide. As expected, B2M was significantly down-regulated in SpCas9WT and SpCas9v4 samples containing the sgRNA compared to samples with no guide (**Figure 4.7c**, red dot). Importantly, all differentially expressed genes (DEGs) for V4 were also DEGs for WT, suggesting that SpCas9v4 and SpCas9WT are comparably specific in this assay (**Figure 4.7c**, purple dots).

4.2 DISCUSSION

Concurrently, to enable compatibility between persistence of expression and immunogenicity, we also developed the LORAX protein engineering platform that can be applied iteratively to tackle particularly challenging multiplexed protein engineering tasks by exploring huge swaths of combinatorial mutation space unapproachable using previous techniques. We demonstrated the power of this technique by creating a Cas9 variant with seven simultaneously deimmunized epitopes which still retains robust functionality in a single round of screening. This opens up a critical door in applying gene editing to long-persistence therapeutic modalities such as AAV or icRNA delivery. Furthermore, while this methodology is particularly suited to the unique challenges of protein deimmunization, it is also applicable to any potential protein engineering goal, so long as there exists an appropriate screening procedure to select for the desired functionality.

Similarly, the versatility of the LORAX platform comes with a set of limitations and tradeoffs that must be managed to leverage its utility. Naturally, library design is of critical importance. Here we have leveraged several features such as Cas9 evolutionary diversity, MHC-binding predictions, HLA allele frequencies, and calculated immunogenicity scores to generate a useful library of variants to test. Other approaches may bring in more sources of information from

places like protein structure[371], coevolutionary epistatic constraints[372], amino acid signaling motifs[373], or T-/B-cell receptor binding repertoires[374], among other possibilities. Another critical factor is careful selection of hits downstream of screening. Here we have developed a network-based method for differentiating spurious from bona fide hits leveraging known aspects of protein epistasis and fitness landscapes. Similar customizations and tweaks relevant to the specific biology of a given problem may yield substantial returns in applying LORAX or other large-scale combinatorial screening methods to various protein engineering challenges.

4.3 METHODS

4.3.1 Computational methods

4.3.1.1 Cas9 alignment and mutation selection

Naturally occurring variation in Cas9 sequence space was explored by aligning BLAST hits of the SpCas9 amino acid sequence. This set was then pruned by removing truncated, duplicated, or engineered sequences, and those sequences whose origin could not be determined. At specified immunogenic epitopes and key anchor residues, top alternative amino acids were obtained using frequency in the alignment weighted by overall sequence identity to the wild type SpCas9 sequence, such that commonly occurring amino acid substitutions appearing in sequences highly similar to the wild-type were prioritized for further analysis and potential inclusion in the LORAX library.

4.3.1.2 HLA frequency estimation and binding predictions

HLA-binding predictions were carried out using netMHC4.1 or netMHCpan3.1. Global HLA allele frequencies were estimated from data at allelefrequencies.net as follows. Data was

divided into 11 geographical regions. Allele frequencies for each of those regions were estimated from all available data from populations therein. These regional frequencies were then averaged weighted by global population contribution. Alleles with greater than 0.001% frequency in the global population, or those with greater than 0.01% in any region, were included for further analysis and predictions.

4.3.1.3 Immunogenicity scores

The vector of predicted nM affinities output by netMHC were first normalized across alleles to account for the fact that some alleles have higher affinity across all peptides, and to allow for the relatively equivalent contribution of all alleles. These values were then transformed using the 1-log(affinity) transformation also borrowed from netMHC such that lower nM affinities will result in larger resulting values. These transformed, normalized affinities are then weighted by population allele frequency and summed across all alleles and epitopes. Finally, the scores are standardized across proteins to facilitate comparison.

4.3.1.4 Base calling and genotyping

Raw reads coming off the MinION flow cell were base-called using Guppy 3.6.0 and aligned to an SpCas9 reference sequence containing non-informative NNN bases at library mutation positions, so as not to bias calling towards wild-type or mutant library members, using Minimap2's map-ont presets. Reads covering the full length of the Cas9 gene with high mapping quality were genotyped at each individual mutation site and tabulated to the corresponding library member. Reads with ambiguous sites were excluded from further analysis.

4.3.1.5 Cluster analysis

Network analysis was performed by first thresholding genotypes to include only those identified as hits from the screen. These were genotypes appearing in the pre-screen plasmid library, both post-screen replicates, and having a fold-change enrichment larger than the wild-type sequence (4.5-fold enrichment). These hits were used to create a graph with nodes corresponding to genotypes and node sizes corresponding to fold change enrichment. Edges were placed between nodes at most 4 mutations distant from each other, and edge weights were defined by 1/d where d is distance between genotypes. Network analysis was done using python bindings of igraph. Plots were generated using the Fruchterman-Reingold force-directed layout algorithm.

4.3.2 Experimental methods

4.3.2.1 Cell culture

HEK293T and HeLa cells were cultured in DMEM supplemented with 10% FBS and 1% Antibiotic-Antimycotic (Thermo Fisher). K562 cells were cultured in RPMI supplemented with 10% FBS and 1% Antibiotic-Antimycotic (Thermo Fisher). All cells were cultured in an incubator at 37°C and 5% CO_2.

DNA transfections were performed by seeding HEK293T cells in 12 well plates at 25% confluency and adding 1 µg of each DNA construct and 4 µL of Lipofectamine 2000 (Thermo Fisher). RNA transfections were performed by adding 1 µg of each RNA construct and 3.5 µL of Lipofectamine MessengerMax (ThermoFisher). Electroporations were performed in K562 cells using the SF Cell Line 4D-Nucleofector X Kit S (Lonza) per manufacturer's protocol.

4.3.2.2 Flow cytometry experiments:

GFP intensity, defined as the median intensity of the cell population, was quantified after transfection using a BD LSRFortessa Cell Analyzer.

4.3.2.3 Lipid nanoparticle formulations:

(6Z,9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-yl-4-(dimethylamino) butanoate (DLin-MC3-DMA) was purchased from BioFine International Inc. 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC) and 1,2-dimyristoyl-rac-glycero-3-methoxypolyethylene glycol-2000 (DMG-PEG-2000) were purchased from Avanti Polar Lipids. Cholesterol was purchased from Sigma-Aldrich. mRNA LNPs were formulated with DLin-MC3-DMA:cholesterol:DSPC:DMG-PEG at a mole ratio of 50:38.5:10:1.5 and a N/P ratio of 5.4. To prepare LNPs, lipids in ethanol and mRNA in 25 mM acetate buffer, pH 4.0 were combined at a flow rate of 1:3 in a PDMS staggered herringbone mixer[375]. The dimensions of the mixer channels were 200 by 100 um, with herringbone structures 30 um high and 50 um wide. Immediately after formulation, 3 volumes of PBS was added and LNPs were purified in 100 kDa MWCO centrifugal filters by exchanging the volume 3 times. Final formulations were passed through a 0.2 um filter. LNPs were stored at 4°C for up to 4 days before use. LNP hydrodynamic diameter and polydispersity index were measured by dynamic light scattering (Malvern NanoZS Zetasizer). The mRNA content and percent encapsulation were measured with a Quant-iT RiboGreen RNA Assay (Invitrogen) with and without Triton X-100 according to the manufacturer's protocol.

4.3.2.4 Animal experiments

All animal procedures were performed in accordance with protocols approved by the Institutional Animal Care and Use Committee of the University of California, San Diego. All mice were acquired from Jackson Labs.

To assess persistence of RNA constructs *in vivo*, 10 μg of circular GFP icRNA or linear GFP icdRNA LNPs were injected retro-orbitally into C57BL/6J mice. After 3 days and 7 days, livers were isolated and placed in RNAlater (Sigma-Aldrich). RNA was later isolated using QIAzol Lysis Reagent and purified using RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol. Amount of circularized RNA were assessed by RT-qPCR.

To investigate the ability of circular icRNA and linear icdRNA COVID RNA to elicit an immune response, BALB/c mice were injected intramuscularly into the gastrocnemius muscle with PBS, 0.2 μg, or 2 μg of Omicron Spike (2P) circular icRNA or linear icdRNA. A booster shot containing the same amount and type of RNA was performed on day 21. Blood draws were performed on days 0, 9, 21, 31, and 42, serum was separated using blood collection tubes (Sarstedt), and antibody production was then assessed by a sandwich enzyme-linked immunosorbent assay (ELISA). ELISA was performed using the ELISA Starter Accessory Kit (Bethyl, E101) per manufacturer's instructions. Briefly, 96-well MaxiSorp well plates were coated with recombinant SARS-COV-2 Spike protein S1, Omicron variant (GenScript Biotech) diluted in 1x coating buffer (Bethyl) to a concentration of 2 μg/mL overnight at 4C. Plates were washed five times with 1x washing buffer (Bethyl), followed by the addition of 1x blocking buffer for 1 hour at RT. Samples were diluted 1:50 in sample/conjugate diluent (Bethyl) for days 9 and 21 and 1:200 for days 31 and 42 and added to the plate for 2 hours at RT. Sample/conjugate diluent was used as a blank. Plates were washed five times with 1x washing buffer and incubated in secondary antibody (horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG antibody, Southern Biotech 1036-05, diluted 1:5000 in sample/conjugate diluent) for 1 hour at RT. After five washes, 50 μL/well TMB One Component HRP Microwell Substrate (Bethyl) was added and incubated at RT in the dark. 50 μL/well of 0.2M H_2SO_4 was added to terminate color development and

absorbance was measured at 450 nm in a SpectraMax iD5 Multi-Mode Microplate Reader (Molecular Devices).

4.3.2.5 Identification of HPRT1 Guide

The lentiCRISPR-v2 plasmid (Addgene #52961) was first digested with Esp3I and a guide targeting the HPRT1 gene was cloned in via Gibson assembly. After lentivirus production, HeLa cells were seeded at 25% confluency in 96 well plates and transduced with virus (lentiCRISPR-v2 with or without HPRT1 guide) and 8 μg/mL polybrene (Millipore). Virus was removed the next day and 2.5 μg/mL puromycin was added to remove cells that did not receive virus two days later. After 2 days of puromycin selection, 0-14 μg/mL 6-TG was added. After 5 days, cells were stained with crystal violet, solubilized using 1% sodium dodecyl sulfate, and absorbance was measured at 595 nm on a plate reader. 6 μg/mL was chosen due to the lack of cells in the negative control.

4.3.2.6 Generation of variant Cas9 library

Cas9 variant sequences were generated by separating the full-length gene sequence into small sections, where each section contained wildtype or variant Cas9 sequences. Degenerate pools of these gBlocks were PCR amplified and annealed together, yielding a final library size of 1,492,992 elements. The lentiCRISPR-v2 plasmid containing the HPRT1 guide was digested with BamHI and XbaI and Gibson assembly was used to clone elements into the vector. The Gibson reactions were then transformed into electrocompetent cells and cultured at 37C overnight. Plasmid DNA was isolated using the Qiagen Plasmid Maxi Kit and library coverage was estimated by calculating the number of colonies found on LB-carbenicillin plates. DNA was then used to create lentivirus containing the variant Cas9 library.

4.3.2.7 Cas9 Screen

HeLa cells were seeded in 15 15-cm plates and transduced with virus containing the variant Cas9 library and 8 µg/mL polybrene. Media was changed the next day and 2.5 µg/mL puromycin was added to remove cells that did not receive virus two days later. 6 µg/mL 6-TG was added to media once cells reached 90% confluency. Media was changed every other day for ten days to allow for selection of cells containing functional Cas9 variants. After ten days, cells were lifted from the plates and DNA was isolated using the DNeasy Blood & Tissue Kit per manufacturer's protocol.

4.3.2.8 Nanopore Sequencing

Pre-screen analysis of the Cas9 variant library elements was performed by amplifying the sequence from the plasmid. 1 µg of the variant Cas9 sequences was used for library preparation using the Ligation Sequencing Kit (Oxford Nanopore Technologies, SQK-LSK109) per manufacturer's instructions. DNA was then loaded into a MinION flow cell (Oxford Nanopore Technologies, R9.4.1). Post-screen analysis of library elements was performed by amplifying the Cas9 sequences from 75 µg of genomic DNA. 1 µg of the variant Cas9 sequences was similarly prepared using the Ligation Sequencing Kit and sequenced on a MinION flow cell.

4.3.2.9 HDR validation

Lentivirus was produced from a plasmid containing a GFP sequence with a stop codon and 68 bp AAVS1 fragment. HEK293T cells were treated with 8 µg/mL polybrene and lentivirus. After puromycin selection to create a stable line, cells were transfected with plasmids containing

variant Cas9 sequences, a guide targeting the AAVS locus and a GFP repair donor plasmid. After 3 days, FACS was performed and percent GFP positive cells were quantified.

4.3.2.10 Genome engineering experiments

To validate variant Cas9 functional cutting, variant Cas9 and guides were transfected into HEK293T cells. After two days, genomic DNA was isolated. Genomic DNA was also isolated after two days from K562 cells after electroporation. To assess activity of CCR5 ZFNs delivered as icRNAs, HEK293Ts were transfected with circular icRNA or linear icdRNA and genomic DNA was isolated after three days. Assessment of GFP ZFN was performed by transfecting HEK293Ts stably expressing a broken GFP with circular icRNA or linear icdRNA and isolating genomic DNA after three days. To assess activity of Cas9 delivered as icRNAs, HEK293Ts and K562 were transfected or nucleofected with Cas9 WT or Cas9 v4 along with a guide RNA (synthesized via Synthego) and genomic DNA was isolated after three days.

4.3.2.11 Epigenome engineering experiments

dCas9-VPR experiments were performed by transfecting HEK293T cells with dCas9wt-VPR or dCas9v4-VPR with or without a gRNA targeting the ASCL1 gene. Likewise, KRAB-dCas9 experiments were performed by transfecting cells with KRAB-dCas9wt or KRAB-dCas9v4 with or without a gRNA targeting the CXCR4 gene. CRISPRoff experiments were performed by transfecting HEK293T cells with circular icRNA CRISPRoffwt or CRISPRoffv4 with or without a gRNA targeting the B2M gene (Synthego). RNA was isolated three days later and repression or activation of genes was assessed by qPCR.

4.3.2.12 Quantification of editing using NGS

After extraction of genomic DNA, PCR was performed to amplify the target site. Amplicons were then indexed using the NEBNext Multiplex Oligos for Illumina kit (NEB). Amplicons were then pooled and sequenced using a Miseq Nano with paired end 150 bp reads. Editing efficiency was quantified using CRISPResso2.

4.3.2.13 Cas9 Specificity

RNA isolated from the CRISPRoff experiment was used to assess specificity. RNAseq libraries were generated from 300 ng of RNA using the NEBNext Poly(A) mRNA magnetic isolation module and NEBNext Ultra II Directional RNA Library Prep kit for Illumina and sequenced on the Illumina NovaSeq 6000 with paired end 100 bp reads. Fastq files were mapped to the reference human genome hg38 using STAR aligner. Differential gene expression was analyzed using the Bioconductor package DESeq2 with the cutoff of $\log_2$(fold change) greater than 0.5 or less than -0.5 and a p-value less than 10^{-3}. To identify differentially expressed genes, CRISPRoff WT and V4 samples containing the B2M guide were compared to samples not receiving the guide.

4.3.2.14 ELISpot assay

TAP-deficient T2 cells were a generous gift from Stephen Schoenberger lab. PBMCs were purchased from StemCell Technologies. All donors contained the HLA-A*0201 allele. Both cell lines were maintained in RPMI1640 media supplemented with 10% FBS, 1% Penicillin-Streptomycin, 10 mM HEPES, and 1 mM sodium pyruvate. On the first day, PBMCs were thawed and rested overnight at a density of 10^6 cells/mL. T2 cells were pulsed with peptides at 10 μg/mL overnight. Peptides were produced from Genscript's Custom Peptide Synthesis service at crude purity. Lastly, 96-well plates (Immobilon-P, Millipore) were coated with 10 μg/mL anti-IFNγ

monoclonal antibody (1-D1K, Mabtech) overnight at 4C. The next day, T2 cells were washed two times and 50,000 T2 cells and 100,000 PBMCs were added to each well. 4 replicates were used per condition. After 22 hours, cells were removed from the plate and 2 μg/mL biotinylated anti-IFNγ secondary antibody (7-B6-1, Mabtech) was added for 2 hours. Plates were washed and 1:1000 Streptavidin-ALP (3310-10-1000, Mabtech) was added for 45 minutes. Plates were washed and color was developed by adding BCIP/NBT-plus substrate (3650-10, Mabtech) for 10 minutes. Plates were thoroughly washed in water, dried at room temperature, and spots were automatically counted using an ELISpot plate reader.

4.3.2.15 Lentivirus production

HEK293FT cells were seeded in 1 15-cm plate and transfected with 36 μL Lipofectamine 2000, 3 μg pMD2.G (Addgene #12259), 12 μg pCMV delta R8.2 (Addgene #12263), and 9 μg of the lentiCRISPR-v2 plasmid. Supernatant containing viral particles was harvested after 48 and 72 hours, filtered with 0.45 μm Steriflip filters (Millipore), concentrated to a final volume of 1 mL using an Amicon Ultra-15 centrifugal filter unit with a 100,000 NMWL cutoff (Millipore), and frozen at -80C.

4.3.2.16 RT-qPCR

cDNA was synthesized from RNA using the Protoscript II First Strand cDNA Synthesis Kit (NEB). qPCR was performed using a CFX Connect Real Time PCR Detection System (Bio-Rad). All samples were run in triplicates and results were normalized against GAPDH expression. Primers for qPCR are listed in **(Table 4.1)**.

Table 4.1: qPCR primers

CXCR4_F	GAAGCTGTTGGCTGAAAAGG
CXCR4_R	CTCACTGACGTTGGCAAAGA
ASCL1_F	CGCGGCCAACAAGAAGATG
ASCL1_R	CGACGAGTAGGATGAGACCG
B2M_F	TATGCCTGCCGTGTGAACCATGT
B2M_R	GGCATCTTCAAACCTCCATGATGCT
GAPDH_F	ACAGTCAGCCGCATCTTCTT
GAPDH_R	ACGACCAAATCCGTTGACTC
GFP_F	ACGTAAACGGCCACAAGTTC
GFP_R	AAGTCGTGCTGCTTCATGTG

4.4 ACKNOWLEDGEMENTS

Chapter 4, in part, is a reprint of the material as it appears in bioRxiv (2022). Aditya Kumar, Nathan Palmer, Katelyn Miyasaki, Emma Finburgh, Yichen Xiang, Andrew Portell, Amir Dailamy, Amanda Suhardjo, Wei Leong Chew, Ester J. Kwon, & Prashant Mali. The dissertation author was the primary researcher and author of this paper. Thanks to members of the Mali lab for discussions, advice and help with experiments. This work was generously supported by UCSD Institutional Funds, NIH grants (R01HG009285, R01CA222826, R01GM123313), Department of Defense Grant (DODPR210085), and a Longevity Impetus Grant from Norn Group. This publication includes data generated at the UC San Diego IGM Genomics Center utilizing an Illumina NovaSeq 6000 that was purchased with funding from a National Institutes of Health SIG grant (S10 OD026929). Some schematics were created using BioRender.

