Title

Developing Epigenetic Therapies for Haploinsufficiency

Author

Punjya, Niraj Rajiv

Publication Date

2023

Peer reviewed|Thesis/dissertation

Developing Epigenetic Therapies for Haploinsufficiency

By

NIRAJ RAJIV PUNJYA
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry, Molecular, Cellular, and Developmental Biology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

David Segal

Kyle Fink

Paul Hagerman

Committee in Charge

2023

ABSTRACT

The rapid development of genome editing technologies in the last couple of decades has opened the door for the creation of novel therapeutics and cures for genetic disorders that were once thought to be untreatable. In this dissertation, we explore the potential to develop these tools for the treatment of three genetic disorders, Neurofibromatosis type 1 (NF1), SYNGAP1-related Intellectual Disability, and 22q11.2 deletion syndrome (22q11.2 DS). Although these disorders are very different from one another in terms of gene targets and patient presentation, they share a common theme: they are all a type of haploinsufficiency where an epigenetic gene therapy may provide immense benefit to patients.

Epigenetic editing using CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) involves modifying the chemical modifications that control gene expression without altering the underlying DNA sequence. CRISPR is a powerful tool for editing the genome, which involves using Cas proteins and their ability to complex with RNA guides to localize to specific regions of the genome. The CRISPR system has been adapted to target epigenetic marks using modified versions of the Cas proteins and fusing it to an epigenetic enzyme that adds or removes chemical marks on DNA or histones, such as DNA methyltransferase or histone deacetylase, respectively. This modified Cas9 protein can be directed to specific regions of the genome using guide RNAs that are designed to target specific sequences.

Alternatively, CRISPR can also be used to recruit proteins that activate or repress gene expression to specific regions of the genome. This involves fusing a modified

Cas9 protein to a transcriptional activator or repressor protein respectively, which can then be targeted to specific regions of the genome using guide RNAs.

Overall, epigenetic editing using CRISPR provides a powerful tool for understanding the role of epigenetic modifications in gene regulation and has potential therapeutic applications in treating diseases that involve aberrant epigenetic marks. In summary, for NF1, we utilized a duplex CRISPR repression system to downregulate disease specific genes, and for SYNGAP1-related intellectual disability and 22q11.2 deletion syndrome, we utilized a CRISPR activation system to upregulate disease specific genes.

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my deepest gratitude and appreciation to the many people who have supported and inspired me throughout my journey in pursuing my PhD.

First and foremost, I am grateful to my lab mates, whose guidance, encouragement, and expertise have been instrumental in shaping my research and helping me to grow as a scientist. Their unwavering support and friendship have made my time in the lab both productive and enjoyable, and I am fortunate to have had the opportunity to work alongside such talented and dedicated individuals. Henriette O'Geen, Ulrika Beitnere, and Antonio Gomez, as senior lab members, were a great source of information helping and guiding me design and plan experiments. I greatly enjoyed working alongside fellow graduate students Nicholas Heath, Oran Gutierrez, Chelsey Lee, and Tomas Francisco Salter-Cid.

I would also like to express my gratitude to my family and friends, who have been a constant source of love, encouragement, and support throughout my academic journey. Their unwavering belief in me has been a source of strength and motivation during the challenging times, and their encouragement and kindness have lifted me up during the moments of doubt and uncertainty. My sister Nirali has been a constant source of inspiration in pursuing research to address genetic disorders.

Furthermore, I am deeply indebted to my advisors, who have provided me with invaluable guidance, feedback, and mentorship throughout my PhD journey. Their expertise, insights, and unwavering commitment to my success have been an inspiration and have helped me to achieve the goals that I set for myself. Thank you Dr. Dave Segal for giving me the opportunity to expand my genome editing toolbelt and learn the nuances of epigenetic editing.

Finally, I would like to acknowledge the support of the broader UC Davis academic community, including colleagues, mentors, and collaborators, who have contributed to my development as a researcher and scholar. I am grateful for their feedback, insights, and support, and I look forward to continuing to collaborate with them in the future.

I am humbled and grateful to have had the support and encouragement of so many remarkable individuals during my PhD journey. I would not have been able to achieve this milestone without their help, and I am truly fortunate to have had such an incredible support system.

TABLE OF CONTENTS

Chapter 1- Introduction

The rapid development of genome editing technologies in the last couple of decades has opened the door for the creation of novel therapeutics and cures for genetic disorders that were once thought to be untreatable. In this dissertation, we explore the potential to develop these tools for the treatment of three genetic disorders, Neurofibromatosis type 1 (NF1), SYNGAP1-related intellectual disability, and 22q11.2 deletion syndrome (22q11.2 DS). Although these disorders are very different from one another in terms of gene targets and patient presentation, they share a common theme: they are all a type of haploinsufficiency where an epigenetic gene therapy may provide immense benefit to patients. In this introduction, first we'll document the history of genome editing, and how this work fits into this narrative. We'll briefly discuss each of these disorders, highlighting what is known about the genetics of each, as well a clinical discussion in terms of what is currently lacking in terms of therapy, and how our approach may help fill this gap. Later, we'll discuss specific details of the CRISPR/Cas technologies we have harnessed to combat these three disorders. In summary, for NF1, we utilized a duplex CRISPR repression system to downregulate disease specific genes, and for Syngap1 and 22q11.2 deletion syndrome, we utilized a CRISPR activation system to upregulate disease specific genes.

History of Genome Editing and Epigenome Editing

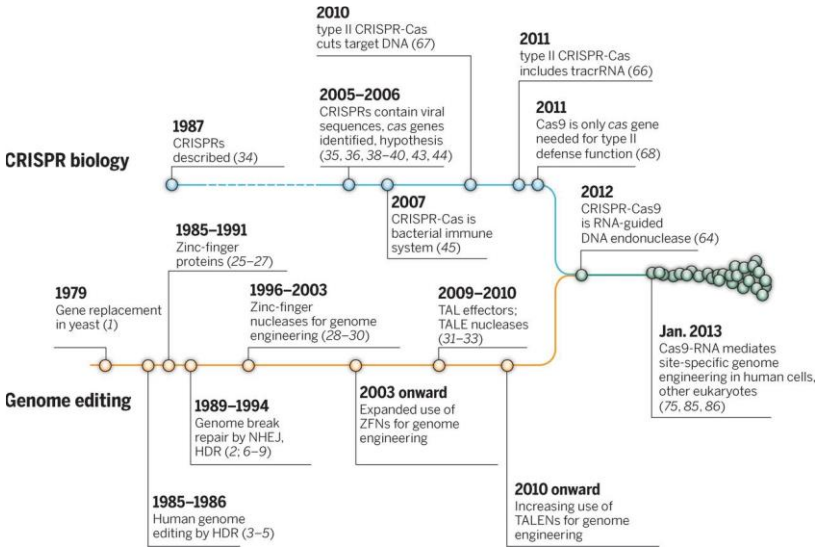
Although any conversation of modern-day genome editing/epigenome editing is dominated by CRISPR-Cas, it's important to reflect on the initial roots that made the ability of manipulate the genome possible.