CHAPTER 5 – CONCLUSIONS

5.1 SUMMARY OF THE WORK

In chapter 1, I present an overview of the state of CRISPR-Cas gene therapy especially focusing on the most promising therapeutic modalities and the significant, though certainly not impassable, obstacles they still face in translation to clinical use. Narrowing in on one major obstacle, namely, the interaction with the immune system, in chapter 2 I take a comprehensive approach to defining this issue, characterizing the immune response to AAV-delivered CRISPR-Cas at both the humoral and cell-mediated levels. I also explore which types of immunity are critical for which components of the therapeutic target, highlighting the fact that neutralizing antibodies against the delivery vehicle can fully inhibit therapeutic effect, and that anti-Cas9 T-cells induced upon administration of a CRISPR therapy can substantially mitigate the efficacy of gene editing.

With this understanding in hand, I set out in chapter 3 to examine how natural orthologs of AAV and Cas can circumvent the immune response due to immune orthogonality: that is, preserved function, but divergent structure, such that MHC-restricted epitopes are not shared among immune orthogonal orthologs. Testing the immune orthogonality of AAV5 and AAV8, as well as SpCas9 and SaCas9 revealed that the use of these immune orthogonal proteins enabled multiple dosing strategies that are inhibited by immunity when using the same protein for consecutive doses.

Building on this, I set out in chapter 4 to begin to approach immune avoidance directly through engineering of Cas9 to delete immunogenic MHC-restricted epitopes. Epitope deletion has been successfully applied to Cas9 previously, but only in a one-by-one manner[241], which is simply insufficient to generally dampen the immune response, especially across patients with

divergent MHC haplotypes. A more comprehensive approach to Cas9 deimmunization requires specialized techniques to screen mutations across the full length of the protein in combinatorial fashion while preserving protein function. To this end, I develop the LORAX suite of protein engineering tools to enable progressive screening of combinations of epitope-abolishing mutations millions of elements at a time. Using this innovation, I identify a Cas9 variant with seven simultaneously de-immunized epitopes which retains near wild-type efficiency both in traditional DNA cutting and gene activation/repression modalities. We additionally validate the immune avoidance of these modified epitopes by showing diminished T-cell responses to the mutated epitopes. This approach is generalizable to a wide variety of protein engineering problems and will help open the door to bringing CRISPR-Cas to the clinic with a greater chance of both safety and efficacy.

5.2 LIMITATIONS

Although this represents a significant step in the right direction, there are limitations to these approaches to immune avoidance. Firstly, abolishing T-cell epitopes may not block B-cell responses, as they are able to recognize antigens in their native form rather than relying on MHC-restricted epitopes. Although there is reason to believe that limiting MHCII-mediated activation of CD4+ helper T-cells can impede the effectiveness of memory B-cell activation, antibody isotype switching and affinity maturation that mediates the stronger secondary immune response upon repeated exposure to an antigen[345]. Secondly, most of this work was performed in the C57B6 or BALB/c mouse lines, which likely have diminished immune responses due to homozygosity created by inbreeding. Extrapolating discoveries in this model to non-human primates and then to humans may require a quantitative reassessment of the degree of engineering required to avoid the immune response. Finally, further advancements to the LORAX platform, especially in the choice

of mutations will likely enhance the utility of this approach by bringing in more sources of information, including structural and coevolutionary attributes among others, into what mutations are likely to be both de-immunizing and non-deleterious to protein function.

5.3 BROADER APPLICATIONS

We are currently witnessing the birth of a new therapeutic paradigm based on CRISPR-Cas to precisely target genetic and epigenetic disease in a powerful and previously inaccessible manner. Given the rapid development and adoption of the CRISPR-Cas system, the explosion of preclinical studies, and the progress in clinical trials in the last few years, the near future will likely see the approval and deployment of several CRISPR-based therapies targeting diverse pathologies, some previously untreatable, and likely encompassing multiple therapeutic and delivery modalities. Despite this exciting prognosis, significant challenges remain, especially precise and high-quantity delivery to target tissues, demonstrated safety both at the level of off-target effects and immune interactions, and scalable, affordable deployment of the drugs. The substantial research efforts mounted towards all these issues have yielded rapid advancements on every front, but much more remains to be done.

Although first-generation gene therapy Zolgensma[4] garnered the epithet of "most expensive drug ever" at $2.125 million, we do not believe this must be a lasting attribute of this class of drugs. Although aspects such as AAV manufacturing challenges, R&D costs, regulatory burdens, and pharmaceutical market economics are important reasons for large price tags, the pace of technical progress can rapidly alter what once seemed set in stone. The precision and deftness with which we can manipulate genomes using CRISPR systems was virtually unthinkable just 17 years ago when the $2.7 billion human genome project was completed, just one year after the acronym CRISPR was first proposed[376]. Now we have a thousand-dollar genome and a million-

dollar gene therapy. Given the explosion of research and development into CRISPR-based gene therapy and the expansion of its application from rare genetic disease to large-market ubiquitous conditions such as chronic pain[377], cardiovascular disease[221,378], and Alzheimer's[7], in the decades to come it is not entirely implausible nor unprecedented to hope for a few thousand-dollar gene therapy.

Optimizing the translation of new technologies to rapidly enable disease treatment is of critical importance but is far from an easy task. Many therapeutic modalities offering huge benefit have been stymied by unforeseen or unappreciated setbacks on the path to approval, often taking decades to address. We have attempted here to outline the current state and translational path forward of CRISPR therapeutics, highlighting areas where additional research is required, or potential challenges may occur in order to help orient thinking and optimize resources to solve these challenges. Expediting the development of these therapies will require vigilance, insight, and cooperation among scientists of multiple backgrounds, pharmaceutical companies, and regulatory stakeholders to ensure that the potential of CRISPR-based drugs may be realized as efficiently and safely as possible.

5.4 ACKNOWLEDGMENTS

Chapter 5, in part, is a reprint of the material as it appears in The CRISPR Journal 3(4): 253–275 (2020). Tay, Lavina Sierra; Palmer, Nathan; Panwala, Rebecca; Chew, Wei Leong; Mali, Prashant. The dissertation author was the primary researcher and author of this paper. This work was generously supported by NIH grants (R01HG009285, RO1CA222826, RO1GM123313) and by the Agency for Science, Technology, and Research under its Industrial Alignment Fund (Pre-Positioning) (H17/01/a0/012).

REFERENCES

1. Dunbar, C. E., High, K. A., Joung, J. K., Kohn, D. B., Ozawa, K. & Sadelain, M. Gene therapy comes of age. *Science* **359**, eaan4672 (2018).

2. Schuessler-Lenz, M., Enzmann, H. & Vamvakas, S. Regulators' Advice Can Make a Difference: European Medicines Agency Approval of Zynteglo for Beta Thalassemia. *Clin. Pharmacol. Ther.* **107**, 492–494 (2020).

3. Russell, S., Bennett, J., Wellman, J. A., Chung, D. C., Yu, Z.-F., Tillman, A., Wittes, J., Pappas, J., Elci, O., McCague, S., Cross, D., Marshall, K. A., Walshire, J., Kehoe, T. L., Reichert, H., Davis, M., Raffini, L., George, L. A., Hudson, F. P., Dingfield, L., Zhu, X., Haller, J. A., Sohn, E. H., Mahajan, V. B., Pfeifer, W., Weckmann, M., Johnson, C., Gewaily, D., Drack, A., Stone, E., Wachtel, K., Simonelli, F., Leroy, B. P., Wright, J. F., High, K. A. & Maguire, A. M. Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65 -mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. *The Lancet* **390**, 849–860 (2017).

4. Mendell, J. R., Al-Zaidy, S., Shell, R., Arnold, W. D., Rodino-Klapac, L. R., Prior, T. W., Lowes, L., Alfano, L., Berry, K., Church, K., Kissel, J. T., Nagendran, S., L'Italien, J., Sproule, D. M., Wells, C., Cardenas, J. A., Heitzer, M. D., Kaspar, A., Corcoran, S., Braun, L., Likhite, S., Miranda, C., Meyer, K., Foust, K. D., Burghes, A. H. M. & Kaspar, B. K. Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. *N Engl J Med* **377**, 1713–1722 (2017).

5. Stadtmauer, E. A., Fraietta, J. A., Davis, M. M., Cohen, A. D., Weber, K. L., Lancaster, E., Mangan, P. A., Kulikovskaya, I., Gupta, M., Chen, F., Tian, L., Gonzalez, V. E., Xu, J., Jung, I., Melenhorst, J. J., Plesa, G., Shea, J., Matlawski, T., Cervini, A., Gaymon, A. L., Desjardins, S., Lamontagne, A., Salas-Mckee, J., Fesnak, A., Siegel, D. L., Levine, B. L., Jadlowsky, J. K., Young, R. M., Chew, A., Hwang, W.-T., Hexner, E. O., Carreno, B. M., Nobles, C. L., Bushman, F. D., Parker, K. R., Qi, Y., Satpathy, A. T., Chang, H. Y., Zhao, Y., Lacey, S. F. & June, C. H. CRISPR-engineered T cells in patients with refractory cancer. *Science* **367**, eaba7365 (2020).

6. Strich, J. R. & Chertow, D. S. CRISPR-Cas Biology and Its Application to Infectious Diseases. *J Clin Microbiol* **57**, e01307-18, /jcm/57/4/JCM.01307-18.atom (2018).

7. Park, H., Oh, J., Shim, G., Cho, B., Chang, Y., Kim, S., Baek, S., Kim, H., Shin, J., Choi, H., Yoo, J., Kim, J., Jun, W., Lee, M., Lengner, C. J., Oh, Y.-K. & Kim, J. In vivo neuronal gene editing via CRISPR–Cas9 amphiphilic nanocomplexes alleviates deficits in mouse models of Alzheimer's disease. *Nat Neurosci* **22**, 524–528 (2019).

8. Xie, C., Zhang, Y.-P., Song, L., Luo, J., Qi, W., Hu, J., Lu, D., Yang, Z., Zhang, J., Xiao, J., Zhou, B., Du, J.-L., Jing, N., Liu, Y., Wang, Y., Li, B.-L., Song, B.-L. & Yan, Y. Genome editing with CRISPR/Cas9 in postnatal mice corrects PRKAG2 cardiac syndrome. *Cell Res* **26**, 1099–1111 (2016).

9. Lu, Y., Xue, J., Deng, T., Zhou, X., Yu, K., Deng, L., Huang, M., Yi, X., Liang, M., Wang, Y., Shen, H., Tong, R., Wang, W., Li, L., Song, J., Li, J., Su, X., Ding, Z., Gong, Y., Zhu, J., Wang, Y., Zou, B., Zhang, Y., Li, Y., Zhou, L., Liu, Y., Yu, M., Wang, Y., Zhang, X., Yin, L., Xia, X., Zeng, Y., Zhou, Q., Ying, B., Chen, C., Wei, Y., Li, W. & Mok, T. Safety and feasibility of CRISPR-edited T cells in patients with refractory non-small-cell lung cancer. *Nat Med* **26**, 732–740 (2020).

10. Lino, C. A., Harper, J. C., Carney, J. P. & Timlin, J. A. Delivering CRISPR: a review of the challenges and approaches. *Drug Delivery* **25**, 1234–1257 (2018).

11. Barnes, C., Scheideler, O. & Schaffer, D. Engineering the AAV capsid to evade immune responses. *Current Opinion in Biotechnology* **60**, 99–103 (2019).

12. Liu, J.-J., Orlova, N., Oakes, B. L., Ma, E., Spinner, H. B., Baney, K. L. M., Chuck, J., Tan, D., Knott, G. J., Harrington, L. B., Al-Shayeb, B., Wagner, A., Brötzmann, J., Staahl, B. T., Taylor, K. L., Desmarais, J., Nogales, E. & Doudna, J. A. CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature* **566**, 218–223 (2019).

13. Pausch, P., Al-Shayeb, B., Bisom-Rapp, E., Tsuchida, C. A., Li, Z., Cress, B. F., Knott, G. J., Jacobsen, S. E., Banfield, J. F. & Doudna, J. A. CRISPR-CasΦ from huge phages is a hypercompact genome editor. *Science* **369**, 333–337 (2020).

14. Sack, B. K. & Herzog, R. W. Evading the immune response upon in vivo gene therapy with viral vectors. *Curr Opin Mol Ther* **11**, 493–503 (2009).

15. Fu, H., Cataldi, M. P., Ware, T. A., Zaraspe, K., Meadows, A. S., Murrey, D. A. & McCarty, D. M. Functional correction of neurological and somatic disorders at later stages of disease in MPS IIIA mice by systemic scAAV9-hSGSH gene delivery. *Molecular Therapy - Methods & Clinical Development* **3**, 16036 (2016).

16. Silva Lima, B. & Videira, M. A. Toxicology and Biodistribution: The Clinical Value of Animal Biodistribution Studies. *Molecular Therapy - Methods & Clinical Development* **8**, 183–197 (2018).

17. Petit, L., Khanna, H. & Punzo, C. Advances in Gene Therapy for Diseases of the Eye. *Human Gene Therapy* **27**, 563–579 (2016).

18. Zheng, C. & Baum, B. J. Evaluation of Promoters for Use in Tissue-Specific Gene Delivery. in *Gene Therapy Protocols* (ed. Doux, J. M.) 205–219 (Humana Press, 2008). doi:10.1007/978-1-60327-248-3_13.

19. Li, Q., Qin, Z., Wang, Q., Xu, T., Yang, Y. & He, Z. Applications of Genome Editing Technology in Animal Disease Modeling and Gene Therapy. *Computational and Structural Biotechnology Journal* **17**, 689–698 (2019).

20. Gonzalez, L., Strbo, N. & Podack, E. R. Humanized mice: novel model for studying mechanisms of human immune-based therapies. *Immunol Res* **57**, 326–334 (2013).

21. Husain, S. R., Han, J., Au, P., Shannon, K. & Puri, R. K. Gene therapy for cancer: regulatory considerations for approval. *Cancer Gene Ther* **22**, 554–563 (2015).

22. Zincarelli, C., Soltys, S., Rengo, G. & Rabinowitz, J. E. Analysis of AAV Serotypes 1–9 Mediated Gene Expression and Tropism in Mice After Systemic Injection. *Molecular Therapy* **16**, 1073–1080 (2008).

23. Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., Church, G. M., Wiedenheft, B., Sternberg, S. H., Doudna, J. A., Bhaya, D., Davison, M., Barrangou, R., Terns, M. P., Terns, R. M., Jinek, M., Gasiunas, G., Barrangou, R., Horvath, P., Siksnys, V., Sapranauskas, R., Brummelkamp, T. R., Bernards, R., Agami, R., Miyagishi, M., Taira, K., Deltcheva, E., Zou, J., Mali, P., Huang, X., Dowey, S. N., Cheng, L., Sanjana, N. E., Lee, J. H., Hockemeyer, D., Kosuri, S., Pattanayak, V., Ramirez, C. L., Joung, J. K., Liu, D. R., King, N. M., Cohen-Haguenauer, O., Kim, Y. G., Cha, J., Chandrasegaran, S., Rebar, E. J., Pabo, C. O., Boch, J., Moscou, M. J., Bogdanove, A. J., Khalil, A. S., Collins, J. J., Purnick, P. E., Weiss, R., Zou, J., Holt, N., Urnov, F. D., Lombardo, A., Li, H., Makarova, K. S., Horvath, P., Barrangou, R., Deveau, H., Ploeg, J. R. van der, Rho, M., Wu, Y. W., Tang, H., Doak, T. G., Ye, Y., Pride, D. T., Garneau, J. E., Esvelt, K. M., Carlson, J. C., Liu, D. R., Barrangou, R., Horvath, P., Kent, W. J., Dreszer, T. R., Karolchik, D., Quinlan, A. R., Hall, I. M., Langmead, B., Trapnell, C., Pop, M., Salzberg, S. L., Lorenz, R., Mathews, D. H., Sabina, J., Zuker, M., Turner, D. H., Thurman, R. E., Xu, Q., Schlabach, M. R., Hannon, G. J. & Elledge, S. J. RNA-guided human genome engineering via Cas9. *Science (New York, N.Y.)* **339**, 823–6 (2013).

24. Jinek, M., East, A., Cheng, A., Lin, S., Ma, E. & Doudna, J. RNA-programmed genome editing in human cells. *eLife* **2013**, (2013).

25. Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. & Zhang, F. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819–823 (2013).

26. Chen, X., Xu, F., Zhu, C., Ji, J., Zhou, X., Feng, X. & Guang, S. Dual sgRNA-directed gene knockout using CRISPR/Cas9 technology in Caenorhabditis elegans. *Sci Rep* **4**, 7581 (2015).

27. O'Connell, M. R., Oakes, B. L., Sternberg, S. H., East-Seletsky, A., Kaplan, M. & Doudna, J. A. Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* **516**, 263–266 (2014).

28. Konermann, S., Lotfy, P., Brideau, N. J., Oki, J., Shokhirev, M. N. & Hsu, P. D. Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. *Cell* **173**, 665-676.e14 (2018).

29. Price, A. A., Sampson, T. R., Ratner, H. K., Grakoui, A. & Weiss, D. S. Cas9-mediated targeting of viral RNA in eukaryotic cells. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6164–6169 (2015).

30. Freije, C. A., Myhrvold, C., Boehm, C. K., Lin, A. E., Welch, N. L., Carter, A., Metsky, H. C., Luo, C. Y., Abudayyeh, O. O., Gootenberg, J. S., Yozwiak, N. L., Zhang, F. & Sabeti, P. C. Programmable Inhibition and Detection of RNA Viruses Using Cas13. *Molecular Cell* **76**, 826-837.e11 (2019).

31. Zhou, H., Su, J., Hu, X., Zhou, C., Li, H., Chen, Z., Xiao, Q., Wang, B., Wu, W., Sun, Y., Zhou, Y., Tang, C., Liu, F., Wang, L., Feng, C., Liu, M., Li, S., Zhang, Y., Xu, H., Yao, H., Shi, L. & Yang, H. Glia-to-Neuron Conversion by CRISPR-CasRx Alleviates Symptoms of Neurological Disease in Mice. *Cell* **181**, 590-603.e16 (2020).

32. DeWitt, M. A., Corn, J. E. & Carroll, D. Genome editing via delivery of Cas9 ribonucleoprotein. *Methods* **121–122**, 9–15 (2017).

33. Hacein-Bey-Abina, S., Garrigue, A., Wang, G. P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., Asnafi, V., MacIntyre, E., Dal Cortivo, L., Radford, I., Brousse, N., Sigaux, F., Moshous, D., Hauer, J., Borkhardt, A., Belohradsky, B. H., Wintergerst, U., Velez, M. C., Leiva, L., Sorensen, R., Wulffraat, N., Blanche, S., Bushman, F. D., Fischer, A. & Cavazzana-Calvo, M. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J. Clin. Invest.* **118**, 3132–3142 (2008).

34. Naldini, L. Genetic engineering of hematopoiesis: current stage of clinical translation and future perspectives. *EMBO Mol Med* **11**, (2019).

35. De Ravin, S. S., Li, L., Wu, X., Choi, U., Allen, C., Koontz, S., Lee, J., Theobald-Whiting, N., Chu, J., Garofalo, M., Sweeney, C., Kardava, L., Moir, S., Viley, A., Natarajan, P., Su, L., Kuhns, D., Zarember, K. A., Peshwa, M. V. & Malech, H. L. CRISPR-Cas9 gene repair of hematopoietic stem cells from patients with X-linked chronic granulomatous disease. *Sci. Transl. Med.* **9**, eaah3480 (2017).

36. Dever, D. P., Bak, R. O., Reinisch, A., Camarena, J., Washington, G., Nicolas, C. E., Pavel-Dinu, M., Saxena, N., Wilkens, A. B., Mantri, S., Uchida, N., Hendel, A., Narla, A., Majeti, R., Weinberg, K. I. & Porteus, M. H. CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. *Nature* **539**, 384–389 (2016).

37. Xu, J., Peng, C., Sankaran, V. G., Shao, Z., Esrick, E. B., Chong, B. G., Ippolito, G. C., Fujiwara, Y., Ebert, B. L., Tucker, P. W. & Orkin, S. H. Correction of Sickle Cell Disease in Adult Mice by Interference with Fetal Hemoglobin Silencing. *Science* **334**, 993–996 (2011).

38. Sankaran, V. G., Menne, T. F., Xu, J., Akie, T. E., Lettre, G., Van Handel, B., Mikkola, H. K. A., Hirschhorn, J. N., Cantor, A. B. & Orkin, S. H. Human Fetal Hemoglobin Expression Is Regulated by the Developmental Stage-Specific Repressor *BCL11A*. *Science* **322**, 1839–1842 (2008).

39. Srivastava, A. In vivo tissue-tropism of adeno-associated viral vectors. *Current Opinion in Virology* **21**, 75–80 (2016).

40. Kanaan, N. M., Sellnow, R. C., Boye, S. L., Coberly, B., Bennett, A., Agbandje-McKenna, M., Sortwell, C. E., Hauswirth, W. W., Boye, S. E. & Manfredsson, F. P. Rationally Engineered AAV Capsids Improve Transduction and Volumetric Spread in the CNS. *Molecular Therapy - Nucleic Acids* **8**, 184–197 (2017).

41. Rosario, A. M., Cruz, P. E., Ceballos-Diaz, C., Strickland, M. R., Siemienski, Z., Pardo, M., Schob, K.-L., Li, A., Aslanidi, G. V., Srivastava, A., Golde, T. E. & Chakrabarty, P. Microglia-specific targeting by novel capsid-modified AAV6 vectors. *Molecular Therapy - Methods & Clinical Development* **3**, 16026 (2016).

42. Bowles, D. E., McPhee, S. W., Li, C., Gray, S. J., Samulski, J. J., Camp, A. S., Li, J., Wang, B., Monahan, P. E., Rabinowitz, J. E., Grieger, J. C., Govindasamy, L., Agbandje-McKenna, M., Xiao, X. & Samulski, R. J. Phase 1 Gene Therapy for Duchenne Muscular Dystrophy Using a Translational Optimized AAV Vector. *Molecular Therapy* **20**, 443–455 (2012).

43. Li, W., Asokan, A., Wu, Z., Van Dyke, T., DiPrimio, N., Johnson, J. S., Govindaswamy, L., Agbandje-McKenna, M., Leichtle, S., Eugene Redmond Jr, D., McCown, T. J., Petermann, K. B., Sharpless, N. E. & Samulski, R. J. Engineering and Selection of Shuffled AAV Genomes: A New Strategy for Producing Targeted Biological Nanoparticles. *Molecular Therapy* **16**, 1252–1260 (2008).

44. Maheshri, N., Koerber, J. T., Kaspar, B. K. & Schaffer, D. V. Directed evolution of adeno-associated virus yields enhanced gene delivery vectors. *Nat Biotechnol* **24**, 198–204 (2006).

45. Asuri, P., Bartel, M. A., Vazin, T., Jang, J.-H., Wong, T. B. & Schaffer, D. V. Directed Evolution of Adeno-associated Virus for Enhanced Gene Delivery and Gene Targeting in Human Pluripotent Stem Cells. *Molecular Therapy* **20**, 329–338 (2012).

46. Li, W., Zhang, L., Johnson, J. S., Zhijian, W., Grieger, J. C., Ping-Jie, X., Drouin, L. M., Agbandje-McKenna, M., Pickles, R. J. & Samulski, R. J. Generation of Novel AAV Variants by Directed Evolution for Improved CFTR Delivery to Human Ciliated Airway Epithelium. *Molecular Therapy* **17**, 2067–2077 (2009).

47. Wang, L., Calcedo, R., Bell, P., Lin, J., Grant, R. L., Siegel, D. L. & Wilson, J. M. Impact of Pre-Existing Immunity on Gene Transfer to Nonhuman Primate Liver with Adeno-Associated Virus 8 Vectors. *Human Gene Therapy* **22**, 1389–1401 (2011).

48. Rivière, C., Danos, O. & Douar, A. M. Long-term expression and repeated administration of AAV type 1, 2 and 5 vectors in skeletal muscle of immunocompetent adult mice. *Gene Ther* **13**, 1300–1308 (2006).

49. Tse, L. V., Klinc, K. A., Madigan, V. J., Castellanos Rivera, R. M., Wells, L. F., Havlik, L. P., Smith, J. K., Agbandje-McKenna, M. & Asokan, A. Structure-guided evolution of antigenically distinct adeno-associated virus variants for immune evasion. *Proc. Natl. Acad. Sci. U.S.A.* **114**, (2017).

50. Lochrie, M. A., Tatsuno, G. P., Christie, B., McDonnell, J. W., Zhou, S., Surosky, R., Pierce, G. F. & Colosi, P. Mutations on the External Surfaces of Adeno-AssociatedVirus Type 2 Capsids That Affect Transduction andNeutralization. *J Virol* **80**, 821–834 (2006).

51. Giles, A. R., Govindasamy, L., Somanathan, S. & Wilson, J. M. Mapping an Adeno-associated Virus 9-Specific Neutralizing Epitope To Develop Next-Generation Gene Delivery Vectors. *J Virol* **92**, e01011-18 (2018).

52. Dauletbekov, D. L., Pfromm, J. K., Fritz, A. K. & Fischer, M. D. Innate Immune Response Following AAV Administration. in *Retinal Degenerative Diseases* (eds. Bowes Rickman, C., Grimm, C., Anderson, R. E., Ash, J. D., LaVail, M. M. & Hollyfield, J. G.) vol. 1185 165–168 (Springer International Publishing, 2019).

53. Butterfield, J. S. S., Biswas, M., Shirley, J. L., Kumar, S. R. P., Sherman, A., Terhorst, C., Ling, C. & Herzog, R. W. TLR9-Activating CpG-B ODN but Not TLR7 Agonists Triggers Antibody Formation to Factor IX in Muscle Gene Transfer. *Human Gene Therapy Methods* **30**, 81–92 (2019).

54. Ashley, S. N., Somanathan, S., Giles, A. R. & Wilson, J. M. TLR9 signaling mediates adaptive immunity following systemic AAV gene therapy. *Cellular Immunology* **346**, 103997 (2019).

55. Faust, S. M., Bell, P., Cutler, B. J., Ashley, S. N., Zhu, Y., Rabinowitz, J. E. & Wilson, J. M. CpG-depleted adeno-associated virus vectors evade immune detection. *J. Clin. Invest.* **123**, 2994–3001 (2013).

56. Shao, W., Earley, L. F., Chai, Z., Chen, X., Sun, J., He, T., Deng, M., Hirsch, M. L., Ting, J., Samulski, R. J. & Li, C. Double-stranded RNA innate immune response activation from long-term adeno-associated virus vector transduction. *JCI Insight* **3**, e120474 (2018).

57. Rabinowitz, J., Chan, Y. K. & Samulski, R. J. Adeno-associated Virus (AAV) versus Immune Response. *Viruses* **11**, 102 (2019).

58. Calcedo, R., Chichester, J. A. & Wilson, J. M. Assessment of Humoral, Innate, and T-Cell Immune Responses to Adeno-Associated Virus Vectors. *Human Gene Therapy Methods* **29**, 86–95 (2018).

59. Xiong, W., Wu, D. M., Xue, Y., Wang, S. K., Chung, M. J., Ji, X., Rana, P., Zhao, S. R., Mai, S. & Cepko, C. L. AAV *cis* -regulatory sequences are correlated with ocular toxicity. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 5785–5794 (2019).

60. Horowitz, E. D., Weinberg, M. S. & Asokan, A. Glycated AAV Vectors: Chemical Redirection of Viral Tissue Tropism. *Bioconjugate Chem.* **22**, 529–532 (2011).

61. Katrekar, D., Moreno, A. M., Chen, G., Worlikar, A. & Mali, P. Oligonucleotide conjugated multi-functional adeno-associated viruses. *Sci Rep* **8**, 3589 (2018).

62. Kelemen, R. E., Mukherjee, R., Cao, X., Erickson, S. B., Zheng, Y. & Chatterjee, A. A Precise Chemical Strategy To Alter the Receptor Specificity of the Adeno-Associated Virus. *Angew. Chem. Int. Ed.* **55**, 10645–10649 (2016).

63. Hudry, E., Martin, C., Gandhi, S., György, B., Scheffer, D. I., Mu, D., Merkel, S. F., Mingozzi, F., Fitzpatrick, Z., Dimant, H., Masek, M., Ragan, T., Tan, S., Brisson, A. R., Ramirez, S. H., Hyman, B. T. & Maguire, C. A. Exosome-associated AAV vector as a robust and convenient neuroscience tool. *Gene Ther* **23**, 380–392 (2016).

64. Meliani, A., Boisgerault, F., Fitzpatrick, Z., Marmier, S., Leborgne, C., Collaud, F., Simon Sola, M., Charles, S., Ronzitti, G., Vignaud, A., van Wittenberghe, L., Marolleau, B., Jouen, F., Tan, S., Boyer, O., Christophe, O., Brisson, A. R., Maguire, C. A. & Mingozzi, F. Enhanced liver gene transfer and evasion of preexisting humoral immunity with exosome-enveloped AAV vectors. *Blood Advances* **1**, 2019–2031 (2017).

65. György, B., Sage, C., Indzhykulian, A. A., Scheffer, D. I., Brisson, A. R., Tan, S., Wu, X., Volak, A., Mu, D., Tamvakologos, P. I., Li, Y., Fitzpatrick, Z., Ericsson, M., Breakefield, X. O., Corey, D. P. & Maguire, C. A. Rescue of Hearing by Gene Delivery to Inner-Ear Hair Cells Using Exosome-Associated AAV. *Molecular Therapy* **25**, 379–391 (2017).

66. Wassmer, S. J., Carvalho, L. S., György, B., Vandenberghe, L. H. & Maguire, C. A. Exosome-associated AAV2 vector mediates robust gene delivery into the murine retina upon intravitreal injection. *Sci Rep* **7**, 45329 (2017).

67. Schiller, L. T., Lemus-Diaz, N., Rinaldi Ferreira, R., Böker, K. O. & Gruber, J. Enhanced Production of Exosome-Associated AAV by Overexpression of the Tetraspanin CD9. *Molecular Therapy - Methods & Clinical Development* **9**, 278–287 (2018).

68. Akinc, A., Maier, M. A., Manoharan, M., Fitzgerald, K., Jayaraman, M., Barros, S., Ansell, S., Du, X., Hope, M. J., Madden, T. D., Mui, B. L., Semple, S. C., Tam, Y. K., Ciufolini, M., Witzigmann, D., Kulkarni, J. A., van der Meel, R. & Cullis, P. R. The Onpattro story and the clinical translation of nanomedicines containing nucleic acid-based drugs. *Nat. Nanotechnol.* **14**, 1084–1087 (2019).

69. Semple, S. C., Akinc, A., Chen, J., Sandhu, A. P., Mui, B. L., Cho, C. K., Sah, D. W. Y., Stebbing, D., Crosley, E. J., Yaworski, E., Hafez, I. M., Dorkin, J. R., Qin, J., Lam, K., Rajeev, K. G., Wong, K. F., Jeffs, L. B., Nechev, L., Eisenhardt, M. L., Jayaraman, M., Kazem, M., Maier, M. A., Srinivasulu, M., Weinstein, M. J., Chen, Q., Alvarez, R., Barros, S. A., De, S., Klimuk, S. K., Borland, T., Kosovrasti, V., Cantley, W. L., Tam, Y. K., Manoharan, M., Ciufolini, M. A., Tracy, M. A., de Fougerolles, A., MacLachlan, I., Cullis, P. R., Madden, T. D. & Hope, M. J. Rational design of cationic lipids for siRNA delivery. *Nat Biotechnol* **28**, 172–176 (2010).

70. Mishra, S., Webster, P. & Davis, M. E. PEGylation significantly affects cellular uptake and intracellular trafficking of non-viral gene delivery particles. *European Journal of Cell Biology* **83**, 97–111 (2004).

71. Finn, J. D., Smith, A. R., Patel, M. C., Shaw, L., Youniss, M. R., van Heteren, J., Dirstine, T., Ciullo, C., Lescarbeau, R., Seitzer, J., Shah, R. R., Shah, A., Ling, D., Growe, J., Pink, M., Rohde, E., Wood, K. M., Salomon, W. E., Harrington, W. F., Dombrowski, C., Strapps, W. R., Chang, Y. & Morrissey, D. V. A Single Administration of CRISPR/Cas9 Lipid Nanoparticles Achieves Robust and Persistent In Vivo Genome Editing. *Cell Reports* **22**, 2227–2235 (2018).

72. Jiang, C., Mei, M., Li, B., Zhu, X., Zu, W., Tian, Y., Wang, Q., Guo, Y., Dong, Y. & Tan, X. A non-viral CRISPR/Cas9 delivery system for therapeutically targeting HBV DNA and pcsk9 in vivo. *Cell Res* **27**, 440–443 (2017).

73. Yin, H., Song, C.-Q., Dorkin, J. R., Zhu, L. J., Li, Y., Wu, Q., Park, A., Yang, J., Suresh, S., Bizhanova, A., Gupta, A., Bolukbasi, M. F., Walsh, S., Bogorad, R. L., Gao, G., Weng, Z., Dong, Y., Koteliansky, V., Wolfe, S. A., Langer, R., Xue, W. & Anderson, D. G. Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nat Biotechnol* **34**, 328–333 (2016).

74. Rensen, P. C. N., de Vrueh, R. L. A., Kuiper, J., Bijsterbosch, M. K., Biessen, E. A. L. & van Berkel, T. J. C. Recombinant lipoproteins: lipoprotein-like lipid particles for drug targeting. *Advanced Drug Delivery Reviews* **47**, 251–276 (2001).

75. Lee, K., Conboy, M., Park, H. M., Jiang, F., Kim, H. J., Dewitt, M. A., Mackley, V. A., Chang, K., Rao, A., Skinner, C., Shobha, T., Mehdipour, M., Liu, H., Huang, W., Lan, F., Bray, N. L., Li, S., Corn, J. E., Kataoka, K., Doudna, J. A., Conboy, I. & Murthy, N. Nanoparticle delivery of Cas9 ribonucleoprotein and donor DNA in vivo induces homology-directed DNA repair. *Nat Biomed Eng* **1**, 889–901 (2017).

76. Ding, Q. Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing. *Circ. Res.* **115**, (2014).

77. Lee, B., Lee, K., Panda, S., Gonzales-Rojas, R., Chong, A., Bugay, V., Park, H. M., Brenner, R., Murthy, N. & Lee, H. Y. Nanoparticle delivery of CRISPR into the brain rescues a mouse model of fragile X syndrome from exaggerated repetitive behaviours. *Nat Biomed Eng* **2**, 497–507 (2018).

78. Chen, G., Abdeen, A. A., Wang, Y., Shahi, P. K., Robertson, S., Xie, R., Suzuki, M., Pattnaik, B. R., Saha, K. & Gong, S. A biodegradable nanocapsule delivers a Cas9 ribonucleoprotein complex for in vivo genome editing. *Nat. Nanotechnol.* **14**, 974–980 (2019).

79. Miyata, K., Oba, M., Nakanishi, M., Fukushima, S., Yamasaki, Y., Koyama, H., Nishiyama, N. & Kataoka, K. Polyplexes from Poly(aspartamide) Bearing 1,2-Diaminoethane Side Chains Induce pH-Selective, Endosomal Membrane Destabilization with Amplified Transfection and Negligible Cytotoxicity. *J. Am. Chem. Soc.* **130**, 16287–16294 (2008).

80. Park, H. M., Liu, H., Wu, J., Chong, A., Mackley, V., Fellmann, C., Rao, A., Jiang, F., Chu, H., Murthy, N. & Lee, K. Extension of the crRNA enhances Cpf1 gene editing in vitro and in vivo. *Nat Commun* **9**, 3313 (2018).

81. Schmeer, M. & Schleef, M. Pharmaceutical Grade Large-Scale Plasmid DNA Manufacturing Process. in *DNA Vaccines* (eds. Rinaldi, M., Fioretti, D. & Iurescia, S.) vol. 1143 219–240 (Springer New York, 2014).

82. Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J. L., Vázquez, E. & Villaverde, A. Microbial factories for recombinant pharmaceuticals. *Microb Cell Fact* **8**, 17 (2009).

83. Zhu, J. Mammalian cell protein expression for biopharmaceutical production. *Biotechnology Advances* **30**, 1158–1170 (2012).

84. Sanchez-Garcia, L., Martín, L., Mangues, R., Ferrer-Miralles, N., Vázquez, E. & Villaverde, A. Recombinant pharmaceuticals from microbial cells: a 2015 update. *Microb Cell Fact* **15**, 33 (2016).

85. Mamat, U., Wilke, K., Bramhill, D., Schromm, A. B., Lindner, B., Kohl, T. A., Corchero, J. L., Villaverde, A., Schaffer, L., Head, S. R., Souvignier, C., Meredith, T. C. & Woodard, R. W. Detoxifying Escherichia coli for endotoxin-free production of recombinant proteins. *Microb Cell Fact* **14**, 57 (2015).

86. Nieß, A., Siemann-Herzberg, M. & Takors, R. Protein production in Escherichia coli is guided by the trade-off between intracellular substrate availability and energy cost. *Microb Cell Fact* **18**, 8 (2019).

87. Sandberg, T. E., Salazar, M. J., Weng, L. L., Palsson, B. O. & Feist, A. M. The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. *Metabolic Engineering* **56**, 1–16 (2019).

88. Qiao, J., Li, W., Lin, S., Sun, W., Ma, L. & Liu, Y. Co-expression of Cas9 and single-guided RNAs in Escherichia coli streamlines production of Cas9 ribonucleoproteins. *Commun Biol* **2**, 161 (2019).

89. Tripathi, N. K. & Shrivastava, A. Recent Developments in Bioprocessing of Recombinant Proteins: Expression Hosts and Process Development. *Front. Bioeng. Biotechnol.* **7**, 420 (2019).

90. Karst, D. J., Steinebach, F. & Morbidelli, M. Continuous integrated manufacturing of therapeutic proteins. *Current Opinion in Biotechnology* **53**, 76–84 (2018).