Our current ability to make and manipulate DNA is a story of more than 60 years of innovation, beginning with the discovery of the DNA double helix. In 1953, James

Watson and Francis Crick pioneered the structure of DNA, postulating the “double helix”. This discovery was only possible with the help of Rosalind Franklin, whose X-ray diffraction images of DNA proteins incepted the idea to Watson and Crick. This landmark discovery by these three scientists paved the road for the modern study of biology and genetics. It is widely considered one of the most significant early events in the field of genetics.[1]

Soon after, enzymes (including polymerases, ligases, and restriction endonucleases) and the polymerase chain reaction (PCR) was discovered, allowing scientists a way to isolate genes and gene fragments, and introduce targeted mutations into genes *in vitro*, in cells, and even in model organisms. Genomic sequencing advances gave us whole-genome sequencing data for many organisms, including humans, and has helped us better understand the genetic landscape of our DNA. Now, the RNA-guided enzyme Cas9, which originates from the CRISPR-Cas adaptive bacterial immune system, is transforming biology by providing a genome engineering tool based on the principles of Waston-Crick base pairing. The simplicity and ease of use of the CRISPR-Cas system has replaced its predecessors, such as Zinc-Finger nucleases (ZFNs) and Transcription Activator Like Effector Nucleases (TALENs) as the premier tool for targeted, genome editing.[5] Now, the CRISPR-Cas system has been adapted to targeted epigenetic editing, by utilizing a DNAase-dead Cas protein, and coupling transcriptional regulators to the machinery.[6]

Figure 1.1- Timeline depicting major milestones in CRISPR Biology for genome editing.



Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014 Nov 28;346(6213):1258096. doi: 10.1126/science.1258096. PMID: 25430774.

Genome Engineering pre-CRISPR

After the discovery of the DNA double helix, researchers and clinicians envisioned the possibility of making specific changes to the genomes of cells and organisms. Not only would this ability be useful for studying the function of unknown genes, it could be used to treat genetic disease.

Early approaches to genome editing utilized the principle of site-specific recognition of DNA sequences. By studying natural DNA repair pathways in simple organisms such as bacteria and yeast, and DNA recombination, researchers discovered the endogenous machinery used to repair double-stranded DNA breaks (DSBs). Thus, a valuable strategy for targeted genomic engineering emerged. If we can introduce a precise break in the DNA at sites where changes need to be introduced, we can effectively manipulate the genome to our liking.[7–11]

The initial approaches to targeted DNA cleavage took advantage of DNA base pair recognition by oligonucleotides or small molecules. It was thought that oligonucleotides coupled to chemical cleavage or cross-linking reagents such as bleomycin and psoralen could be useful for site-specific chromosome modification in yeast and mammalian cells.[12,13] The first instance of being able to introduce targeted double stranded breaks occurred with the development of peptide nucleic acids and polyamides.[12] Here, specific chromosomal loci could be targeted and cleaved if the chemical recognition agent was coupled with a cleavage agent, such as bleomycin. Furthermore, self-splicing introns utilized nucleic acid base pairing to change sequences at the DNA or RNA level.[14–16] Although these approaches were fairly primitive in their ability to introduce targeted DSBs, they demonstrated the promise for base pairing as a means of site-specific genome modification. Such principles were expanded on with homing endonucleases, zinc-finger proteins, and transcription activator–like (TAL) effectors nucleases (TALENs).

Rare-cutting nucleases, such as homing endonucleases, were the earliest used for site-specific DNA cleavage and integration of a DNA sequence flanked by arms of homology to the target site. Using this approach, researchers could incorporate specific genetic information into the genome at sites recognized by the homing endonuclease.[17,18]

Concurrently, another form of targeted DNA binding was being discovered: Zinc-finger mediated DNA binding. Here, the Zinc-finger DNA recognition domain could be coupled to a sequenced-independent nuclease domain of the restriction enzyme FokI, creating a site-specific nuclease.[19] When designed to recognize a chromosomal

sequence, such zinc finger nucleases (ZFNs) were found to be effective at inducing genomic sequence changes in *Drosophila* and mammalian cells.[20,21] ZFNs turned out to be effective genome editing tools for some experiments, but were not widely adopted primarily due to the inherent difficulty in designing and validating such proteins for a specific region of the genome. This limitation opened the door for TALEs, which occur naturally in bacteria that infect plants.[22–24]

Similar to ZFNs, TAL effectors could be coupled to FokI, creating TAL effector nucleases (TALENs). TALENs were much easier to produce and validate than ZFNs, creating the possibility for genome editing at many more regions of the genome. Nevertheless, TALENs were still difficult to design, synthesize, and validate, especially when compared to the next generation of genome editors, the CRISPR/Cas system.

Genome Engineering with CRISPR-Cas

CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated proteins) is a revolutionary tool for genome editing. It is a precise and efficient way to modify DNA sequences and has been widely used in research and has potential applications in medicine, agriculture, and many other fields.

The CRISPR-Cas system is based on a natural process that bacteria use to defend themselves against viruses. The bacteria capture pieces of the virus's DNA and integrate them into their own genome. The integrated DNA then serves as a memory of the virus, allowing the bacteria to recognize and destroy it if it infects again. Scientists have harnessed this natural process to create a tool for precise genome editing.[25,26]

The basic process of CRISPR-Cas genome editing involves the use of two components: a guide RNA (gRNA) and an endonuclease enzyme, usually Cas9. The gRNA contains a sequence that is complementary to the target DNA sequence, and the endonuclease cleaves the DNA at the targeted location. By designing the gRNA to target specific sequences, scientists can make precise cuts in the DNA and introduce desired changes, such as the addition, deletion, or substitution of specific base pairs.[27,28]

One of the key advantages of CRISPR-Cas genome editing is its efficiency. It can be used to target multiple sites in the genome simultaneously.[29] Additionally, the system can be adapted to various organisms, including plants, animals, and humans, making it a versatile tool for research and therapeutic applications.

In the field of medicine, CRISPR-Cas has shown promising results in treating a wide range of diseases, including genetic disorders, cancers, and infectious diseases. For example, scientists have used the system to correct mutations in human cells that cause genetic diseases such as sickle cell anemia, X-Linked severe combined immunodeficiency, and muscular dystrophy.[30–32] In the case of cancer, CRISPR-Cas can be used to precisely target and destroy cancer cells while leaving healthy cells untouched.[33]

However, the use of CRISPR-Cas in human therapies also raises ethical concerns. The precise but occasional off-target nature of the system means that it has the potential to introduce unintended mutations into the genome, and the long-term consequences of these mutations are not yet fully understood. There are also concerns

about the potential for the technology to be used for germline editing, which would result in changes that are passed down to future generations.