91. Seki, A. & Rutz, S. Optimized RNP transfection for highly efficient CRISPR/Cas9-mediated gene knockout in primary T cells. *Journal of Experimental Medicine* **215**, 985–997 (2018).

92. Hendel, A., Bak, R. O., Clark, J. T., Kennedy, A. B., Ryan, D. E., Roy, S., Steinfeld, I., Lunstad, B. D., Kaiser, R. J., Wilkens, A. B., Bacchetta, R., Tsalenko, A., Dellinger, D.,

Bruhn, L. & Porteus, M. H. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat Biotechnol* **33**, 985–989 (2015).

93. Schumann, K., Lin, S., Boyer, E., Simeonov, D. R., Subramaniam, M., Gate, R. E., Haliburton, G. E., Ye, C. J., Bluestone, J. A., Doudna, J. A. & Marson, A. Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10437–10442 (2015).

94. Kelley, M. L., Strezoska, Ž., He, K., Vermeulen, A. & Smith, A. van B. Versatility of chemically synthesized guide RNAs for CRISPR-Cas9 genome editing. *Journal of Biotechnology* **233**, 74–83 (2016).

95. Schubert, M. S., Cedrone, E., Neun, B., Behlke, M. A. & Dobrovolskaia, M. A. Chemical Modification of CRISPR gRNAs Eliminate type I Interferon Responses in Human Peripheral Blood Mononuclear Cells. *J Cytokine Biol* **03**, (2018).

96. Basila, M., Kelley, M. L. & Smith, A. van B. Minimal 2'-O-methyl phosphorothioate linkage modification pattern of synthetic guide RNAs for increased stability and efficient CRISPR-Cas9 gene editing avoiding cellular toxicity. *PLoS ONE* **12**, e0188593 (2017).

97. Scott, T., Soemardy, C. & Morris, K. V. Development of a Facile Approach for Generating Chemically Modified CRISPR/Cas9 RNA. *Molecular Therapy - Nucleic Acids* **19**, 1176–1185 (2020).

98. Xu, L., Wang, J., Liu, Y., Xie, L., Su, B., Mou, D., Wang, L., Liu, T., Wang, X., Zhang, B., Zhao, L., Hu, L., Ning, H., Zhang, Y., Deng, K., Liu, L., Lu, X., Zhang, T., Xu, J., Li, C., Wu, H., Deng, H. & Chen, H. CRISPR-Edited Stem Cells in a Patient with HIV and Acute Lymphocytic Leukemia. *N Engl J Med* **381**, 1240–1247 (2019).

99. Allay, J. A., Sleep, S., Long, S., Tillman, D. M., Clark, R., Carney, G., Fagone, P., McIntosh, J. H., Nienhuis, A. W., Davidoff, A. M., Nathwani, A. C. & Gray, J. T. Good Manufacturing Practice Production of Self-Complementary Serotype 8 Adeno-Associated Viral Vector for a Hemophilia B Clinical Trial. *Human Gene Therapy* **22**, 595–604 (2011).

100. Powers, A. D., Piras, B. A., Clark, R. K., Lockey, T. D. & Meagher, M. M. Development and Optimization of AAV hFIX Particles by Transient Transfection in an iCELLis ® Fixed-Bed Bioreactor. *Human Gene Therapy Methods* **27**, 112–121 (2016).

101. Grieger, J. C., Soltys, S. M. & Samulski, R. J. Production of Recombinant Adeno-associated Virus Vectors Using Suspension HEK293 Cells and Continuous Harvest of Vector From the Culture Media for GMP FIX and FLT1 Clinical Vector. *Molecular Therapy* **24**, 287–297 (2016).

102. Schmeer, M., Buchholz, T. & Schleef, M. Plasmid DNA Manufacturing for Indirect and Direct Clinical Applications. *Human Gene Therapy* **28**, 856–861 (2017).

103. Flotte, T. R., Trapnell, B. C., Humphries, M., Carey, B., Calcedo, R., Rouhani, F., Campbell-Thompson, M., Yachnis, A. T., Sandhaus, R. A., McElvaney, N. G., Mueller, C.,

Messina, L. M., Wilson, J. M., Brantly, M., Knop, D. R., Ye, G. & Chulay, J. D. Phase 2 Clinical Trial of a Recombinant Adeno-Associated Viral Vector Expressing α_1 - Antitrypsin: Interim Results. *Human Gene Therapy* **22**, 1239–1247 (2011).

104. Thomas, D. L., Wang, L., Niamke, J., Liu, J., Kang, W., Scotti, M. M., Ye, G.-J., Veres, G. & Knop, D. R. Scalable Recombinant Adeno-Associated Virus Production Using Recombinant Herpes Simplex Virus Type 1 Coinfection of Suspension-Adapted Mammalian Cells. *Human Gene Therapy* **20**, 861–870 (2009).

105. Adamson-Small, L., Potter, M., Falk, D. J., Cleaver, B., Byrne, B. J. & Clément, N. A scalable method for the production of high-titer and high-quality adeno-associated type 9 vectors using the HSV platform. *Molecular Therapy - Methods & Clinical Development* **3**, 16031 (2016).

106. Urabe, M., Ding, C. & Kotin, R. M. Insect Cells as a Factory to Produce Adeno-Associated Virus Type 2 Vectors. *Human Gene Therapy* **13**, 1935–1943 (2002).

107. Clément, N. & Grieger, J. C. Manufacturing of recombinant adeno-associated viral vectors for clinical trials. *Molecular Therapy - Methods & Clinical Development* **3**, 16002 (2016).

108. Smith, R. H., Levy, J. R. & Kotin, R. M. A Simplified Baculovirus-AAV Expression Vector System Coupled With One-step Affinity Purification Yields High-titer rAAV Stocks From Insect Cells. *Molecular Therapy* **17**, 1888–1896 (2009).

109. Rumachik, N. G., Malaker, S. A., Poweleit, N., Maynard, L. H., Adams, C. M., Leib, R. D., Cirolia, G., Thomas, D., Stamnes, S., Holt, K., Sinn, P., May, A. P. & Paulk, N. K. *Methods Matter -- Standard Production Platforms for Recombinant AAV Produce Chemically and Functionally Distinct Vectors*. http://biorxiv.org/lookup/doi/10.1101/640169 (2019) doi:10.1101/640169.

110. Wright, J. F. & Zelenaia, O. Vector Characterization Methods for Quality Control Testing of Recombinant Adeno-Associated Viruses. in *Viral Vectors for Gene Therapy* (eds. Merten, O.-W. & Al-Rubeai, M.) vol. 737 247–278 (Humana Press, 2011).

111. Ylä-Herttuala, S. Endgame: Glybera Finally Recommended for Approval as the First Gene Therapy Drug in the European Union. *Molecular Therapy* **20**, 1831–1832 (2012).

112. Flanigan, K. M., Campbell, K., Viollet, L., Wang, W., Gomez, A. M., Walker, C. M. & Mendell, J. R. Anti-Dystrophin T Cell Responses in Duchenne Muscular Dystrophy: Prevalence and a Glucocorticoid Treatment Effect. *Human Gene Therapy* **24**, 797–806 (2013).

113. Mendell, J. R., Campbell, K., Rodino-Klapac, L., Sahenk, Z., Shilling, C., Lewis, S., Bowles, D., Gray, S., Li, C., Galloway, G., Malik, V., Coley, B., Clark, K. R., Li, J., Xiao, X., Samulski, J., McPhee, S. W., Samulski, R. J. & Walker, C. M. Dystrophin Immunity in Duchenne's Muscular Dystrophy. *N Engl J Med* **363**, 1429–1437 (2010).

114. Tabebordbar, M., Zhu, K., Cheng, J. K. W., Chew, W. L., Widrick, J. J., Yan, W. X., Maesner, C., Wu, E. Y., Xiao, R., Ran, F. A., Cong, L., Zhang, F., Vandenberghe, L. H., Church, G. M. & Wagers, A. J. In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* **351**, 407–411 (2016).

115. Bengtsson, N. E., Hall, J. K., Odom, G. L., Phelps, M. P., Andrus, C. R., Hawkins, R. D., Hauschka, S. D., Chamberlain, J. R. & Chamberlain, J. S. Muscle-specific CRISPR/Cas9 dystrophin gene editing ameliorates pathophysiology in a mouse model for Duchenne muscular dystrophy. *Nat Commun* **8**, 14454 (2017).

116. Zhang, Y., Long, C., Li, H., McAnally, J. R., Baskin, K. K., Shelton, J. M., Bassel-Duby, R. & Olson, E. N. CRISPR-Cpf1 correction of muscular dystrophy mutations in human cardiomyocytes and mice. *Sci. Adv.* **3**, e1602814 (2017).

117. Ryu, S.-M., Koo, T., Kim, K., Lim, K., Baek, G., Kim, S.-T., Kim, H. S., Kim, D., Lee, H., Chung, E. & Kim, J.-S. Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy. *Nat Biotechnol* **36**, 536–539 (2018).

118. Le Guiner, C., Servais, L., Montus, M., Larcher, T., Fraysse, B., Moullec, S., Allais, M., François, V., Dutilleul, M., Malerba, A., Koo, T., Thibaut, J.-L., Matot, B., Devaux, M., Le Duff, J., Deschamps, J.-Y., Barthelemy, I., Blot, S., Testault, I., Wahbi, K., Ederhy, S., Martin, S., Veron, P., Georger, C., Athanasopoulos, T., Masurier, C., Mingozzi, F., Carlier, P., Gjata, B., Hogrel, J.-Y., Adjali, O., Mavilio, F., Voit, T., Moullier, P. & Dickson, G. Long-term microdystrophin gene therapy is effective in a canine model of Duchenne muscular dystrophy. *Nat Commun* **8**, 16105 (2017).

119. Nelson, C. E., Wu, Y., Gemberling, M. P., Oliver, M. L., Waller, M. A., Bohning, J. D., Robinson-Hamm, J. N., Bulaklak, K., Castellanos Rivera, R. M., Collier, J. H., Asokan, A. & Gersbach, C. A. Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy. *Nat Med* **25**, 427–432 (2019).

120. Amoasii, L., Hildyard, J. C. W., Li, H., Sanchez-Ortiz, E., Mireault, A., Caballero, D., Harron, R., Stathopoulou, T.-R., Massey, C., Shelton, J. M., Bassel-Duby, R., Piercy, R. J. & Olson, E. N. Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy. *Science* **362**, 86–91 (2018).

121. Zhang, Y., Li, H., Min, Y.-L., Sanchez-Ortiz, E., Huang, J., Mireault, A. A., Shelton, J. M., Kim, J., Mammen, P. P. A., Bassel-Duby, R. & Olson, E. N. Enhanced CRISPR-Cas9 correction of Duchenne muscular dystrophy in mice by a self-complementary AAV delivery system. *Sci. Adv.* **6**, eaay6812 (2020).

122. Lee, S. H., Kim, Y. S., Nah, S. K., Kim, H. J., Park, H. Y., Yang, J. Y., Park, K. & Park, T. K. Transduction Patterns of Adeno-associated Viral Vectors in a Laser-Induced Choroidal Neovascularization Mouse Model. *Molecular Therapy - Methods & Clinical Development* **9**, 90–98 (2018).

123. Fauser, S., Luberichs, J. & Schüttauf, F. Genetic animal models for retinal degeneration. *Surv Ophthalmol* **47**, 357–367 (2002).

124. Acland, G. M., Aguirre, G. D., Ray, J., Zhang, Q., Aleman, T. S., Cideciyan, A. V., Pearce-Kelling, S. E., Anand, V., Zeng, Y., Maguire, A. M., Jacobson, S. G., Hauswirth, W. W. & Bennett, J. Gene therapy restores vision in a canine model of childhood blindness. *Nat Genet* **28**, 92–95 (2001).

125. Aguirre, G. K., Komáromy, A. M., Cideciyan, A. V., Brainard, D. H., Aleman, T. S., Roman, A. J., Avants, B. B., Gee, J. C., Korczykowski, M., Hauswirth, W. W., Acland, G. M., Aguirre, G. D. & Aguirre, G. K. Canine and Human Visual Cortex Intact and Responsive Despite Early Retinal Blindness from RPE65 Mutation. *PLoS Med* **4**, e230 (2007).

126. Maguire, A. M., Simonelli, F., Pierce, E. A., Pugh, E. N., Mingozzi, F., Bennicelli, J., Banfi, S., Marshall, K. A., Testa, F., Surace, E. M., Rossi, S., Lyubarsky, A., Arruda, V. R., Konkle, B., Stone, E., Sun, J., Jacobs, J., Dell'Osso, L., Hertle, R., Ma, J., Redmond, T. M., Zhu, X., Hauck, B., Zelenaia, O., Shindler, K. S., Maguire, M. G., Wright, J. F., Volpe, N. J., McDonnell, J. W., Auricchio, A., High, K. A. & Bennett, J. Safety and Efficacy of Gene Transfer for Leber's Congenital Amaurosis. *N Engl J Med* **358**, 2240–2248 (2008).

127. Hauswirth, W. W., Aleman, T. S., Kaushal, S., Cideciyan, A. V., Schwartz, S. B., Wang, L., Conlon, T. J., Boye, S. L., Flotte, T. R., Byrne, B. J. & Jacobson, S. G. Treatment of Leber Congenital Amaurosis Due to *RPE65* Mutations by Ocular Subretinal Injection of Adeno-Associated Virus Gene Vector: Short-Term Results of a Phase I Trial. *Human Gene Therapy* **19**, 979–990 (2008).

128. Bainbridge, J. W. B., Smith, A. J., Barker, S. S., Robbie, S., Henderson, R., Balaggan, K., Viswanathan, A., Holder, G. E., Stockman, A., Tyler, N., Petersen-Jones, S., Bhattacharya, S. S., Thrasher, A. J., Fitzke, F. W., Carter, B. J., Rubin, G. S., Moore, A. T. & Ali, R. R. Effect of Gene Therapy on Visual Function in Leber's Congenital Amaurosis. *N Engl J Med* **358**, 2231–2239 (2008).

129. Amado, D., Mingozzi, F., Hui, D., Bennicelli, J. L., Wei, Z., Chen, Y., Bote, E., Grant, R. L., Golden, J. A., Narfstrom, K., Syed, N. A., Orlin, S. E., High, K. A., Maguire, A. M. & Bennett, J. Safety and Efficacy of Subretinal Readministration of a Viral Vector in Large Animals to Treat Congenital Blindness. *Sci. Transl. Med.* **2**, (2010).

130. Annear, M. J., Bartoe, J. T., Barker, S. E., Smith, A. J., Curran, P. G., Bainbridge, J. W., Ali, R. R. & Petersen-Jones, S. M. Gene therapy in the second eye of RPE65-deficient dogs improves retinal function. *Gene Ther* **18**, 53–61 (2011).

131. Bennett, J., Ashtari, M., Wellman, J., Marshall, K. A., Cyckowski, L. L., Chung, D. C., McCague, S., Pierce, E. A., Chen, Y., Bennicelli, J. L., Zhu, X., Ying, G., Sun, J., Wright, J. F., Auricchio, A., Simonelli, F., Shindler, K. S., Mingozzi, F., High, K. A. & Maguire, A. M. AAV2 Gene Therapy Readministration in Three Adults with Congenital Blindness. *Sci. Transl. Med.* **4**, (2012).

132. Boye, S. E., Alexander, J. J., Boye, S. L., Witherspoon, C. D., Sandefer, K. J., Conlon, T. J., Erger, K., Sun, J., Ryals, R., Chiodo, V. A., Clark, M. E., Girkin, C. A., Hauswirth, W. W. & Gamlin, P. D. The Human Rhodopsin Kinase Promoter in an AAV5 Vector Confers

Rod- and Cone-Specific Expression in the Primate Retina. *Human Gene Therapy* **23**, 1101–1115 (2012).

133. Sun, X., Pawlyk, B., Xu, X., Liu, X., Bulgakov, O. V., Adamian, M., Sandberg, M. A., Khani, S. C., Tan, M.-H., Smith, A. J., Ali, R. R. & Li, T. Gene therapy with a promoter targeting both rods and cones rescues retinal degeneration caused by AIPL1 mutations. *Gene Ther* **17**, 117–131 (2010).

134. Beltran, W. A., Cideciyan, A. V., Lewin, A. S., Iwabe, S., Khanna, H., Sumaroka, A., Chiodo, V. A., Fajardo, D. S., Román, A. J., Deng, W.-T., Swider, M., Alemán, T. S., Boye, S. L., Genini, S., Swaroop, A., Hauswirth, W. W., Jacobson, S. G. & Aguirre, G. D. Gene therapy rescues photoreceptor blindness in dogs and paves the way for treating human X-linked retinitis pigmentosa. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 2132–2137 (2012).

135. Daiger, S. P., Bowne, S. J. & Sullivan, L. S. Genes and Mutations Causing Autosomal Dominant Retinitis Pigmentosa. *Cold Spring Harb Perspect Med* **5**, a017129 (2015).

136. Bakondi, B., Lv, W., Lu, B., Jones, M. K., Tsai, Y., Kim, K. J., Levy, R., Akhtar, A. A., Breunig, J. J., Svendsen, C. N. & Wang, S. In Vivo CRISPR/Cas9 Gene Editing Corrects Retinal Dystrophy in the S334ter-3 Rat Model of Autosomal Dominant Retinitis Pigmentosa. *Molecular Therapy* **24**, 556–563 (2016).

137. Courtney, D. G., Moore, J. E., Atkinson, S. D., Maurizi, E., Allen, E. H. A., Pedrioli, D. M. L., McLean, W. H. I., Nesbit, M. A. & Moore, C. B. T. CRISPR/Cas9 DNA cleavage at SNP-derived PAM enables both in vitro and in vivo KRT12 mutation-specific targeting. *Gene Ther* **23**, 108–112 (2016).

138. Li, P., Kleinstiver, B. P., Leon, M. Y., Prew, M. S., Navarro-Gomez, D., Greenwald, S. H., Pierce, E. A., Joung, J. K. & Liu, Q. Allele-Specific CRISPR-Cas9 Genome Editing of the Single-Base P23H Mutation for Rhodopsin-Associated Dominant Retinitis Pigmentosa. *The CRISPR Journal* **1**, 55–64 (2018).

139. Mao, H., James, T., Schwein, A., Shabashvili, A. E., Hauswirth, W. W., Gorbatyuk, M. S. & Lewin, A. S. AAV Delivery of Wild-Type Rhodopsin Preserves Retinal Function in a Mouse Model of Autosomal Dominant Retinitis Pigmentosa. *Human Gene Therapy* **22**, 567–575 (2011).

140. Heier, J. S., Kherani, S., Desai, S., Dugel, P., Kaushal, S., Cheng, S. H., Delacono, C., Purvis, A., Richards, S., Le-Halpere, A., Connelly, J., Wadsworth, S. C., Varona, R., Buggage, R., Scaria, A. & Campochiaro, P. A. Intravitreous injection of AAV2-sFLT01 in patients with advanced neovascular age-related macular degeneration: a phase 1, open-label trial. *The Lancet* **390**, 50–61 (2017).

141. Chung, S. H., Mollhoff, I. N., Nguyen, U., Nguyen, A., Stucka, N., Tieu, E., Manna, S., Meleppat, R. K., Zhang, P., Nguyen, E. L., Fong, J., Zawadzki, R. & Yiu, G. Factors Impacting Efficacy of AAV-Mediated CRISPR-Based Genome Editing for Treatment of Choroidal Neovascularization. *Molecular Therapy - Methods & Clinical Development* **17**, 409–417 (2020).

142. Chadderton, N., Palfi, A., Millington-Ward, S., Gobbo, O., Overlack, N., Carrigan, M., O'Reilly, M., Campbell, M., Ehrhardt, C., Wolfrum, U., Humphries, P., Kenna, P. F. & Jane Farrar, G. Intravitreal delivery of AAV-NDI1 provides functional benefit in a murine model of Leber hereditary optic neuropathy. *Eur J Hum Genet* **21**, 62–68 (2013).

143. Kotterman, M. A., Yin, L., Strazzeri, J. M., Flannery, J. G., Merigan, W. H. & Schaffer, D. V. Antibody neutralization poses a barrier to intravitreal adeno-associated viral vector gene delivery to non-human primates. *Gene Ther* **22**, 116–126 (2015).

144. MacLachlan, T. K., Lukason, M., Collins, M., Munger, R., Isenberger, E., Rogers, C., Malatos, S., DuFresne, E., Morris, J., Calcedo, R., Veres, G., Scaria, A., Andrews, L. & Wadsworth, S. Preclinical Safety Evaluation of AAV2-sFLT01— A Gene Therapy for Age-related Macular Degeneration. *Molecular Therapy* **19**, 326–334 (2011).

145. Seitz, I. P., Michalakis, S., Wilhelm, B., Reichel, F. F., Ochakovski, G. A., Zrenner, E., Ueffing, M., Biel, M., Wissinger, B., Bartz-Schmidt, K. U., Peters, T., Fischer, M. D., & for the RD-CURE Consortium. Superior Retinal Gene Transfer and Biodistribution Profile of Subretinal Versus Intravitreal Delivery of AAV8 in Nonhuman Primates. *Invest. Ophthalmol. Vis. Sci.* **58**, 5792 (2017).

146. Ochakovski, G. A., Bartz-Schmidt, K. U. & Fischer, M. D. Retinal Gene Therapy: Surgical Vector Delivery in the Translation to Clinical Trials. *Front. Neurosci.* **11**, (2017).

147. Takahashi, K., Igarashi, T., Miyake, K., Kobayashi, M., Yaguchi, C., Iijima, O., Yamazaki, Y., Katakai, Y., Miyake, N., Kameya, S., Shimada, T., Takahashi, H. & Okada, T. Improved Intravitreal AAV-Mediated Inner Retinal Gene Transduction after Surgical Internal Limiting Membrane Peeling in Cynomolgus Monkeys. *Molecular Therapy* **25**, 296–302 (2017).

148. Kay, C. N., Ryals, R. C., Aslanidi, G. V., Min, S. H., Ruan, Q., Sun, J., Dyka, F. M., Kasuga, D., Ayala, A. E., Van Vliet, K., Agbandje-McKenna, M., Hauswirth, W. W., Boye, S. L. & Boye, S. E. Targeting Photoreceptors via Intravitreal Delivery Using Novel, Capsid-Mutated AAV Vectors. *PLoS ONE* **8**, e62097 (2013).

149. D'Avola, D., López-Franco, E., Sangro, B., Pañeda, A., Grossios, N., Gil-Farina, I., Benito, A., Twisk, J., Paz, M., Ruiz, J., Schmidt, M., Petry, H., Harper, P., de Salamanca, R. E., Fontanellas, A., Prieto, J. & González-Aseguinolaza, G. Phase I open label liver-directed gene therapy clinical trial for acute intermittent porphyria. *Journal of Hepatology* **65**, 776–783 (2016).

150. Nathwani, A. C., Reiss, U. M., Tuddenham, E. G. D., Rosales, C., Chowdary, P., McIntosh, J., Della Peruta, M., Lheriteau, E., Patel, N., Raj, D., Riddell, A., Pie, J., Rangarajan, S., Bevan, D., Recht, M., Shen, Y.-M., Halka, K. G., Basner-Tschakarjan, E., Mingozzi, F., High, K. A., Allay, J., Kay, M. A., Ng, C. Y. C., Zhou, J., Cancio, M., Morton, C. L., Gray, J. T., Srivastava, D., Nienhuis, A. W. & Davidoff, A. M. Long-Term Safety and Efficacy of Factor IX Gene Therapy in Hemophilia B. *N Engl J Med* **371**, 1994–2004 (2014).

151. Jacobs, F., Gordts, S., Muthuramu, I. & De Geest, B. The Liver as a Target Organ for Gene Therapy: State of the Art, Challenges, and Future Perspectives. *Pharmaceuticals* **5**, 1372–1392 (2012).

152. Glover, D. J., Lipps, H. J. & Jans, D. A. Towards safe, non-viral therapeutic gene expression in humans. *Nat Rev Genet* **6**, 299–310 (2005).

153. Izembart, A., Aguado, E., Gauthier, O., Aubert, D., Moullier, P. & Ferry, N. *In Vivo* Retrovirus-Mediated Gene Transfer to the Liver of Dogs Results in Transient Expression and Induction of a Cytotoxic Immune Response. *Human Gene Therapy* **10**, 2917–2925 (1999).

154. Gao, G.-P., Alvira, M. R., Wang, L., Calcedo, R., Johnston, J. & Wilson, J. M. Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11854–11859 (2002).

155. Cooper, M., Nayak, S., Hoffman, B. E., Terhorst, C., Cao, O. & Herzog, R. W. Improved Induction of Immune Tolerance to Factor IX by Hepatic AAV-8 Gene Transfer. *Human Gene Therapy* **20**, 767–776 (2009).

156. Davidoff, A. M., Gray, J. T., Ng, C. Y. C., Zhang, Y., Zhou, J., Spence, Y., Bakar, Y. & Nathwani, A. C. Comparison of the ability of adeno-associated viral vectors pseudotyped with serotype 2, 5, and 8 capsid proteins to mediate efficient transduction of the liver in murine and nonhuman primate models. *Molecular Therapy* **11**, 875–888 (2005).

157. Calcedo, R., Vandenberghe, L. H., Gao, G., Lin, J. & Wilson, J. M. Worldwide Epidemiology of Neutralizing Antibodies to Adeno-Associated Viruses. *J INFECT DIS* **199**, 381–390 (2009).

158. Boutin, S., Monteilhet, V., Veron, P., Leborgne, C., Benveniste, O., Montus, M. F. & Masurier, C. Prevalence of Serum IgG and Neutralizing Factors Against Adeno-Associated Virus (AAV) Types 1, 2, 5, 6, 8, and 9 in the Healthy Population: Implications for Gene Therapy Using AAV Vectors. *Human Gene Therapy* **21**, 704–712 (2010).

159. Verdera, H. C., Kuranda, K. & Mingozzi, F. AAV Vector Immunogenicity in Humans: A Long Journey to Successful Gene Transfer. *Molecular Therapy* **28**, 723–746 (2020).

160. Wang, L., Yang, Y., Breton, C. A., White, J., Zhang, J., Che, Y., Saveliev, A., McMenamin, D., He, Z., Latshaw, C., Li, M. & Wilson, J. M. CRISPR/Cas9-mediated in vivo gene targeting corrects hemostasis in newborn and adult factor IX–knockout mice. *Blood* **133**, 2745–2752 (2019).

161. Chen, H., Shi, M., Gilam, A., Zheng, Q., Zhang, Y., Afrikanova, I., Li, J., Gluzman, Z., Jiang, R., Kong, L.-J. & Chen-Tsai, R. Y. Hemophilia A ameliorated in mice by CRISPR-based in vivo genome editing of human Factor VIII. *Sci Rep* **9**, 16838 (2019).

162. Wang, L., Yang, Y., Breton, C., Bell, P., Li, M., Zhang, J., Che, Y., Saveliev, A., He, Z., White, J., Latshaw, C., Xu, C., McMenamin, D., Yu, H., Morizono, H., Batshaw, M. L. &

Wilson, J. M. A mutation-independent CRISPR-Cas9–mediated gene targeting approach to treat a murine model of ornithine transcarbamylase deficiency. *Sci. Adv.* **6**, eaax5701 (2020).

163. Yang, Y., Wang, L., Bell, P., McMenamin, D., He, Z., White, J., Yu, H., Xu, C., Morizono, H., Musunuru, K., Batshaw, M. L. & Wilson, J. M. A dual AAV system enables the Cas9-mediated correction of a metabolic liver disease in newborn mice. *Nat Biotechnol* **34**, 334–338 (2016).

164. Kwon, J. H., Lee, Y. M., Cho, J.-H., Kim, G.-Y., Anduaga, J., Starost, M. F., Mansfield, B. C. & Chou, J. Y. Liver-directed gene therapy for murine glycogen storage disease type Ib. *Human Molecular Genetics* **26**, 4395–4405 (2017).

165. Jeyakumar, J., Kia, A., McIntosh, J., Verhoef, D., Kalcheva, P., Hosseini, P., Sheridan, R., Corbau, R. & Nathwani, A. Liver-directed gene therapy corrects Fabry disease in mice. *Molecular Genetics and Metabolism* **126**, S80 (2019).

166. Miranda, C. J., Canavese, M., Chisari, E., Pandya, J., Cocita, C., Portillo, M., McIntosh, J., Kia, A., Foley, J. H., Dane, A., Jeyakumar, J. M., Sheridan, R., Corbau, R. & Nathwani, A. C. Liver-Directed AAV Gene Therapy for Gaucher Disease. *Blood* **134**, 3354–3354 (2019).

167. Sands, M. S. & Davidson, B. L. Gene therapy for lysosomal storage diseases. *Molecular Therapy* **13**, 839–849 (2006).

168. Wraith, J. E., Scarpa, M., Beck, M., Bodamer, O. A., De Meirleir, L., Guffon, N., Meldgaard Lund, A., Malm, G., Van der Ploeg, A. T. & Zeman, J. Mucopolysaccharidosis type II (Hunter syndrome): a clinical review and recommendations for treatment in the era of enzyme replacement therapy. *Eur J Pediatr* **167**, 267–277 (2008).

169. Guffon, N., Bertrand, Y., Forest, I., Fouilhoux, A. & Froissart, R. Bone Marrow Transplantation in Children with Hunter Syndrome: Outcome after 7 to 17 Years. *The Journal of Pediatrics* **154**, 733–737 (2009).

170. Wyatt, K., Henley, W., Anderson, L., Anderson, R., Nikolaou, V., Stein, K., Klinger, L., Hughes, D., Waldek, S., Lachmann, R., Mehta, A., Vellodi, A. & Logan, S. The effectiveness and cost-effectiveness of enzyme and substrate replacement therapies: a longitudinal cohort study of people with lysosomal storage disorders. *Health Technol Assess* **16**, (2012).

171. Boelens, J. J., Wynn, R. F., O'Meara, A., Veys, P., Bertrand, Y., Souillet, G., Wraith, J. E., Fischer, A., Cavazzana-Calvo, M., Sykora, K. W., Sedlacek, P., Rovelli, A., Uiterwaal, C. S. P. M. & Wulffraat, N. Outcomes of hematopoietic stem cell transplantation for Hurler's syndrome in Europe: a risk factor analysis for graft failure. *Bone Marrow Transplant* **40**, 225–233 (2007).

172. Boelens, J. J., Orchard, P. J. & Wynn, R. F. Transplantation in inborn errors of metabolism: current considerations and future perspectives. *Br J Haematol* **167**, 293–303 (2014).

173. Aldenhoven, M., Boelens, J. & de Koning, T. J. The Clinical Outcome of Hurler Syndrome after Stem Cell Transplantation. *Biology of Blood and Marrow Transplantation* **14**, 485–498 (2008).

174. Ou, L., Przybilla, M. J., Tăbăran, A.-F., Overn, P., O'Sullivan, M. G., Jiang, X., Sidhu, R., Kell, P. J., Ory, D. S. & Whitley, C. B. A novel gene editing system to treat both Tay–Sachs and Sandhoff diseases. *Gene Ther* **27**, 226–236 (2020).

175. Ou, L., DeKelver, R. C., Rohde, M., Tom, S., Radeke, R., St. Martin, S. J., Santiago, Y., Sproul, S., Przybilla, M. J., Koniar, B. L., Podetz-Pedersen, K. M., Laoharawee, K., Cooksley, R. D., Meyer, K. E., Holmes, M. C., McIvor, R. S., Wechsler, T. & Whitley, C. B. ZFN-Mediated In Vivo Genome Editing Corrects Murine Hurler Syndrome. *Molecular Therapy* **27**, 178–187 (2019).

176. Schuh, R. S., Poletto, É., Pasqualim, G., Tavares, A. M. V., Meyer, F. S., Gonzalez, E. A., Giugliani, R., Matte, U., Teixeira, H. F. & Baldo, G. In vivo genome editing of mucopolysaccharidosis I mice using the CRISPR/Cas9 system. *Journal of Controlled Release* **288**, 23–33 (2018).

177. Zhang, H., Yang, B., Mu, X., Ahmed, S. S., Su, Q., He, R., Wang, H., Mueller, C., Sena-Esteves, M., Brown, R., Xu, Z. & Gao, G. Several rAAV Vectors Efficiently Cross the Blood–brain Barrier and Transduce Neurons and Astrocytes in the Neonatal Mouse Central Nervous System. *Molecular Therapy* **19**, 1440–1448 (2011).

178. Duque, S., Joussemet, B., Riviere, C., Marais, T., Dubreil, L., Douar, A.-M., Fyfe, J., Moullier, P., Colle, M.-A. & Barkats, M. Intravenous Administration of Self-complementary AAV9 Enables Transgene Delivery to Adult Motor Neurons. *Molecular Therapy* **17**, 1187–1196 (2009).

179. Samaranch, L., Salegio, E. A., San Sebastian, W., Kells, A. P., Foust, K. D., Bringas, J. R., Lamarre, C., Forsayeth, J., Kaspar, B. K. & Bankiewicz, K. S. Adeno-Associated Virus Serotype 9 Transduction in the Central Nervous System of Nonhuman Primates. *Human Gene Therapy* **23**, 382–389 (2012).

180. Passini, M. A. & Wolfe, J. H. Widespread Gene Delivery and Structure-Specific Patterns of Expression in the Brain after Intraventricular Injections of Neonatal Mice with an Adeno-Associated Virus Vector. *J Virol* **75**, 12382–12392 (2001).

181. Haurigot, V., Marcó, S., Ribera, A., Garcia, M., Ruzo, A., Villacampa, P., Ayuso, E., Añor, S., Andaluz, A., Pineda, M., García-Fructuoso, G., Molas, M., Maggioni, L., Muñoz, S., Motas, S., Ruberte, J., Mingozzi, F., Pumarola, M. & Bosch, F. Whole body correction of mucopolysaccharidosis IIIA by intracerebrospinal fluid gene therapy. *J. Clin. Invest.* **123**, 3254–3271 (2013).

182. Souweidane, M. M., Fraser, J. F., Arkin, L. M., Sondhi, D., Hackett, N. R., Kaminsky, S. M., Heier, L., Kosofsky, B. E., Worgall, S., Crystal, R. G. & Kaplitt, M. G. Gene therapy for late infantile neuronal ceroid lipofuscinosis: neurosurgical considerations: Clinical article. *PED* **6**, 115–122 (2010).

183. Meyer, K., Ferraiuolo, L., Schmelzer, L., Braun, L., McGovern, V., Likhite, S., Michels, O., Govoni, A., Fitzgerald, J., Morales, P., Foust, K. D., Mendell, J. R., Burghes, A. H. M. & Kaspar, B. K. Improving Single Injection CSF Delivery of AAV9-mediated Gene Therapy for SMA: A Dose–response Study in Mice and Nonhuman Primates. *Molecular Therapy* **23**, 477–487 (2015).

184. Federici, T., Taub, J. S., Baum, G. R., Gray, S. J., Grieger, J. C., Matthews, K. A., Handy, C. R., Passini, M. A., Samulski, R. J. & Boulis, N. M. Robust spinal motor neuron transduction following intrathecal delivery of AAV9 in pigs. *Gene Ther* **19**, 852–859 (2012).

185. Ragagnin, A. M. G., Shadfar, S., Vidal, M., Jamali, M. S. & Atkin, J. D. Motor Neuron Susceptibility in ALS/FTD. *Front. Neurosci.* **13**, 532 (2019).

186. Rosen, D. R., Siddique, T., Patterson, D., Figlewicz, D. A., Sapp, P., Hentati, A., Donaldson, D., Goto, J., O'Regan, J. P., Deng, H.-X., Rahmani, Z., Krizus, A., McKenna-Yasek, D., Cayabyab, A., Gaston, S. M., Berger, R., Tanzi, R. E., Halperin, J. J., Herzfeldt, B., Van den Bergh, R., Hung, W.-Y., Bird, T., Deng, G., Mulder, D. W., Smyth, C., Laing, N. G., Soriano, E., Pericak–Vance, M. A., Haines, J., Rouleau, G. A., Gusella, J. S., Horvitz, H. R. & Brown, R. H. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362**, 59–62 (1993).

187. Gaj, T., Ojala, D. S., Ekman, F. K., Byrne, L. C., Limsirichai, P. & Schaffer, D. V. In vivo genome editing improves motor function and extends survival in a mouse model of ALS. *Sci. Adv.* **3**, eaar3952 (2017).

188. Duan, W., Guo, M., Yi, L., Liu, Y., Li, Z., Ma, Y., Zhang, G., Liu, Y., Bu, H., Song, X. & Li, C. The deletion of mutant SOD1 via CRISPR/Cas9/sgRNA prolongs survival in an amyotrophic lateral sclerosis mouse model. *Gene Ther* **27**, 157–169 (2020).

189. Lim, C. K. W., Gapinske, M., Brooks, A. K., Woods, W. S., Powell, J. E., Zeballos C., M. A., Winter, J., Perez-Pinera, P. & Gaj, T. Treatment of a Mouse Model of ALS by In Vivo Base Editing. *Molecular Therapy* **28**, 1177–1189 (2020).

190. Ekman, F. K., Ojala, D. S., Adil, M. M., Lopez, P. A., Schaffer, D. V. & Gaj, T. CRISPR-Cas9-Mediated Genome Editing Increases Lifespan and Improves Motor Deficits in a Huntington's Disease Mouse Model. *Molecular Therapy - Nucleic Acids* **17**, 829–839 (2019).

191. György, B., Lööv, C., Zaborowski, M. P., Takeda, S., Kleinstiver, B. P., Commins, C., Kastanenka, K., Mu, D., Volak, A., Giedraitis, V., Lannfelt, L., Maguire, C. A., Joung, J. K., Hyman, B. T., Breakefield, X. O. & Ingelsson, M. CRISPR/Cas9 Mediated Disruption of the Swedish APP Allele as a Therapeutic Approach for Early-Onset Alzheimer's Disease. *Molecular Therapy - Nucleic Acids* **11**, 429–440 (2018).