Overall, CRISPR-Cas is a powerful tool for genome editing, with a wide range of potential applications in fields such as medicine, agriculture, and biotechnology. However, it is important to carefully consider the ethical implications of this technology and ensure that it is used responsibly and for the benefit of humanity. Further research is needed to fully understand the potential of CRISPR-Cas and to ensure that it is used in a safe and effective manner. In this dissertation, we explore the potential of modifying the canonical genome editing capabilities of the CRISPR-Cas system to perform epigenetic editing.

Epigenetic Editing with the CRISPR-Cas System

Epigenetic editing refers to the manipulation of specific epigenetic marks (i.e., chemical modifications of DNA and/or its associated proteins) with the goal of altering gene expression. This is achieved through the use of molecular tools that can target and modify specific epigenetic marks, leading to changes in the regulation of gene expression. Unlike traditional genetic editing, which involves altering the DNA sequence itself, epigenetic editing does not involve a permanent change to the underlying genetic code.

Epigenetic marks play a crucial role in the regulation of gene expression and cellular behavior, and their misregulation has been linked to a wide range of diseases, including cancer, neurological disorders, and cardiovascular diseases. By modifying

specific epigenetic marks, scientists aim to restore normal cellular behavior and reverse the effects of disease-causing epigenetic changes.[6]

One of the main molecular tools used for epigenetic editing is the CRISPR-Cas system, which has been adapted for use as an epigenetic editor. This system allows for precise, targeted modification of specific epigenetic marks. Other epigenetic editing tools include the use of small molecules, such as small interfering RNA (siRNA) and chemicals that can modify specific epigenetic marks.[34]

Epigenetic editing has the potential to treat a wide range of diseases that have previously been difficult to treat with traditional genetic approaches. However, the field of epigenetic editing is still in its early stages, and there are still many technical and scientific challenges that need to be addressed. For example, the specificity and precision of epigenetic editing tools are still a concern, as well as their potential off-target effects. Additionally, the long-term consequences of epigenetic editing are still not well understood, and further research is needed to fully understand the safety and efficacy of this approach.

Despite these challenges, the potential of epigenetic editing is substantial, and its continued development holds great promise for the treatment of a wide range of diseases. Further research into the biology of epigenetic marks and the development of more advanced molecular tools will likely lead to a deeper understanding of the underlying mechanisms of disease and the development of more effective therapies. In conclusion, epigenetic editing is a promising new approach in the field of molecular biology that offers the potential to treat a wide range of diseases without causing

permanent changes to the DNA sequence, offering a safer alternative to traditional genetic editing.

With the advent CRISPR-Cas, virtually any DNA sequence can be targeted with new customizable epigenome-engineering tools that regulate gene expression by modifying transcription and/or epigenetic state. Making site-specific alterations to the epigenomic landscape in eukaryotic cells is a powerful strategy for interrogating the mechanistic relationships among chromatin state, gene regulation, and cell phenotype. Furthermore, control over gene regulation is a valuable tool for applications such as gene therapy and reprogramming cell fate.[35]

The Two Flavors of Epigenetic Editing: CRISPR Activation and CRISPR Repression

Epigenetic editing differs from canonical gene editing in that permanent changes to the genome are not made. Instead, locus-specific targeting of epigenetic enzymes is used to rewrite the local epigenetic landscape of a genomic site. Such editing can lead to transcriptional reprogramming, altering how genes are expressed.

Another key difference between epigenetic editing and gene editing is the use of a DNAse-dead Cas protein (dCas9) instead of the fully functional DNA-cleaving protein. With dCas9, the DNAse region of the protein is mutated. Instead of introducing a double stranded break at a specific region of interest, transcriptional reprogrammers are fused to the dCas9 protein. The dCas9 still has the ability to complex with the single-guide RNA, allowing the dCas9-sgRNA complex to act as a shuttle for recruiting the transcriptional reprogrammers to a specific region of the genome, altering the epigenetic

landscape of that region.[35] dCas9 is localized to a target sequence by the single-guide RNA, which is an engineered nucleic acid consisting of 18-25 nucleotides followed by a constant region that complexes with dCas9. dCas9 binds target genomic sequences that are complementary to this 18-25 nucleotide single-guide. Target-site flexibility for CRISPR-Cas9 is determined by the protospacer-adjacent motif (PAM), a Cas9 recognition site that immediately follows the protospacer target site in the genome.[27,36,37]

Implementing CRISPR–dCas9 greatly accelerated the advancement of epigenetic editing, yielding preclinical therapeutic successes using a variety of epigenetic enzymes. There are two main flavors of epigenetic editing with the CRISPR-Cas0 system, defined either by the need to upregulate genes (CRISPR activation) or downregulate genes (CRISPR repression).

CRISPR Activation

With CRISPR activation, the direct fusion of potent transcriptional effector domains to dCas proteins can induce transcriptional activation when targeted to endogenous genes. Shortly after the modular nature of ZFP recognition of DNA was identified, ZFPs were linked to VP16 to create a programmable transcriptional activator.[38–40] VP16 is a viral activation domain that recruits Pol II transcriptional machinery.[41] VP64, a tetramer of VP16 domains, has been linked to DBDs to activate coding and noncoding genes via targeting of promoters and regulatory elements.[42–47] Although VP64 does not directly modify chromatin, it recruits remodeling factors and has been linked to increased chromatin accessibility and to the deposition of activating

histone marks, including acetylation of the lysine 27 residue of histone subunit 3 (H3K27ac) and methylation of the lysine 4 residue of histone subunit 3 (H3K4me).[48–50] Another commonly used activator is the p65 subunit of the human NF-κB complex, which has been tethered to ZFPs, TALEs and dCas9 .[51] Gene induction by VP64 and p65 is generally strongest when effector domains are targeted upstream of transcription start sites (TSSs) and in promoter regions, although targeting downstream of TSSs and at distal enhancers can also be effective.[51]

Recruitment of multiple dCas9-VP64 fusions to a single target locus is often required to elicit a robust transcriptional response.[52] Recently, next-generation activators have been developed that outperform the original dCas9-VP64 fusion by recruiting multiple effector domains to a single dCas9-gRNA complex.[53] For example, the SunTag system recruits multiple VP64 activators to dCas9, resulting in stronger activation with a single gRNA than achieved with dCas9-VP64 fusions.[54] Repurposing the gRNA as a scaffold for the recruitment of activation domains p65 and HSF1 via MS2-targeting aptamers also improves transcriptional activation.[51] Fusion of VPR (the product of the tandem fusion of VP64, p65 and RTA, a transactivation domain from **γ**-herpes viruses to the C terminus of dCas9 also increases transcriptional activation in human cells compared to dCas9-VP64.[55] These improved methods enable the use of single gRNAs to achieve robust activation, potentiating genome-wide gene activation screens. For the purposes of CRISPR activation in this thesis (Syngap1 and 22q11.2 DS), we utilized the dCas9-VPR system.

Aberrant gene regulation is often associated with pathological states, either as a symptom or as a cause of underlying disease.[56] Site-specific epigenome editing

provides the opportunity to study the contributions of gene regulation to disease and is an exciting potential avenue for treatment.

Targeted activation of compensatory genes can mitigate the symptoms of diseases that have no cure. For instance, engineered activators can induce the expression of developmentally silenced fetal γ-globin and counteract the loss of functional β-globin in sickle cell anemia and β-thalassemia.[57] Transcriptional modulation strategies can also generate protective effects for regenerative medicine applications. The activation of endogenous vascular endothelial growth factor (VEGF) with engineered ZFPs has been proposed to generate neovasculature in subjects with diabetic neuropathy and peripheral arterial disease, as well as to enhance wound healing.[58] Targeting the VEGF promoter with engineered activators activates all VEGF isoforms, which can result in more mature vasculature formation compared to that achieved via exogenous delivery of a single VEGF isoform. Similarly, engineered ZFP-p65 fusions upregulate pigment epithelium–derived factor and prevent neovascularization in mouse models, suggesting a potential anti-angiogenic treatment for choroidal neovascularization such as that seen in acute macular degeneration.[59] Finally, targeted activation of the aberrantly silenced frataxin gene with TALE-VP64 fusions is being explored to treat Friedreich's ataxia.[60]

CRISPR Repression

The ability to repress genes is possible with technologies suchas RNA interference (RNAi). However, site-specific gene silencing with engineered dCas9-repressor fusions offers an alternative to RNAi, which has been limited by inefficient

knockdown, off-target effects and toxicity associated with oversaturation of endogenous microRNA pathways.[34,61] Programmable dCas9-repressors can also target any portion of the genome, facilitating silencing of regulatory elements and noncoding genes that cannot be targeted by RNAi. Genes and regulatory elements can also be disrupted with site-specific nucleases, but nuclease-mediated genome editing is a stochastic and often inefficient process, in contrast to the generally more uniform gene repression techniques.

Localization of just a dCas9 protein without an effector domain to promoter regions or regions downstream of the TSS can silence gene expression by steric interference of transcription factor binding and RNA polymerase elongation.[36,38,62,63] For example, dCas9 targeted to the Nanog enhancer disrupts binding of endogenous activating transcription factors and silences Nanog expression.[64] However, gene repression by steric hindrance alone is often not sufficient for robust silencing. Effectors that recruit endogenous epigenetic modifiers of histone marks and DNA methylation, leading to chromatin condensation, typically generate more potent silencing.[65–68] The silencing domain most commonly used with DNA-binding domains (DBDs) is the Krüppel-associated box (KRAB), a naturally occurring motif in mammalian zinc-finger transcription factors.[69–71] Localizing KRAB to DNA initiates a heterochromatin-forming complex that includes the histone methyltransferase SETDB1 and the histone deacetylase (HDAC) NuRD complex.[72–75] In contrast to engineered activator platforms that benefit from multiplexing for potent activation, KRAB fusions do not seem to act synergistically and can readily achieve tenfold or greater repression of endogenous genes with recruitment of a single effector.[64–66,69] In addition to being able to silence

genes from promoters, KRAB can effectively repress distal and proximal gene regulatory elements including enhancers.[76]

Alternatively, effector domains that directly catalyze repressive DNA methylation or histone modifications can be fused to dCas9 to create custom epigenetic silencing proteins. Synthetic ZFPs tethered to DNMT3a catalyze DNA methylation and suppress transcription from endogenous gene promoters.[58,77–79] LSD1 has been tethered to TALEs and dCas9 DBDs to remove H3K4me from active enhancers and suppress downstream target expression.[80] A variety of ZFP- and TALE-based histone methyltransferase fusions have also been created that repress endogenous gene transcription by depositing H3K9 monomethylation, dimethylation and trimethylation at promoter regions.[67,81,82]

Each type of epigenetic repressor provides unique advantages conferred by its mechanism of action. The KRAB domain acts as a recruiter, and the complex of heterochromatin-modifying enzymes that it localizes to target DNA probably contributes to its versatility and potency for silencing protein-coding genes, noncoding RNA and regulatory elements.[49,65,69,71] Lastly, the temporal stability and heritability of silencing are important considerations. Silencing induced by the Kap1 complex associated with the KRAB domain can persist through cell replication, and H3K9 methylation and HP1 localization are properties of constitutive heterochromatin.[83] Silencing with KRAB- and H3K9 methyltransferase-based repressors, however, has been reversible after removal of the dCas9-repressor fusion in some studies.[69,81] Alternatively, targeted DNA methylation marks have the potential to be inherited by daughter cells and persist long

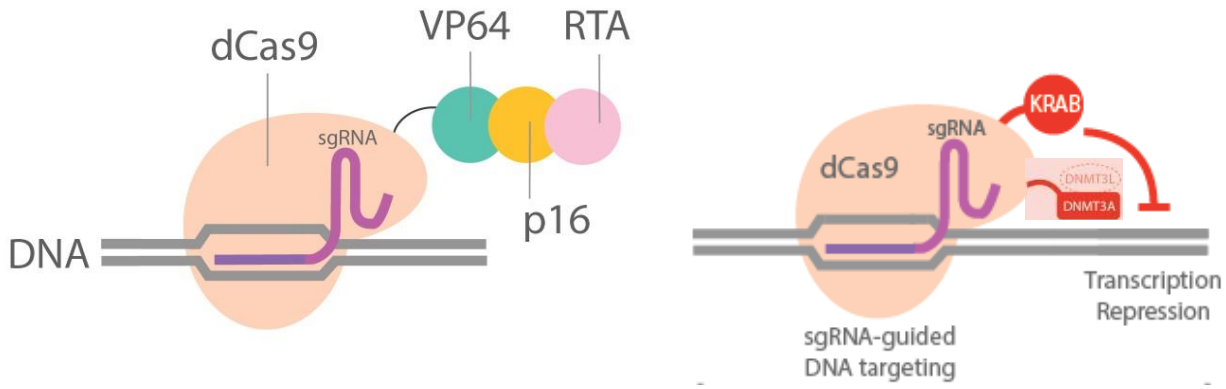
term—a benefit for applications in which stable and heritable suppression is desired.[77,79] An important goal for future work is to generate a better understanding of how the heritability and persistence of gene modulation and epigenetic marks are affected by the duration and magnitude of the activity of the epigenetic effector, the local chromatin environment, the cell type, the presence of endogenous cofactors, and the identity of the epigenetic marks being created.