192. Roca, C., Motas, S., Marcó, S., Ribera, A., Sánchez, V., Sánchez, X., Bertolin, J., León, X., Pérez, J., Garcia, M., Villacampa, P., Ruberte, J., Pujol, A., Haurigot, V. & Bosch, F.

Disease correction by AAV-mediated gene therapy in a new mouse model of mucopolysaccharidosis type IIID. *Human Molecular Genetics* **26**, 1535–1551 (2017).

193. Passini, M. A., Macauley, S. L., Huff, M. R., Taksir, T. V., Bu, J., Wu, I.-H., Piepenhagen, P. A., Dodge, J. C., Shihabuddin, L. S., O'Riordan, C. R., Schuchman, E. H. & Stewart, G. R. AAV Vector-Mediated Correction of Brain Pathology in a Mouse Model of Niemann–Pick A Disease. *Molecular Therapy* **11**, 754–762 (2005).

194. Ohno, M., Cole, S. L., Yasvoina, M., Zhao, J., Citron, M., Berry, R., Disterhoft, J. F. & Vassar, R. BACE1 gene deletion prevents neuron loss and memory deficits in 5XFAD APP/PS1 transgenic mice. *Neurobiology of Disease* **26**, 134–145 (2007).

195. McGreevy, J. W., Hakim, C. H., McIntosh, M. A. & Duan, D. Animal models of Duchenne muscular dystrophy: from basic mechanisms to gene therapy. *Disease Models & Mechanisms* **8**, 195–213 (2015).

196. Tu, Z., Yang, W., Yan, S., Guo, X. & Li, X.-J. CRISPR/Cas9: a powerful genetic engineering tool for establishing large animal models of neurodegenerative diseases. *Mol Neurodegeneration* **10**, 35 (2015).

197. Moretti, A., Fonteyne, L., Giesert, F., Hoppmann, P., Meier, A. B., Bozoglu, T., Baehr, A., Schneider, C. M., Sinnecker, D., Klett, K., Fröhlich, T., Rahman, F. A., Haufe, T., Sun, S., Jurisch, V., Kessler, B., Hinkel, R., Dirschinger, R., Martens, E., Jilek, C., Graf, A., Krebs, S., Santamaria, G., Kurome, M., Zakhartchenko, V., Campbell, B., Voelse, K., Wolf, A., Ziegler, T., Reichert, S., Lee, S., Flenkenthaler, F., Dorn, T., Jeremias, I., Blum, H., Dendorfer, A., Schnieke, A., Krause, S., Walter, M. C., Klymiuk, N., Laugwitz, K. L., Wolf, E., Wurst, W. & Kupatt, C. Somatic gene editing ameliorates skeletal and cardiac muscle failure in pig and human models of Duchenne muscular dystrophy. *Nat Med* **26**, 207–214 (2020).

198. Li, X.-J., Chang, R., Liu, X. & Li, S. Transgenic animal models for study of the pathogenesis of Huntington’s disease and therapy. *DDDT* 2179 (2015) doi:10.2147/DDDT.S58470.

199. Menon, K. P., Tieu, P. T. & Neufeld, E. F. Architecture of the canine IDUA gene and mutation underlying canine mucopolysaccharidosis I. *Genomics* **14**, 763–768 (1992).

200. Ellinwood, N. M., Vite, C. H. & Haskins, M. E. Gene therapy for lysosomal storage diseases: the lessons and promise of animal models. *J. Gene Med.* **6**, 481–506 (2004).

201. Vite, C. H., McGowan, J. C., Niogi, S. N., Passini, M. A., Drobatz, K. J., Haskins, M. E. & Wolfe, J. H. Effective gene therapy for an inherited CNS disease in a large animal model. *Ann Neurol.* **57**, 355–364 (2005).

202. Vite, C. H., Passini, M. A., Haskins, M. E. & Wolfe, J. H. Adeno-associated virus vector-mediated transduction in the cat brain. *Gene Ther* **10**, 1874–1881 (2003).

203. Matsuzaki, Y., Konno, A., Mochizuki, R., Shinohara, Y., Nitta, K., Okada, Y. & Hirai, H. Intravenous administration of the adeno-associated virus-PHP.B capsid fails to upregulate transduction efficiency in the marmoset brain. *Neuroscience Letters* **665**, 182–188 (2018).

204. Yuasa, K., Yoshimura, M., Urasawa, N., Ohshima, S., Howell, J. M., Nakamura, A., Hijikata, T., Miyagoe-Suzuki, Y. & Takeda, S. Injection of a recombinant AAV serotype 2 into canine skeletal muscles evokes strong immune responses against transgene products. *Gene Ther* **14**, 1249–1260 (2007).

205. Wang, Z., Allen, J. M., Riddell, S. R., Gregorevic, P., Storb, R., Tapscott, S. J., Chamberlain, J. S. & Kuhr, C. S. Immunity to Adeno-Associated Virus-Mediated Gene Transfer in a Random-Bred Canine Model of Duchenne Muscular Dystrophy. *Human Gene Therapy* **18**, 18–26 (2007).

206. Wang, Z., Kuhr, C. S., Allen, J. M., Blankinship, M., Gregorevic, P., Chamberlain, J. S., Tapscott, S. J. & Storb, R. Sustained AAV-mediated Dystrophin Expression in a Canine Model of Duchenne Muscular Dystrophy with a Brief Course of Immunosuppression. *Molecular Therapy* **15**, 1160–1166 (2007).

207. Nathwani, A. C., Gray, J. T., McIntosh, J., Ng, C. Y. C., Zhou, J., Spence, Y., Cochrane, M., Gray, E., Tuddenham, E. G. D. & Davidoff, A. M. Safe and efficient transduction of the liver after peripheral vein infusion of self-complementary AAV vector results in stable therapeutic expression of human FIX in nonhuman primates. *Blood* **109**, 1414–1421 (2007).

208. Li, H., Lasaro, M. O., Jia, B., Lin, S. W., Haut, L. H., High, K. A. & Ertl, H. C. Capsid-specific T-cell Responses to Natural Infections With Adeno-associated Viruses in Humans Differ From Those of Nonhuman Primates. *Molecular Therapy* **19**, 2021–2030 (2011).

209. Cai, Y., Cheng, T., Yao, Y., Li, X., Ma, Y., Li, L., Zhao, H., Bao, J., Zhang, M., Qiu, Z. & Xue, T. In vivo genome editing rescues photoreceptor degeneration via a Cas9/RecA-mediated homology-directed repair pathway. *Sci. Adv.* **5**, eaav3335 (2019).

210. Latella, M. C., Di Salvo, M. T., Cocchiarella, F., Benati, D., Grisendi, G., Comitato, A., Marigo, V. & Recchia, A. In vivo Editing of the Human Mutant Rhodopsin Gene by Electroporation of Plasmid-based CRISPR/Cas9 in the Mouse Retina. *Molecular Therapy - Nucleic Acids* **5**, e389 (2016).

211. Yu, W., Mookherjee, S., Chaitankar, V., Hiriyanna, S., Kim, J.-W., Brooks, M., Ataeijannati, Y., Sun, X., Dong, L., Li, T., Swaroop, A. & Wu, Z. Nrl knockdown by AAV-delivered CRISPR/Cas9 prevents retinal degeneration in mice. *Nat Commun* **8**, 14716 (2017).

212. Maeder, M. L., Stefanidakis, M., Wilson, C. J., Baral, R., Barrera, L. A., Bounoutas, G. S., Bumcrot, D., Chao, H., Ciulla, D. M., DaSilva, J. A., Dass, A., Dhanapal, V., Fennell, T. J., Friedland, A. E., Giannoukos, G., Gloskowski, S. W., Glucksmann, A., Gotta, G. M., Jayaram, H., Haskett, S. J., Hopkins, B., Horng, J. E., Joshi, S., Marco, E., Mepani, R., Reyon, D., Ta, T., Tabbaa, D. G., Samuelsson, S. J., Shen, S., Skor, M. N., Stetkiewicz, P., Wang, T., Yudkoff, C., Myer, V. E., Albright, C. F. & Jiang, H. Development of a gene-

editing approach to restore vision loss in Leber congenital amaurosis type 10. *Nat Med* **25**, 229–233 (2019).

213. Jo, D. H., Song, D. W., Cho, C. S., Kim, U. G., Lee, K. J., Lee, K., Park, S. W., Kim, D., Kim, J. H., Kim, J.-S., Kim, S., Kim, J. H. & Lee, J. M. CRISPR-Cas9–mediated therapeutic editing of *Rpe65* ameliorates the disease phenotypes in a mouse model of Leber congenital amaurosis. *Sci. Adv.* **5**, eaax1210 (2019).

214. Koo, T., Park, S. W., Jo, D. H., Kim, D., Kim, J. H., Cho, H.-Y., Kim, J., Kim, J. H. & Kim, J.-S. CRISPR-LbCpf1 prevents choroidal neovascularization in a mouse model of age-related macular degeneration. *Nat Commun* **9**, 1855 (2018).

215. Yang, S., Chang, R., Yang, H., Zhao, T., Hong, Y., Kong, H. E., Sun, X., Qin, Z., Jin, P., Li, S. & Li, X.-J. CRISPR/Cas9-mediated gene editing ameliorates neurotoxicity in mouse model of Huntington's disease. *Journal of Clinical Investigation* **127**, 2719–2724 (2017).

216. Colasante, G., Lignani, G., Brusco, S., Di Berardino, C., Carpenter, J., Giannelli, S., Valassina, N., Bido, S., Ricci, R., Castoldi, V., Marenna, S., Church, T., Massimino, L., Morabito, G., Benfenati, F., Schorge, S., Leocani, L., Kullmann, D. M. & Broccoli, V. dCas9-Based Scn1a Gene Activation Restores Inhibitory Interneuron Excitability and Attenuates Seizures in Dravet Syndrome Mice. *Molecular Therapy* **28**, 235–253 (2020).

217. Matharu, N., Rattanasopha, S., Tamura, S., Maliskova, L., Wang, Y., Bernard, A., Hardin, A., Eckalbar, W. L., Vaisse, C. & Ahituv, N. CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency. *Science* **363**, eaau0629 (2019).

218. Stephens, C. J., Lauron, E. J., Kashentseva, E., Lu, Z. H., Yokoyama, W. M. & Curiel, D. T. Long-term correction of hemophilia B using adenoviral delivery of CRISPR/Cas9. *Journal of Controlled Release* **298**, 128–141 (2019).

219. Guan, Y., Ma, Y., Li, Q., Sun, Z., Ma, L., Wu, L., Wang, L., Zeng, L., Shao, Y., Chen, Y., Ma, N., Lu, W., Hu, K., Han, H., Yu, Y., Huang, Y., Liu, M. & Li, D. CRISPR /Cas9-mediated somatic correction of a novel coagulator factor IX gene mutation ameliorates hemophilia in mouse. *EMBO Mol Med* **8**, 477–488 (2016).

220. Chadwick, A. C., Evitt, N. H., Lv, W. & Musunuru, K. Reduced Blood Lipid Levels With In Vivo CRISPR-Cas9 Base Editing of ANGPTL3. *Circulation* **137**, 975–977 (2018).

221. Carreras, A., Pane, L. S., Nitsch, R., Madeyski-Bengtson, K., Porritt, M., Akcakaya, P., Taheri-Ghahfarokhi, A., Ericson, E., Bjursell, M., Perez-Alcazar, M., Seeliger, F., Althage, M., Knöll, R., Hicks, R., Mayr, L. M., Perkins, R., Lindén, D., Borén, J., Bohlooly-Y, M. & Maresca, M. In vivo genome and base editing of a human PCSK9 knock-in hypercholesterolemic mouse model. *BMC Biol* **17**, 4 (2019).

222. Villiger, L., Grisch-Chan, H. M., Lindsay, H., Ringnalda, F., Pogliano, C. B., Allegri, G., Fingerhut, R., Häberle, J., Matos, J., Robinson, M. D., Thöny, B. & Schwank, G. Treatment of a metabolic liver disease by in vivo genome base editing in adult mice. *Nat Med* **24**, 1519–1525 (2018).

223. Shao, Y., Wang, L., Guo, N., Wang, S., Yang, L., Li, Y., Wang, M., Yin, S., Han, H., Zeng, L., Zhang, L., Hui, L., Ding, Q., Zhang, J., Geng, H., Liu, M. & Li, D. Cas9-nickase–mediated genome editing corrects hereditary tyrosinemia in rats. *Journal of Biological Chemistry* **293**, 6883–6892 (2018).

224. Song, C.-Q., Jiang, T., Richter, M., Rhym, L. H., Koblan, L. W., Zafra, M. P., Schatoff, E. M., Doman, J. L., Cao, Y., Dow, L. E., Zhu, L. J., Anderson, D. G., Liu, D. R., Yin, H. & Xue, W. Adenine base editing in an adult mouse model of tyrosinaemia. *Nat Biomed Eng* **4**, 125–130 (2020).

225. Wang, D., Li, J., Song, C.-Q., Tran, K., Mou, H., Wu, P.-H., Tai, P. W. L., Mendonca, C. A., Ren, L., Wang, B. Y., Su, Q., Gessler, D. J., Zamore, P. D., Xue, W. & Gao, G. Cas9-mediated allelic exchange repairs compound heterozygous recessive mutations in mice. *Nat Biotechnol* **36**, 839–842 (2018).

226. Amoasii, L., Long, C., Li, H., Mireault, A. A., Shelton, J. M., Sanchez-Ortiz, E., McAnally, J. R., Bhattacharyya, S., Schmidt, F., Grimm, D., Hauschka, S. D., Bassel-Duby, R. & Olson, E. N. Single-cut genome editing restores dystrophin expression in a new mouse model of muscular dystrophy. *Sci. Transl. Med.* **9**, eaan8081 (2017).

227. Min, Y.-L., Li, H., Rodriguez-Caycedo, C., Mireault, A. A., Huang, J., Shelton, J. M., McAnally, J. R., Amoasii, L., Mammen, P. P. A., Bassel-Duby, R. & Olson, E. N. CRISPR-Cas9 corrects Duchenne muscular dystrophy exon 44 deletion mutations in mice and human cells. *Sci. Adv.* **5**, eaav4324 (2019).

228. Nelson, C. E., Hakim, C. H., Ousterout, D. G., Thakore, P. I., Moreb, E. A., Rivera, R. M. C., Madhavan, S., Pan, X., Ran, F. A., Yan, W. X., Asokan, A., Zhang, F., Duan, D. & Gersbach, C. A. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* **351**, 403–407 (2016).

229. Hakim, C. H., Wasala, N. B., Nelson, C. E., Wasala, L. P., Yue, Y., Louderman, J. A., Lessa, T. B., Dai, A., Zhang, K., Jenkins, G. J., Nance, M. E., Pan, X., Kodippili, K., Yang, N. N., Chen, S., Gersbach, C. A. & Duan, D. AAV CRISPR editing rescues cardiac and muscle function for 18 months in dystrophic mice. *JCI Insight* **3**, e124297 (2018).

230. Long, C., Amoasii, L., Mireault, A. A., McAnally, J. R., Li, H., Sanchez-Ortiz, E., Bhattacharyya, S., Shelton, J. M., Bassel-Duby, R. & Olson, E. N. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* **351**, 400–403 (2016).

231. Xu, L., Park, K. H., Zhao, L., Xu, J., El Refaey, M., Gao, Y., Zhu, H., Ma, J. & Han, R. CRISPR-mediated Genome Editing Restores Dystrophin Expression and Function in mdx Mice. *Molecular Therapy* **24**, 564–569 (2016).

232. Koo, T., Lu-Nguyen, N. B., Malerba, A., Kim, E., Kim, D., Cappellari, O., Cho, H.-Y., Dickson, G., Popplewell, L. & Kim, J.-S. Functional Rescue of Dystrophin Deficiency in Mice Caused by Frameshift Mutations Using Campylobacter jejuni Cas9. *Molecular Therapy* **26**, 1529–1538 (2018).

233. Gao, X., Tao, Y., Lamas, V., Huang, M., Yeh, W.-H., Pan, B., Hu, Y.-J., Hu, J. H., Thompson, D. B., Shu, Y., Li, Y., Wang, H., Yang, S., Xu, Q., Polley, D. B., Liberman, M. C., Kong, W.-J., Holt, J. R., Chen, Z.-Y. & Liu, D. R. Treatment of autosomal dominant hearing loss by in vivo delivery of genome editing agents. *Nature* **553**, 217–221 (2018).

234. György, B., Nist-Lund, C., Pan, B., Asai, Y., Karavitaki, K. D., Kleinstiver, B. P., Garcia, S. P., Zaborowski, M. P., Solanes, P., Spataro, S., Schneider, B. L., Joung, J. K., Géléoc, G. S. G., Holt, J. R. & Corey, D. P. Allele-specific gene editing prevents deafness in a model of dominant progressive hearing loss. *Nat Med* **25**, 1123–1130 (2019).

235. Moreno, A. M., Palmer, N., Alemán, F., Chen, G., Pla, A., Jiang, N., Leong Chew, W., Law, M. & Mali, P. Immune-orthogonal orthologues of AAV capsids and of Cas9 circumvent the immune response to the administration of gene therapy. *Nat Biomed Eng* **3**, 806–816 (2019).

236. Wagner, D. L., Amini, L., Wendering, D. J., Burkhardt, L.-M., Akyüz, L., Reinke, P., Volk, H.-D. & Schmueck-Henneresse, M. High prevalence of Streptococcus pyogenes Cas9-reactive T cells within the adult human population. *Nat Med* **25**, 242–248 (2019).

237. Simhadri, V. L., McGill, J., McMahon, S., Wang, J., Jiang, H. & Sauna, Z. E. Prevalence of Pre-existing Antibodies to CRISPR-Associated Nuclease Cas9 in the USA Population. *Molecular Therapy - Methods & Clinical Development* **10**, 105–112 (2018).

238. Charlesworth, C. T., Deshpande, P. S., Dever, D. P., Camarena, J., Lemgart, V. T., Cromer, M. K., Vakulskas, C. A., Collingwood, M. A., Zhang, L., Bode, N. M., Behlke, M. A., Dejene, B., Cieniewicz, B., Romano, R., Lesch, B. J., Gomez-Ospina, N., Mantri, S., Pavel-Dinu, M., Weinberg, K. I. & Porteus, M. H. Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat Med* **25**, 249–254 (2019).

239. Crudele, J. M. & Chamberlain, J. S. Cas9 immunity creates challenges for CRISPR gene editing therapies. *Nat Commun* **9**, 3497 (2018).

240. Li, A., Tanner, M. R., Lee, C. M., Hurley, A. E., De Giorgi, M., Jarrett, K. E., Davis, T. H., Doerfler, A. M., Bao, G., Beeton, C. & Lagor, W. R. AAV-CRISPR Gene Editing Is Negated by Pre-existing Immunity to Cas9. *Molecular Therapy* **28**, 1432–1441 (2020).

241. Ferdosi, S. R., Ewaisha, R., Moghadam, F., Krishna, S., Park, J. G., Ebrahimkhani, M. R., Kiani, S. & Anderson, K. S. Multifunctional CRISPR-Cas9 with engineered immunosilenced human T cell epitopes. *Nat Commun* **10**, 1842 (2019).