Targeted repressors can also suppress detrimental gene products associated with disease progression. Engineered ZFP repressors have been designed to silence oncogenes and have been effective at slowing the growth of cancer cells in mouse models.[78] Synthetic ZFP repressors were also engineered to selectively silence mutant *Htt* in a mouse model of Huntington's disease.[65] In this study, several ZFP lengths were tested to develop a repressor with specific activity at the longer CAG repeats found in disease-causing mutant *Htt* alleles. This strategy may also be applicable for treating other gain-of-function genetic diseases, such as fragile X syndrome and myotonic dystrophy, and was recently evaluated in cells from individuals with facioscapulohumeral muscular dystrophy. Pharmacologic modulation of DNA methylation and histone modification has shown preclinical promise for treating tumor progression and neurodegenerative disorders such as Alzheimer's, Parkinson's and Huntington's disease.[84] However, the small-molecule drugs used in these therapies typically act by broadly inhibiting the enzymatic activity of epigenetic effectors, and doses are often limited by toxicity after systemic administration. Programmable epigenetic modification with designer DBDs could potentiate targeted therapy by

treating the specific histone and DNA methylation marks that contribute to the progression of such diseases.

For the NF1 part of this thesis, we employed a CRISPR repression system termed CRISP-Off, which consists of a dCas9 protein fused to a KRAB and DNMT3A/L domains.

Figure1.2 - CRISPR Activation (Left) utilizes a dCas9 protein fused to transcription upregulators such as VP64, p16, and RTA (VPR). CRISPR Repression (Right) utilizes a dCas9 protein fused to transcriptional repressors, such as KRAB and DNMT3A/L.



REFERENCES

1. History of Genetic Engineering and the Rise of Genome Editing Tools. *Synthego* https://www.synthego.com/learn/genome-engineering-history.
2. Cyclic polyamides for recognition in the minor groove of DNA. | PNAS. https://www.pnas.org/doi/abs/10.1073/pnas.92.22.10389.
3. Peptide nucleic acid-targeted mutagenesis of a chromosomal gene in mouse cells | PNAS. https://www.pnas.org/doi/abs/10.1073/pnas.95.4.1398.
4. Gottesfeld, J. M., Neely, L., Trauger, J. W., Baird, E. E. & Dervan, P. B. Regulation of gene expression by small molecules. *Nature* **387**, 202–205 (1997).
5. Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).
6. Gjaltema, R. A. F. & Rots, M. G. Advances of epigenetic editing. *Curr. Opin. Chem. Biol.* **57**, 75–81 (2020).
7. Scherer, S. & Davis, R. W. Replacement of chromosome segments with altered DNA sequences constructed in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 4951–4955 (1979).
8. Rong, Y. S. & Golic, K. G. Gene targeting by homologous recombination in Drosophila. *Science* **288**, 2013–2018 (2000).
9. Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. & Kucherlapati, R. S. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* **317**, 230–234 (1985).
10. Thomas, K. R., Folger, K. R. & Capecchi, M. R. High frequency targeting of genes to specific sites in the mammalian genome. *Cell* **44**, 419–428 (1986).
11. Mansour, S. L., Thomas, K. R. & Capecchi, M. R. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature* **336**, 348–352 (1988).
12. Felsenfeld: Formation of a three-stranded polynucleotide... - Google Scholar. https://scholar.google.com/scholar_lookup?title=Formation+of+a+three-stranded+polynucleotide+molecule&author=G.+Felsenfeld&author=D.+R.+Davies&author=A.+Rich&publication_year=1957&journal=J.+Am.+Chem.+Soc.&pages=2023-2024&doi=10.1021%2Fja01565a074.
13. Varshavsky, A. Discovering the RNA Double Helix and Hybridization. *Cell* **127**, 1295–1297 (2006).
14. Yang, J., Zimmerly, S., Perlman, P. S. & Lambowitz, A. M. Efficient integration of an intron RNA into double-stranded DNA by reverse splicing. *Nature* **381**, 332–335 (1996).
15. Zimmerly, S. *et al.* A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell* **83**, 529–538 (1995).
16. Sullenger, B. A. & Cech, T. R. Ribozyme-mediated repair of defective mRNA by targeted, trans-splicing. *Nature* **371**, 619–622 (1994).
17. Jacquier, A. & Dujon, B. An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell* **41**, 383–394 (1985).
18. Chevalier, B. S. *et al.* Design, activity, and structure of a highly specific artificial endonuclease. *Mol. Cell* **10**, 895–905 (2002).

19. Kim, Y. G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1156–1160 (1996).
20. Bibikova, M., Beumer, K., Trautman, J. K. & Carroll, D. Enhancing gene targeting with designed zinc finger nucleases. *Science* **300**, 764 (2003).
21. Bibikova, M., Golic, M., Golic, K. G. & Carroll, D. Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics* **161**, 1169–1175 (2002).
22. Christian, M. *et al.* Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**, 757–761 (2010).
23. Moscou, M. J. & Bogdanove, A. J. A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501 (2009).
24. Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–1512 (2009).
25. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiol. Read. Engl.* **151**, 2551–2561 (2005).
26. Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* **1**, 7 (2006).
27. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
28. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602–607 (2011).
29. Campa, C. C., Weisbach, N. R., Santinha, A. J., Incarnato, D. & Platt, R. J. Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nat. Methods* 1–7 (2019) doi:10.1038/s41592-019-0508-6.
30. Lattanzi, A. *et al.* Development of β-globin gene correction in human hematopoietic stem cells as a potential durable treatment for sickle cell disease. *Sci. Transl. Med.* **13**, eabf2444 (2021).
31. Pavel-Dinu, M. *et al.* Gene correction for SCID-X1 in long-term hematopoietic stem cells. *Nat. Commun.* **10**, 1634 (2019).
32. Marsh, S., Hanson, B., Wood, M. J. A., Varela, M. A. & Roberts, T. C. Application of CRISPR-Cas9-Mediated Genome Editing for the Treatment of Myotonic Dystrophy Type 1. *Mol. Ther. J. Am. Soc. Gene Ther.* **28**, 2527–2539 (2020).
33. Chen, M. *et al.* CRISPR-Cas9 for cancer therapy: Opportunities and challenges. *Cancer Lett.* **447**, 48–55 (2019).
34. Jackson, A. L. *et al.* Widespread siRNA 'off-target' transcript silencing mediated by seed region sequence complementarity. *RNA N. Y. N* **12**, 1179–1187 (2006).
35. Thakore, P. I., Black, J. B., Hilton, I. B. & Gersbach, C. A. Editing the epigenome: technologies for programmable transcription and epigenetic modulation. *Nat. Methods* **13**, 127–137 (2016).
36. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).

37. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014).
38. Choo, Y., Sánchez-García, I. & Klug, A. In vivo repression by a site-specific DNA-binding protein designed against an oncogenic sequence. *Nature* **372**, 642–645 (1994).
39. Pavletich, N. P. & Pabo, C. O. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science* **252**, 809–817 (1991).
40. Liu, Q., Segal, D. J., Ghiara, J. B. & Barbas, C. F. Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 5525–5530 (1997).
41. Ingles, C. J., Shales, M., Cress, W. D., Triezenberg, S. J. & Greenblatt, J. Reduced binding of TFIID to transcriptionally compromised mutants of VP16. *Nature* **351**, 588–590 (1991).
42. Zhang, F. *et al.* Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.* **29**, 149–153 (2011).
43. Beerli, R. R., Dreier, B. & Barbas, C. F. Positive and negative regulation of endogenous genes by designed transcription factors. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 1495–1500 (2000).
44. Maeder, M. L. *et al.* Robust, synergistic regulation of human gene expression using TALE activators. *Nat. Methods* **10**, 243–245 (2013).
45. Perez-Pinera, P. *et al.* Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nat. Methods* **10**, 239–242 (2013).
46. Maeder, M. L. *et al.* CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* **10**, 977–979 (2013).
47. Mali, P. *et al.* CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **31**, 833–838 (2013).
48. Gao, X. *et al.* Reprogramming to pluripotency using designer TALE transcription factors targeting enhancers. *Stem Cell Rep.* **1**, 183–197 (2013).
49. Hilton, I. B. *et al.* Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* **33**, 510–517 (2015).
50. Polstein, L. R. *et al.* Genome-wide specificity of DNA binding, gene regulation, and chromatin remodeling by TALE- and CRISPR/Cas9-based transcriptional activators. *Genome Res.* **25**, 1158–1169 (2015).
51. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588 (2015).
52. Cheng, A. W. *et al.* Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.* **23**, 1163–1171 (2013).
53. Chakraborty, S. *et al.* A CRISPR/Cas9-based system for reprogramming cell lineage specification. *Stem Cell Rep.* **3**, 940–947 (2014).
54. Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S. & Vale, R. D. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**, 635–646 (2014).

55. Chavez, A. *et al.* Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* **12**, 326–328 (2015).
56. Kelly, T. K., De Carvalho, D. D. & Jones, P. A. Epigenetic modifications as therapeutic targets. *Nat. Biotechnol.* **28**, 1069–1078 (2010).
57. Gräslund, T., Li, X., Magnenat, L., Popkov, M. & Barbas, C. F. Exploring strategies for the design of artificial transcription factors: targeting sites proximal to known regulatory regions for the induction of gamma-globin expression and the treatment of sickle cell disease. *J. Biol. Chem.* **280**, 3707–3714 (2005).
58. Siddique, A. N. *et al.* Targeted methylation and gene silencing of VEGF-A in human cells by using a designed Dnmt3a-Dnmt3L single-chain fusion protein with increased DNA methylation activity. *J. Mol. Biol.* **425**, 479–491 (2013).
59. Yokoi, K. *et al.* Gene transfer of an engineered zinc finger protein enhances the anti-angiogenic defense system. *Mol. Ther. J. Am. Soc. Gene Ther.* **15**, 1917–1923 (2007).
60. Chapdelaine, P., Coulombe, Z., Chikh, A., Gérard, C. & Tremblay, J. P. A Potential New Therapeutic Approach for Friedreich Ataxia: Induction of Frataxin Expression With TALE Proteins. *Mol. Ther. Nucleic Acids* **2**, e119 (2013).
61. Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using CRISPR-Cas9. *Nat. Rev. Genet.* **16**, 299–311 (2015).
62. Li, Y. *et al.* Modular construction of mammalian gene circuits using TALE transcriptional repressors. *Nat. Chem. Biol.* **11**, 207–213 (2015).
63. Larson, M. H. *et al.* CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat. Protoc.* **8**, 2180–2196 (2013).
64. Gao, X. *et al.* Comparison of TALE designer transcription factors and the CRISPR/dCas9 in regulation of gene expression by targeting enhancers. *Nucleic Acids Res.* **42**, e155 (2014).
65. Garriga-Canut, M. *et al.* Synthetic zinc finger repressors reduce mutant huntingtin expression in the brain of R6/2 mice. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E3136-3145 (2012).
66. Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013).
67. Snowden, A. W., Gregory, P. D., Case, C. C. & Pabo, C. O. Gene-specific targeting of H3K9 methylation is sufficient for initiating repression in vivo. *Curr. Biol. CB* **12**, 2159–2166 (2002).
68. Kearns, N. A. *et al.* Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat. Methods* **12**, 401–403 (2015).
69. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647–661 (2014).
70. Beerli, R. R., Segal, D. J., Dreier, B. & Barbas III, C. F. Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14628–14633 (1998).
71. Cong, L., Zhou, R., Kuo, Y.-C., Cunniff, M. & Zhang, F. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat. Commun.* **3**, 968 (2012).

72. Sripathy, S. P., Stevens, J. & Schultz, D. C. The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. *Mol. Cell. Biol.* **26**, 8623–8638 (2006).
73. Groner, A. C. *et al.* KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS Genet.* **6**, e1000869 (2010).
74. Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. & Rauscher, F. J. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* **16**, 919–932 (2002).
75. Reynolds, N. *et al.* NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. *EMBO J.* **31**, 593–605 (2012).
76. Thakore, P. I. *et al.* Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods* **12**, 1143–1149 (2015).
77. Rivenbark, A. G. *et al.* Epigenetic reprogramming of cancer cells via targeted DNA methylation. *Epigenetics* **7**, 350–360 (2012).
78. Stolzenburg, S. *et al.* Targeted silencing of the oncogenic transcription factor SOX2 in breast cancer. *Nucleic Acids Res.* **40**, 6725–6740 (2012).
79. Stolzenburg, S. *et al.* Stable oncogenic silencing in vivo by programmable and targeted de novo DNA methylation in breast cancer. *Oncogene* **34**, 5427–5435 (2015).
80. Mendenhall, E. M. *et al.* Locus-specific editing of histone modifications at endogenous enhancers. *Nat. Biotechnol.* **31**, 1133–1136 (2013).
81. Kungulovski, G. *et al.* Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained. *Epigenetics Chromatin* **8**, 12 (2015).
82. Konermann, S. *et al.* Optical control of mammalian endogenous transcription and epigenetic states. *Nature* **500**, 472–476 (2013).
83. Hathaway, N. A. *et al.* Dynamics and memory of heterochromatin in living cells. *Cell* **149**, 1447–1460 (2012).
84. Wang, X. *et al.* Intraosseous delivery of lentiviral vectors targeting factor VIII expression in platelets corrects murine hemophilia A. *Mol. Ther. J. Am. Soc. Gene Ther.* **23**, 617–626 (2015).