242. McGarry, M. E., Illek, B., Ly, N. P., Zlock, L., Olshansky, S., Moreno, C., Finkbeiner, W. E. & Nielson, D. W. In vivo and in vitro ivacaftor response in cystic fibrosis patients with residual CFTR function: N-of-1 studies. *Pediatr Pulmonol.* **52**, 472–479 (2017).

243. Peabody Lever, J. E., Mutyam, V., Hathorne, H. Y., Peng, N., Sharma, J., Edwards, L. J. & Rowe, S. M. Ataluren/ivacaftor combination therapy: Two N-of-1 trials in cystic fibrosis patients with nonsense mutations. *Pediatr Pulmonol* **55**, 1838–1842 (2020).

244. Kim, J., Hu, C., Moufawad El Achkar, C., Black, L. E., Douville, J., Larson, A., Pendergast, M. K., Goldkind, S. F., Lee, E. A., Kuniholm, A., Soucy, A., Vaze, J., Belur, N. R., Fredriksen, K., Stojkovska, I., Tsytsykova, A., Armant, M., DiDonato, R. L., Choi, J., Cornelissen, L., Pereira, L. M., Augustine, E. F., Genetti, C. A., Dies, K., Barton, B., Williams, L., Goodlett, B. D., Riley, B. L., Pasternak, A., Berry, E. R., Pflock, K. A., Chu, S., Reed, C., Tyndall, K., Agrawal, P. B., Beggs, A. H., Grant, P. E., Urion, D. K., Snyder, R. O., Waisbren, S. E., Poduri, A., Park, P. J., Patterson, A., Biffi, A., Mazzulli, J. R., Bodamer, O., Berde, C. B. & Yu, T. W. Patient-Customized Oligonucleotide Therapy for a Rare Genetic Disease. *N Engl J Med* **381**, 1644–1652 (2019).

245. Marks, P. & Witten, C. Toward a new framework for the development of individualized therapies. *Gene Ther* **28**, 615–617 (2021).

246. Stunnenberg, B. C., Deinum, J., Nijenhuis, T., Huysmans, F., van der Wilt, G. J., van Engelen, B. G. M. & van Agt, F. N-of-1 Trials: Evidence-Based Clinical Care or Medical Research that Requires IRB Approval? A Practical Flowchart Based on an Ethical Framework. *Healthcare* **8**, 49 (2020).

247. Blackston, J., Chapple, A., McGree, J., McDonald, S. & Nikles, J. Comparison of Aggregated N-of-1 Trials with Parallel and Crossover Randomized Controlled Trials Using Simulation Studies. *Healthcare* **7**, 137 (2019).

248. Mingozzi, F. & High, K. A. Immune responses to AAV vectors: overcoming barriers to successful gene therapy. *Blood* **122**, 23–36 (2013).

249. Zaldumbide, A. & Hoeben, R. C. How not to be seen: Immune-evasion strategies in gene therapy. *Gene Therapy* vol. 15 239–246 (2008).

250. Yang, Y., Li, Q., Ertl, H. C. & Wilson, J. M. Cellular and humoral immune responses to viral antigens create barriers to lung-directed gene therapy with recombinant adenoviruses. *Journal of virology* **69**, 2004–15 (1995).

251. Jawa, V., Cousens, L. P., Awwad, M., Wakshull, E., Kropshofer, H. & De Groot, A. S. T-cell dependent immunogenicity of protein therapeutics: Preclinical assessment and mitigation. *Clinical Immunology* vol. 149 534–555 (2013).

252. Mays, L. E. & Wilson, J. M. The Complex and Evolving Story of T cell Activation to AAV Vector-encoded Transgene Products. *Molecular Therapy* **19**, 16–27 (2011).

253. Basner-Tschakarjan, E., Bijjiga, E. & Martino, A. T. Pre-clinical assessment of immune responses to adeno-associated virus (AAV) vectors. *Frontiers in Immunology* vol. 5 (2014).

254. Ertl, H. C. J. & High, K. A. Impact of AAV Capsid-Specific T-Cell Responses on Design and Outcome of Clinical Gene Transfer Trials with Recombinant Adeno-Associated Viral Vectors: An Evolving Controversy. *Human Gene Therapy* **28**, 328–337 (2017).

255. Kotterman, M. A., Chalberg, T. W. & Schaffer, D. V. Viral Vectors for Gene Therapy: Translational and Clinical Outlook. *Annual Review of Biomedical Engineering* **17**, 63–89 (2015).

256. Mingozzi, F. & High, K. A. Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nature Reviews Genetics* **12**, 341–355 (2011).

257. Manno, C. S., Arruda, V. R., Pierce, G. F., Glader, B., Ragni, M., Rasko, J., Ozelo, M. C., Hoots, K., Blatt, P., Konkle, B., Dake, M., Kaye, R., Razavi, M., Zajko, A., Zehnder, J., Nakai, H., Chew, A., Leonard, D., Wright, J. F., Lessard, R. R., Sommer, J. M., Tigges, M., Sabatino, D., Luk, A., Jiang, H., Mingozzi, F., Couto, L., Ertl, H. C., High, K. A. & Kay, M. A. Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nature Medicine* **12**, 342–347 (2006).

258. Chew, W. L. Immunity to CRISPR Cas9 and Cas12a therapeutics. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* **10**, e1408 (2018).

259. Jacobs, F., Gordts, S. C., Muthuramu, I. & De Geest, B. The liver as a target organ for gene therapy: state of the art, challenges, and future perspectives. *Pharmaceuticals (Basel, Switzerland)* **5**, 1372–92 (2012).

260. Kok, C. Y., Cunningham, S. C., Carpenter, K. H., Dane, A. P., Siew, S. M., Logan, G. J., Kuchel, P. W. & Alexander, I. E. Adeno-associated Virus-mediated Rescue of Neonatal Lethality in Argininosuccinate Synthetase-deficient Mice. *Molecular Therapy* **21**, 1823–1831 (2013).

261. Courtenay-Luck, N. S., Epenetos, A. A. & Moore, R. Development of primary and secondary immune responses to mouse monoclonal antibodies used in the diagnosis and therapy of malignant neoplasms. *Cancer Research* **46**, 6489–6493 (1986).

262. Wagner, J. a, Messner, a H., Moran, M. L., Daifuku, R., Kouyama, K., Desch, J. K., Manley, S., Norbash, a M., Conrad, C. K., Friborg, S., Reynolds, T., Guggino, W. B., Moss, R. B., Carter, B. J., Wine, J. J., Flotte, T. R. & Gardner, P. Safety and biological efficacy of an adeno-associated virus vector-cystic fibrosis transmembrane regulator (AAV-CFTR) in the cystic fibrosis maxillary sinus. *The Laryngoscope* **109**, 266–74 (1999).

263. Song, S., Morgan, M., Ellis, T., Poirier, A., Chesnut, K., Wang, J., Brantly, M., Muzyczka, N., Byrne, B. J., Atkinson, M. & Flotte, T. R. Sustained secretion of human alpha-1-antitrypsin from murine muscle transduced with adeno-associated virus vectors. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 14384–8 (1998).

264. Chirmule, N., Xiao, W., Truneh, A., Schnell, M. A., Hughes, J. V, Zoltick, P. & Wilson, J. M. Humoral Immunity to Adeno-Associated Virus Type 2 Vectors following Administration to Murine and Nonhuman Primate Muscle. *Journal of Virology* **74**, 2420–2425 (2000).

265. Fields, P. a, Arruda, V. R., Armstrong, E., Chu, K., Mingozzi, F., Hagstrom, J. N., Herzog, R. W. & High, K. a. Risk and prevention of anti-factor IX formation in AAV-mediated gene transfer in the context of a large deletion of F9. *Molecular therapy : the journal of the American Society of Gene Therapy* **4**, 201–210 (2001).

266. Herzog, R. W., Fields, P. a, Arruda, V. R., Brubaker, J. O., Armstrong, E., McClintock, D., Bellinger, D. a, Couto, L. B., Nichols, T. C. & High, K. a. Influence of vector dose on factor IX-specific T and B cell responses in muscle-directed gene therapy. *Human gene therapy* **13**, 1281–91 (2002).

267. Lozier, J. N., Tayebi, N. & Zhang, P. Mapping of genes that control the antibody response to human factor IX in mice. *Blood* **105**, 1029–1035 (2005).

268. Zhang, H. G., High, K. A., Wu, Q., Yang, P., Schlachterman, A., Yu, S., Yi, N., Hsu, H. C. & Mountz, J. D. Genetic analysis of the antibody response to AAV2 and factor IX. *Mol.Ther.* **11**, 866–874 (2005).

269. Tam, H. H., Melo, M. B., Kang, M., Pelet, J. M., Ruda, V. M., Foley, M. H., Hu, J. K., Kumari, S., Crampton, J., Baldeon, A. D., Sanders, R. W., Moore, J. P., Crotty, S., Langer, R., Anderson, D. G., Chakraborty, A. K. & Irvine, D. J. Sustained antigen availability during germinal center initiation enhances antibody responses to vaccination. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E6639–E6648 (2016).

270. Chew, W. L., Tabebordbar, M., Cheng, J. K. W., Mali, P., Wu, E. Y., Ng, A. H. M., Zhu, K., Wagers, A. J. & Church, G. M. A multifunctional AAV–CRISPR–Cas9 and its host response. *Nature Methods* **13**, 868–874 (2016).

271. Benveniste, O., Boutin, S., Monteilhet, V., Veron, P., Leborgne, C., Montus, M. F. & Masurier, C. Prevalence of Serum IgG and Neutralizing Factors Against Adeno-Associated Virus (AAV) Types 1,2,5,6,8, and 9 in the Healthy Population: Implications for Gene Therapy Using AAV Vectors. *Human Gene Therapy* **21**, 704–712 (2010).

272. Gao, G.-P., Alvira, M. R., Wang, L., Calcedo, R., Johnston, J. & Wilson, J. M. Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proceedings of the National Academy of Sciences* **99**, 11854–11859 (2002).

273. Jooss, K., Yang, Y., Fisher, K. J. & Wilson, J. M. Transduction of Dendritic Cells by DNA Viral Vectors Directs the Immune Response to Transgene Products in Muscle Fibers. *Journal of Virology* **72**, 4212–4223 (1998).

274. Gernoux, G., Guilbaud, M., Dubreil, L., Larcher, T., Babarit, C., Ledevin, M., Jaulin, N., Planel, P., Moullier, P. & Adjali, O. Early Interaction of Adeno-Associated Virus Serotype 8 Vector with the Host Immune System Following Intramuscular Delivery Results in Weak but Detectable Lymphocyte and Dendritic Cell Transduction. *Human Gene Therapy* **26**, 1–13 (2015).

275. Zhu, J., Huang, X. & Yang, Y. The TLR9-MyD88 pathway is critical for adaptive immune responses to adeno-associated virus gene therapy vectors in mice. *The Journal of Clinical Investigation* **119**, 2388–2398 (2009).

276. Gernoux, G., Wilson, J. M. & Mueller, C. Regulatory and Exhausted T Cell Responses to AAV Capsid. *Human Gene Therapy* **28**, 338–349 (2017).

277. Ding, Q., Strong, A., Patel, K. M., Ng, S.-L., Gosis, B. S., Regan, S. N., Rader, D. J. & Musunuru, K. Permanent Alteration of PCSK9 With In Vivo CRISPR-Cas9 Genome Editing. *Circulation research* **115**, 488–492 (2014).

278. Wagner, D. L., Amini, L., Wendering, D. J., Burkhardt, L.-M., Akyüz, L., Reinke, P., Volk, H.-D. & Schmueck-Henneresse, M. High prevalence of Streptococcus pyogenes Cas9-reactive T cells within the adult human population. *Nature Medicine* 1 (2018) doi:10.1038/s41591-018-0204-6.

279. McIntosh, J. H., Cochrane, M., Cobbold, S., Waldmann, H., Nathwani, S. A., Davidoff, A. M. & Nathwani, A. C. Successful attenuation of humoral immunity to viral capsid and transgenic protein following AAV-mediated gene transfer with a non-depleting CD4 antibody and cyclosporine. *Gene Ther* **19**, 78–85 (2012).

280. Mingozzi, F., Chen, Y., Edmonson, S. C., Zhou, S., Thurlings, R. M., Tak, P. P., High, K. A. & Vervoordeldonk, M. J. Prevalence and pharmacological modulation of humoral immunity to AAV vectors in gene transfer to synovial tissue. *Gene Ther* **20**, 417–424 (2013).

281. Mingozzi, F., Chen, Y., Murphy, S. L., Edmonson, S. C., Tai, A., Price, S. D., Metzger, M. E., Zhou, S., Wright, J. F., Donahue, R. E., Dunbar, C. E. & High, K. A. Pharmacological Modulation of Humoral Immunity in a Nonhuman Primate Model of AAV Gene Transfer for Hemophilia B. *Molecular Therapy* **20**, 1410–1416 (2017).

282. Unzu, C., Hervás-Stubbs, S., Sampedro, A., Mauleón, I., Mancheño, U., Alfaro, C., de Salamanca, R. E., Benito, A., Beattie, S. G., Petry, H., Prieto, J., Melero, I. & Fontanellas, A. Transient and intensive pharmacological immunosuppression fails to improve AAV-based liver gene transfer in non-human primates. *Journal of Translational Medicine* **10**, 122 (2012).

283. Riechmann, L., Clark, M., Waldmann, H. & Winter, G. Reshaping human antibodies for therapy. *Nature* vol. 332 323–7 (1988).

284. Amabile, A., Migliara, A., Capasso, P., Biffi, M., Cittaro, D., Naldini, L. & Lombardo, A. Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. *Cell* **167**, 219-232.e14 (2016).

285. Hinderer, C., Katz, N., Buza, E. L., Dyer, C., Goode, T., Bell, P., Richman, L. & Wilson, J. M. Severe toxicity in nonhuman primates and piglets following high-dose intravenous administration of an AAV vector expressing human SMN. *Human Gene Therapy* hum.2018.015 (2018) doi:10.1089/hum.2018.015.

286. Adamopoulou, E., Tenzer, S., Hillen, N., Klug, P., Rota, I. A., Tietz, S., Gebhardt, M., Stevanovic, S., Schild, H., Tolosa, E., Melms, A. & Stoeckle, C. Exploring the MHC-peptide matrix of central tolerance in the human thymus. *Nature Communications* **4**, 2039 (2013).

287. Baker, M. P., Reynolds, H. M., Lumicisi, B. & Bryson, C. J. Immunogenicity of protein therapeutics: The key causes, consequences and challenges. *Self Nonself* **1**, 314–322 (2010).

288. Charlesworth, C. T., Deshpande, P. S., Dever, D. P., Dejene, B., Gomez-Ospina, N., Mantri, S., Pavel-Dinu, M., Camarena, J., Weinberg, K. I. & Porteus, M. H. Identification of Pre-Existing Adaptive Immunity to Cas9 Proteins in Humans. doi:10.1101/243345.

289. EL-Manzalawy, Y., Dobbs, D. & Honavar, V. Predicting linear B-cell epitopes using string kernels. *Journal of molecular recognition : JMR* **21**, 243–255 (2008).

290. Larsen, J. E. P., Lund, O. & Nielsen, M. Improved method for predicting linear B-cell epitopes. *Immunome Research* **2**, 2 (2006).

291. Sollner, J., Grohmann, R., Rapberger, R., Perco, P., Lukas, A. & Mayer, B. Analysis and prediction of protective continuous B-cell epitopes on pathogen proteins. *Immunome Research* **4**, 1 (2008).

292. Dalkas, G. A. & Rooman, M. SEPIa, a knowledge-driven algorithm for predicting conformational B-cell epitopes from the amino acid sequence. *BMC Bioinformatics* **18**, 95 (2017).

293. Sun, P., Ju, H., Liu, Z., Ning, Q., Zhang, J., Zhao, X., Huang, Y., Ma, Z. & Li, Y. Bioinformatics resources and tools for conformational B-cell epitope prediction. *Computational and Mathematical Methods in Medicine* vol. 2013 (2013).

294. Andreatta, M. & Nielsen, M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics (Oxford, England)* **32**, 511–517 (2015).

295. Andreatta, M., Karosiene, E., Rasmussen, M., Stryhn, A., Buus, S. & Nielsen, M. Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification. *Immunogenetics* **67**, 641–650 (2015).

296. Vita, R., Overton, J. A., Greenbaum, J. A., Ponomarenko, J., Clark, J. D., Cantrell, J. R., Wheeler, D. K., Gabbard, J. L., Hix, D., Sette, A. & Peters, B. The immune epitope database (IEDB) 3.0. *Nucleic acids research* **43**, D405-12 (2015).

297. Truong, D.-J. J., Kühner, K., Kühn, R., Werfel, S., Engelhardt, S., Wurst, W. & Ortiz, O. Development of an intein-mediated split–Cas9 system for gene therapy. *Nucleic Acids Research* **43**, 6450–6458 (2015).

298. Moreno, A. M., Fu, X., Zhu, J., Katrekar, D., Shih, Y.-R. V, Marlett, J., Cabotaje, J., Tat, J., Naughton, J., Lisowski, L., Varghese, S., Zhang, K. & Mali, P. In Situ Gene Therapy via

AAV-CRISPR-Cas9-Mediated Targeted Gene Regulation. *Molecular therapy : the journal of the American Society of Gene Therapy* **0**, (2018).

299. Grieger, J. C., Choi, V. W. & Samulski, R. J. Production and characterization of adeno-associated viral vectors. *Nat. Protocols* **1**, 1412–1428 (2006).

300. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Research Notes* **9**, 88 (2016).

301. Pinello, L., Canver, M. C., Hoban, M. D., Orkin, S. H., Kohn, D. B., Bauer, D. E. & Yuan, G.-C. Analyzing CRISPR genome-editing experiments with CRISPResso. *Nature Biotechnology* **34**, 695–697 (2016).

302. Clemente, T., Dominguez, M. R., Vieira, N. J., Rodrigues, M. M. & Amarante-Mendes, G. P. In vivo assessment of specific cytotoxic T lymphocyte killing. *Methods* **61**, 105–109 (2013).

303. Sathish, J. G., Sethu, S., Bielsky, M.-C., de Haan, L., French, N. S., Govindappa, K., Green, J., Griffiths, C. E. M., Holgate, S., Jones, D., Kimber, I., Moggs, J., Naisbitt, D. J., Pirmohamed, M., Reichmann, G., Sims, J., Subramanyam, M., Todd, M. D., Van Der Laan, J. W., Weaver, R. J. & Park, B. K. Challenges and approaches for the development of safer immunomodulatory biologics. *Nat Rev Drug Discov* **12**, 306–324 (2013).

304. Harding, F. A., Stickler, M. M., Razo, J. & DuBridge, R. B. The immunogenicity of humanized and fully human antibodies: Residual immunogenicity resides in the CDR regions. *mAbs* **2**, 256–265 (2010).

305. De Groot, a S., Knopp, P. M. & Martin, W. De-immunization of therapeutic proteins by T-cell epitope modification. *Developments in biologicals* **122**, 171–194 (2005).