Chapter 2- Developing an Epigenetic Therapy for Neurofibromatosis Type 1 (NF1)

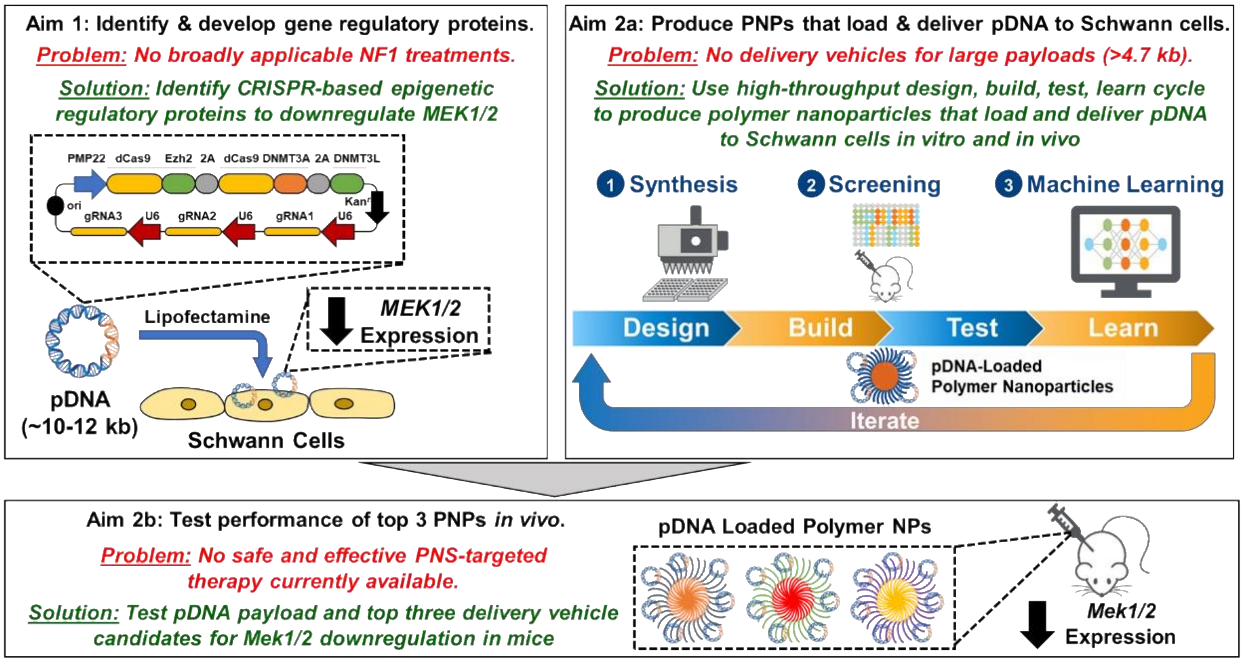
ABSTRACT

Neurofibromatosis type 1 (NF1) is a genetic disorder that affects the nervous system, skin, and other organs. It is caused by mutations in the NF1 gene, which helps regulate cell growth and division. NF1 can cause a wide range of symptoms, including the development of tumors on or under the skin, freckling of the skin in certain patterns, and abnormalities in bone growth. It can also lead to learning disabilities, seizures, vision and hearing problems, and other neurological issues. Symptoms can vary widely between individuals and can range from mild to severe. There is no cure for NF1, but treatment focuses on managing symptoms and preventing complications. This may involve surgery to remove tumors, medications to manage pain or other symptoms, and monitoring for complications such as high blood pressure or vision problems. An epigenetic therapy can provide a way to address the disorder.

Despite recent advances in genome editing, two key challenges remain to translate a NF1 gene therapy to the clinic: 1) development of a broadly applicable therapeutic approach for all individuals with NF1 rather than mutation-specific approaches, and 2) the ability to load and deliver large payloads to specific cell types, such as Schwann cells. This project leverages a multidisciplinary collaboration between Battelle Memorial Institute (BMI) and UC Davis (UCD) to overcome these challenges. The overall goal of this proposal is to create targeted epigenetic regulatory proteins (ERPs) to be transient therapeutic payloads that cause persistent down-regulation of mitogen-activated protein kinases 1 and 2 (MEK1/2) and to deliver plasmid DNA (pDNA) encoding these ERPs to Schwann cells (SCs) in both human and mouse models. In the work described here, we designed, developed, and screened 20 guide

RNAs to identify the top two strongest repressor candidates. Once the guide RNAs were ready, we used qPCR to measure the repression capabilities of the different RNA guide candidates, and mapped the theoretical plasmid payloads (*e.g.*, 12 kbp and 14.5 kbp) to be used for this project.

Figure 2.1- A graphical abstract of the aims of this project. This chapter addresses work done towards Aim1, identifying and developing gene regulatory proteins to reduce *MEK1/2* expression. This project is a collaboration between the Segal Lab at UC Davis and the Battelle institute.



INTRODUCTION

Neurofibromatosis type 1 (NF1) is a complex autosomal dominant disorder caused by germline mutations in the *NF1* tumor suppressor gene. Nearly all individuals with neurofibromatosis type 1 develop pigmentary lesions (café-au-lait macules, skinfold freckling and Lisch nodules) and dermal neurofibromas. Some individuals develop skeletal abnormalities (scoliosis, tibial pseudarthrosis and orbital dysplasia), brain tumors (optic pathway gliomas and glioblastoma), peripheral nerve tumors (spinal neurofibromas, plexiform neurofibromas and malignant peripheral nerve sheath tumors), learning disabilities, attention deficits, and social and behavioral problems, which can negatively affect quality of life.[1] With the identification of *NF1* and the generation of accurate preclinical mouse strains that model some of these clinical features, therapies that target the underlying molecular and cellular pathophysiology for neurofibromatosis type 1 are becoming available. Although no single treatment exists, current clinical management strategies include early detection of disease phenotypes (risk assessment) and biologically targeted therapies. Similarly, new medical and behavioral interventions are emerging to improve the quality of life of patients. Although considerable progress has been made in understanding this condition, numerous challenges remain. This chapter of the thesis presents work towards the development of an epigenetic approach to treating the disorder, by CRISPR repression of the *Map2k1* and *Map2k2* genes, which act downstream of the NF1 protein in the Ras-ERK pathway. Such an approach can potentially be used in conjunction with existing therapies or NF1 specific therapies to address the disorder.