306. Tangri, S., Mothé, B. R., Eisenbraun, J., Sidney, J., Southwood, S., Briggs, K., Zinckgraf, J., Bilsel, P., Newman, M., Chesnut, R., LiCalsi, C. & Sette, A. Rationally Engineered Therapeutic Proteins with Reduced Immunogenicity. *The Journal of Immunology* **174**, 3187 LP – 3196 (2005).

307. Ferdosi, S. R., Ewaisha, R., Moghadam, F., Krishna, S., Park, J. G., Ebrahimkhani, M. R., Kiani, S. & Anderson, K. S. Multifunctional CRISPR/Cas9 with engineered immunosilenced human T cell epitopes. *bioRxiv* (2018).

308. Salvat, R. S., Choi, Y., Bishop, A., Bailey-Kellogg, C. & Griswold, K. E. Protein deimmunization via structure-based design enables efficient epitope deletion at high mutational loads. *Biotechnology and Bioengineering* **112**, 1306–1318 (2015).

309. Armstrong, J. K., Hempel, G., Koling, S., Chan, L. S., Fisher, T., Meiselman, H. J. & Garratty, G. Antibody against poly(ethylene glycol) adversely affects PEG-asparaginase therapy in acute lymphoblastic leukemia patients. *Cancer* **110**, 103–111 (2007).

310. Ganson, N. J., Kelly, S. J., Scarlett, E., Sundy, J. S. & Hershfield, M. S. Control of hyperuricemia in subjects with refractory gout, and induction of antibody against poly(ethylene glycol) (PEG), in a phase I trial of subcutaneous PEGylated urate oxidase. *Arthritis Research & Therapy* **8**, R12–R12 (2006).

311. Veronese, F. M. & Mero, A. The impact of PEGylation on biological therapies. *BioDrugs* vol. 22 315–329 (2008).

312. Jevševar, S., Kunstelj, M. & Porekar, V. G. PEGylation of therapeutic proteins. *Biotechnology Journal* **5**, 113–128 (2010).

313. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. A Programmable Dual-RNA – Guided DNA Endonuclease in Adaptice Bacterial Immunity. *Science (New York, N.Y.)* **337**, 816–822 (2012).

314. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences* **109**, E2579–E2586 (2012).

315. Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., Zhang, F., Porteus, M. H., Baltimore, D., Miller, J. C., Sander, J. D., Wood, A. J., Christian, M., Zhang, F., Miller, J. C., Reyon, D., Boch, J., Moscou, M. J., Bogdanove, A. J., Stoddard, B. L., Jinek, M., Gasiunas, G., Barrangou, R., Horvath, P., Siksnys, V., Garneau, J. E., Deveau, H., Garneau, J. E., Moineau, S., Horvath, P., Barrangou, R., Makarova, K. S., Bhaya, D., Davison, M., Barrangou, R., Deltcheva, E., Sapranauskas, R., Magadán, A. H., Dupuis, M. E., Villion, M., Moineau, S., Deveau, H., Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J., Almendros, C., Jinek, M., Doudna, J. A., Malone, C. D., Hannon, G. J., Meister, G., Tuschl, T., Certo, M. T., Mali, P., Carr, P. A., Church, G. M., Barrangou, R., Horvath, P., Brouns, S. J. & Guschin, D. Y. Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, N.Y.)* **339**, 819–23 (2013).

316. Ran, F. A., Cong, L., Yan, W. X., Scott, D. a., Gootenberg, J. S., Kriz, A. J., Zetsche, B., Shalem, O., Wu, X., Makarova, K. S., Koonin, E. V., Sharp, P. a. & Zhang, F. In vivo genome editing using Staphylococcus aureus Cas9. *Nature* **520**, 186–190 (2015).

317. Mali, P., Esvelt, K. M. & Church, G. M. Cas9 as a versatile tool for engineering biology. *Nature Methods* **10**, 957–963 (2013).

318. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).

319. Kelton, W. J., Pesch, T., Matile, S. & Reddy, S. T. Surveying the Delivery Methods of CRISPR/Cas9 for ex vivo Mammalian Cell Engineering. *CHIMIA International Journal for Chemistry* **70**, 439–442 (2016).

320. Cho, S. W., Kim, S., Kim, J.-S. J. M. & Kim, J.-S. J. M. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nature biotechnology* **31**, 230–2 (2013).

321. Moreno, A. M. & Mali, P. Therapeutic genome engineering via CRISPR-Cas systems. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* vol. 9 (2017).

322. Koonin, E. V., Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Current Opinion in Microbiology* vol. 37 67–78 (2017).

323. Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., Backofen, R. & Koonin, E. V. An updated evolutionary classification of CRISPR–Cas systems. *Nature reviews. Microbiology* **13**, 722–736 (2015).

324. Chylinski, K., Makarova, K. S., Charpentier, E. & Koonin, E. V. Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Research* vol. 42 6091–6105 (2014).

325. Shmakov, S., Smargon, A., Scott, D., Cox, D., Pyzocha, N., Yan, W., Abudayyeh, O. O., Gootenberg, J. S., Makarova, K. S., Wolf, Y. I., Severinov, K., Zhang, F. & Koonin, E. V. Diversity and evolution of class 2 CRISPR–Cas systems. *Nature Reviews Microbiology* **15**, 169–182 (2017).

326. Crawley, A. B., Henriksen, J. R. & Barrangou, R. CRISPRdisco: An Automated Pipeline for the Discovery and Analysis of CRISPR-Cas Systems. *The CRISPR Journal* **1**, 171–181 (2018).

327. Simhadri, V. L., Mcgill, J., Mcmahon, S., Wang, J., Jiang, H. & Sauna, Z. E. Prevalence of Pre-existing Antibodies to CRISPR-associated Nuclease Cas9 in the US Population. *Molecular Therapy: Methods & Clinical Development* (2018) doi:10.1016/j.omtm.2018.06.006.

328. Kurosaki, T., Kometani, K. & Ise, W. Memory B cells. *Nature Reviews Immunology* **15**, 149–159 (2015).

329. Zabel, F., Fettelschoss, A., Vogel, M., Johansen, P., Kündig, T. M. & Bachmann, M. F. Distinct T helper cell dependence of memory B-cell proliferation versus plasma cell differentiation. *Immunology* **150**, 329–342 (2017).

330. Zinn, E., Pacouret, S., Khaychuk, V., Turunen, H. T., Carvalho, L. S., Andres-Mateos, E., Shah, S., Shelke, R., Maurer, A. C., Plovie, E., Xiao, R. & Vandenberghe, L. H. In Silico Reconstruction of the Viral Evolutionary Lineage Yields a Potent Gene Therapy Vector. *Cell Reports* **12**, 1056–1068 (2017).

331. Calcedo, R. & Wilson, J. M. AAV Natural Infection Induces Broad Cross-Neutralizing Antibody Responses to Multiple AAV Serotypes in Chimpanzees. *Human Gene Therapy. Clinical Development* **27**, 79–82 (2016).

332. Harbison, C. E., Weichert, W. S., Gurda, B. L., Chiorini, J. A., Agbandje-McKenna, M. & Parrish, C. R. Examining the cross-reactivity and neutralization mechanisms of a panel of mabs against adeno-associated virus serotypes 1 and 5. *Journal of General Virology* **93**, (2012).

333. Majowicz, A., Salas, D., Zabaleta, N., Rodríguez-Garcia, E., González-Aseguinolaza, G., Petry, H. & Ferreira, V. Successful Repeated Hepatic Gene Delivery in Mice and Non-human Primates Achieved by Sequential Administration of AAV5^{ch} and AAV1. *Molecular Therapy* **25**, 1831–1842 (2017).

334. Vollmers, C., Sit, R. V., Weinstein, J. A., Dekker, C. L. & Quake, S. R. Genetic measurement of memory B-cell recall using antibody repertoire sequencing. *Proceedings of the National Academy of Sciences* **110**, 13463–13468 (2013).

335. Ruppert, J., Sidney, J., Celis, E., Kubo, R. T., Grey, H. M. & Sette, A. Prominent role of secondary anchor residues in peptide binding to HLA-A2.1 molecules. *Cell* **74**, 929–937 (2017).

336. Zhang, S.-Q., Parker, P., Ma, K.-Y., He, C., Shi, Q., Cui, Z., Williams, C. M., Wendel, B. S., Meriwether, A. I., Salazar, M. A. & Jiang, N. Direct measurement of T cell receptor affinity and sequence from naïve antiviral T cells. *Science translational medicine* **8**, 341ra77 (2016).

337. Liepe, J., Marino, F., Sidney, J., Jeko, A., Bunting, D. E., Sette, A., Kloetzel, P. M., Stumpf, M. P. H., Heck, A. J. R. & Mishto, M. A large fraction of HLA class I ligands are proteasome-generated spliced peptides. *Science* **354**, (2016).

338. Fonfara, I., Le Rhun, A., Chylinski, K., Makarova, K. S., Lécrivain, A.-L., Bzdrenga, J., Koonin, E. V & Charpentier, E. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Research* **42**, 2577–2590 (2014).

339. Shmakov, S., Smargon, A., Scott, D., Cox, D., Pyzocha, N., Yan, W., Abudayyeh, O. O., Gootenberg, J. S., Makarova, K. S., Wolf, Y. I., Severinov, K., Zhang, F. & Koonin, E. V. Diversity and evolution of class 2 CRISPR–Cas systems. *Nature Reviews Microbiology* **15**, 169–182 (2017).

340. Burstein, D., Harrington, L. B., Strutt, S. C., Probst, A. J., Anantharaman, K., Thomas, B. C., Doudna, J. A. & Banfield, J. F. New CRISPR–Cas systems from uncultivated microbes. *Nature* **542**, 237–241 (2016).

341. Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., Zhang, F. & Koonin,

E. V. Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. *Molecular Cell* **60**, 385–397 (2015).

342. Baker, M. P., Reynolds, H. M., Lumicisi, B. & Bryson, C. J. Immunogenicity of protein therapeutics: The key causes, consequences and challenges. *Self Nonself* **1**, 314–322 (2010).

343. Jawa, V., Terry, F., Gokemeijer, J., Mitra-Kaushik, S., Roberts, B. J., Tourdot, S. & De Groot, A. S. T-Cell Dependent Immunogenicity of Protein Therapeutics Pre-clinical Assessment and Mitigation-Updated Consensus and Review 2020. *Front Immunol* **11**, 1301 (2020).

344. King, C., Garza, E. N., Mazor, R., Linehan, J. L., Pastan, I., Pepper, M. & Baker, D. Removing T-cell epitopes with computational protein design. *Proc Natl Acad Sci U S A* **111**, 8577–8582 (2014).

345. Mazor, R., Crown, D., Addissie, S., Jang, Y., Kaplan, G. & Pastan, I. Elimination of murine and human T-cell epitopes in recombinant immunotoxin eliminates neutralizing and anti-drug antibodies in vivo. *Cell Mol Immunol* **14**, 432–442 (2017).

346. Andreatta, M. & Nielsen, M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* **32**, 511–517 (2016).

347. Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T., Zheng, Z., Gonzales, A. P. W., Li, Z., Peterson, R. T., Yeh, J.-R. J., Aryee, M. J. & Joung, J. K. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485 (2015).

348. Charles, E. J., Kim, S. E., Knott, G. J., Smock, D., Doudna, J. & Savage, D. F. *Engineering improved Cas13 effectors for targeted post-transcriptional regulation of gene expression*. http://biorxiv.org/lookup/doi/10.1101/2021.05.26.445687 (2021) doi:10.1101/2021.05.26.445687.

349. Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., Zeina, C. M., Gao, X., Rees, H. A., Lin, Z. & Liu, D. R. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63 (2018).

350. Allen, B. D., Nisthal, A. & Mayo, S. L. Experimental library screening demonstrates the successful application of computational protein design to large structural ensembles. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19838–19843 (2010).

351. Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z. & Joung, J. K. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).

352. Cao, J., Novoa, E. M., Zhang, Z., Chen, W. C. W., Liu, D., Choi, G. C. G., Wong, A. S. L., Wehrspaun, C., Kellis, M. & Lu, T. K. High-throughput 5′ UTR engineering for enhanced protein production in non-viral gene therapies. *Nat Commun* **12**, 4138 (2021).

353. Sun, M. G. F., Seo, M.-H., Nim, S., Corbi-Verge, C. & Kim, P. M. Protein engineering by highly parallel screening of computationally designed variants. *Sci. Adv.* **2**, e1600692 (2016).

354. Walton, R. T., Christie, K. A., Whittaker, M. N. & Kleinstiver, B. P. Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* **368**, 290–296 (2020).

355. Zinsli, L. V., Stierlin, N., Loessner, M. J. & Schmelcher, M. Deimmunization of protein therapeutics - Recent advances in experimental and computational epitope prediction and deletion. *Comput Struct Biotechnol J* **19**, 315–329 (2021).

356. Griswold, K. E. & Bailey-Kellogg, C. Design and engineering of deimmunized biotherapeutics. *Curr Opin Struct Biol* **39**, 79–88 (2016).

357. Doud, M. B., Lee, J. M. & Bloom, J. D. How single mutations affect viral escape from broad and narrow antibodies to H1 influenza hemagglutinin. *Nat Commun* **9**, 1386 (2018).

358. Gasiunas, G., Young, J. K., Karvelis, T., Kazlauskas, D., Urbaitis, T., Jasnauskaite, M., Grusyte, M. M., Paulraj, S., Wang, P.-H., Hou, Z., Dooley, S. K., Cigan, M., Alarcon, C., Chilcoat, N. D., Bigelyte, G., Curcuru, J. L., Mabuchi, M., Sun, Z., Fuchs, R. T., Schildkraut, E., Weigele, P. R., Jack, W. E., Robb, G. B., Venclovas, Č. & Siksnys, V. A catalogue of biochemically diverse CRISPR-Cas9 orthologs. *Nat Commun* **11**, 5512 (2020).

359. Takeuchi, N., Wolf, Y. I., Makarova, K. S. & Koonin, E. V. Nature and intensity of selection pressure on CRISPR-associated genes. *J Bacteriol* **194**, 1216–1225 (2012).

360. Osipovitch, D. C., Parker, A. S., Makokha, C. D., Desrosiers, J., Kett, W. C., Moise, L., Bailey-Kellogg, C. & Griswold, K. E. Design and analysis of immune-evading enzymes for ADEPT therapy. *Protein Eng Des Sel* **25**, 613–623 (2012).

361. Wang, Y., Zhao, Y., Bollas, A., Wang, Y. & Au, K. F. Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol* **39**, 1348–1365 (2021).

362. Rang, F. J., Kloosterman, W. P. & de Ridder, J. From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. *Genome Biol* **19**, 90 (2018).

363. Schubert, B., Schärfe, C., Dönnes, P., Hopf, T., Marks, D. & Kohlbacher, O. Population-specific design of de-immunized protein biotherapeutics. *PLoS Comput Biol* **14**, e1005983 (2018).

364. Liao, S., Tammaro, M. & Yan, H. Enriching CRISPR-Cas9 targeted cells by co-targeting the HPRT gene. *Nucleic Acids Res* **43**, e134 (2015).

365. Yang, F., Fang, H., Wang, D., Chen, Y., Zhai, Y., Zhou, B.-B. S. & Li, H. HPRT1 activity loss is associated with resistance to thiopurine in ALL. *Oncotarget* **9**, 2268–2278 (2018).

366. Meini, M.-R., Tomatis, P. E., Weinreich, D. M. & Vila, A. J. Quantitative Description of a Protein Fitness Landscape Based on Molecular Features. *Mol Biol Evol* **32**, 1774–1787 (2015).

367. Clement, K., Rees, H., Canver, M. C., Gehrke, J. M., Farouni, R., Hsu, J. Y., Cole, M. A., Liu, D. R., Joung, J. K., Bauer, D. E. & Pinello, L. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat Biotechnol* **37**, 224–226 (2019).

368. Ninkovic, T., Kinarsky, L., Engelmann, K., Pisarev, V., Sherman, S., Finn, O. J. & Hanisch, F.-G. Identification of O-glycosylated decapeptides within the MUC1 repeat domain as potential MHC class I (A2) binding epitopes. *Mol Immunol* **47**, 131–140 (2009).

369. Nuñez, J. K., Chen, J., Pommier, G. C., Cogan, J. Z., Replogle, J. M., Adriaens, C., Ramadoss, G. N., Shi, Q., Hung, K. L., Samelson, A. J., Pogson, A. N., Kim, J. Y. S., Chung, A., Leonetti, M. D., Chang, H. Y., Kampmann, M., Bernstein, B. E., Hovestadt, V., Gilbert, L. A. & Weissman, J. S. Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. *Cell* **184**, 2503-2519.e17 (2021).

370. Amabile, A., Migliara, A., Capasso, P., Biffi, M., Cittaro, D., Naldini, L. & Lombardo, A. Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. *Cell* **167**, 219-232.e14 (2016).

371. Stiffler, M. A., Poelwijk, F. J., Brock, K. P., Stein, R. R., Riesselman, A., Teyra, J., Sidhu, S. S., Marks, D. S., Gauthier, N. P. & Sander, C. Protein Structure from Experimental Evolution. *Cell Syst* **10**, 15-24.e5 (2020).

372. Green, A. G., Elhabashy, H., Brock, K. P., Maddamsetti, R., Kohlbacher, O. & Marks, D. S. Large-scale discovery of protein interactions at residue resolution using co-evolution calculated from genomic sequences. *Nat Commun* **12**, 1396 (2021).

373. Saylor, K., Gillam, F., Lohneis, T. & Zhang, C. Designs of Antigen Structure and Composition for Improved Protein-Based Vaccine Efficacy. *Front Immunol* **11**, 283 (2020).

374. Joglekar, A. V., Leonard, M. T., Jeppson, J. D., Swift, M., Li, G., Wong, S., Peng, S., Zaretsky, J. M., Heath, J. R., Ribas, A., Bethune, M. T. & Baltimore, D. T cell antigen discovery via signaling and antigen-presenting bifunctional receptors. *Nat Methods* **16**, 191–198 (2019).

375. Chen, D., Love, K. T., Chen, Y., Eltoukhy, A. A., Kastrup, C., Sahay, G., Jeon, A., Dong, Y., Whitehead, K. A. & Anderson, D. G. Rapid discovery of potent siRNA-containing lipid nanoparticles enabled by controlled microfluidic formulation. *J Am Chem Soc* **134**, 6948–6951 (2012).

376. Jansen, Ruud., Embden, Jan. D. A. van, Gaastra, Wim. & Schouls, Leo. M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* **43**, 1565–1575 (2002).

377. Moreno, A. M., Catroli, G. F., Alemán, F., Pla, A., Woller, S. A., Hu, M., Yaksh, T. & Mali, P. *Long-lasting Analgesia via Targeted* in vivo *Epigenetic Repression of Nav1.7*. http://biorxiv.org/lookup/doi/10.1101/711812 (2019) doi:10.1101/711812.

378. Chadwick, A. C., Wang, X. & Musunuru, K. In Vivo Base Editing of PCSK9 (Proprotein Convertase Subtilisin/Kexin Type 9) as a Therapeutic Alternative to Genome Editing. *ATVB* **37**, 1741–1747 (2017).