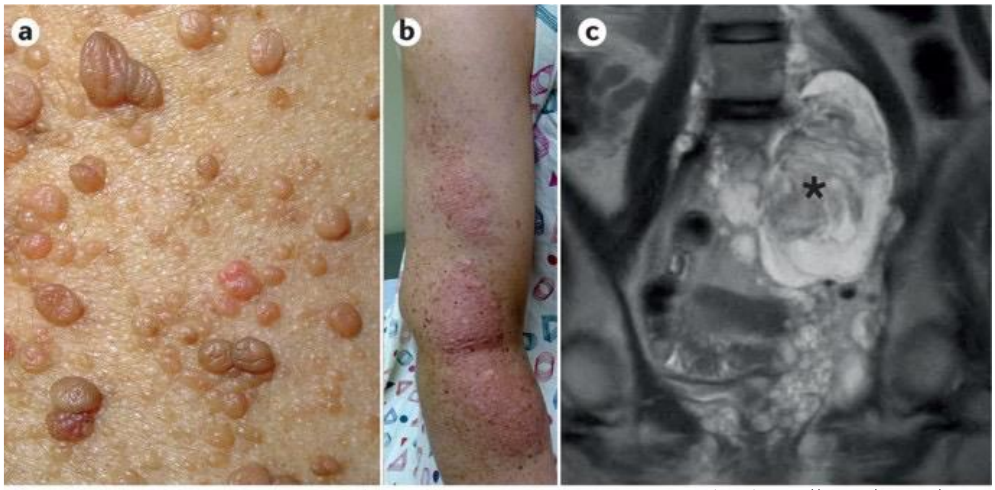
Overview of NF1

NF1 is an autosomal dominant disorder that affects many organ systems. It is one of three conditions described under the broad heading of the neurofibromatosis. The other two, neurofibromatosis type 2 and schwannomatosis, are clinically and genetically distinct from neurofibromatosis type 1.[2]

The defining feature of neurofibromatosis type 1 is the neurofibroma, a nerve sheath tumor that forms in association with spinal, peripheral, or cranial nerves. An example of these neurofibromas is shown in **Figure 2.2**. Other symptoms include pigmentary abnormalities, low-grade gliomas and skeletal dysplasias, as well as the involvement of numerous other organ systems. The condition is gradually progressive over the lifetime of an individual, although the specific manifestations, rate of progression and severity of complications vary widely. At present, no definitive treatment is available, and clinical management is typically limited to surveillance and symptomatic treatment, usually surgical, for specific complications.

The gene responsible for neurofibromatosis type 1 (*NF1*, which encodes neurofibromin) was identified in 1990, and its function and role in the formation of tumors and the other manifestations of neurofibromatosis type 1 have been under intensive study.[3] With an increasing understanding of the mechanisms that underlie the pathogenesis of neurofibromatosis type 1 clinical features, numerous targeted therapies have emerged, which are now being evaluated in preclinical models and in phase II clinical trials.

Figure 2.2- Neurofibromas can grow as discrete nodules, which can be found on the skin as dermal neurofibromas (a). Neurofibromas can also involve multiple branches of larger nerves, referred to as plexiform neurofibromas, which can be seen on the skin (b) or can be internal. C shows an MRI scan of an intra-abdominal plexiform neurofibroma (asterisk).



Gutmann, D., Ferner, R., Listernick, R. *et al.* Neurofibromatosis type 1. *Nat Rev Dis Primers* **3**, 17004 (2017). https://doi.org/10.1038/nrdp.2017.4

The average global prevalence of neurofibromatosis type 1 is ~1 case per 3,000 individuals.[4] 50% of cases of neurofibromatosis type 1 are familial (inherited) and the remainder arise from a *de novo NF1* mutation. The life expectancy of individuals with neurofibromatosis type 1 is reduced by ~8–21 years and an excess of deaths occurs in younger individuals (<40 years of age), compared with the general population; the most common cause of early death is malignant neoplasm. Individuals have an increased risk for malignant and non-malignant conditions compared with the general population.[4–9]

NF1 Genetics

Neurofibromatosis type 1 is an autosomal dominant genetic condition, such that all people with a germline *NF1* mutation have this disease. However, patients can show extreme variability in their clinical features, even in individuals from the same family with

an identical germline *NF1* mutation. In addition to the more common generalized neurofibromatosis type 1, some people harbor features of neurofibromatosis type 1 that are restricted to one segment of their body ('segmental' or mosaic neurofibromatosis type 1), which probably arises from a somatic mutation in *NF1* that occurs during fetal development. In each of these situations, the *NF1* mutation is thought to result in loss of neurofibromin function from that mutated allele.

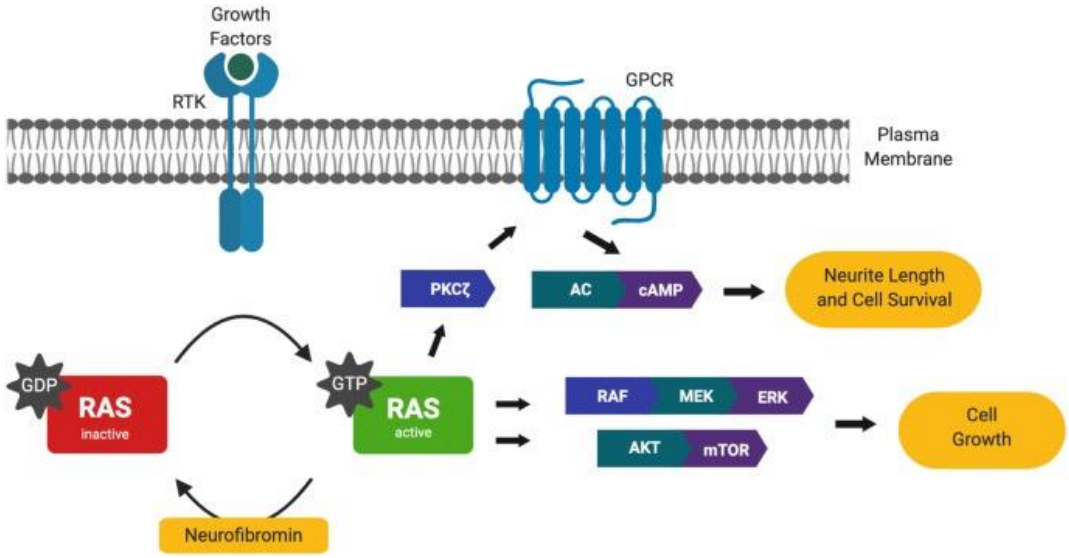
The complexity of molecular testing to identify causative mutations in *NF1* is related to the large size of the gene (~60 exons), the relative lack of mutation hotspots and the diversity of the pathogenetic mutations. A multi-step approach is required, with analysis of blood genomic DNA and mRNA, as well as fluorescent *in situ* hybridization testing for whole *NF1* deletions.[10] This strategy identifies >95% of causative mutations, but in people with segmental neurofibromatosis type 1, analysis of affected tissues is necessary as the *NF1* mutation is not usually detected in the blood.

To date, >7,000 people with neurofibromatosis type 1 have undergone genetic testing, and >3,000 different germline *NF1* mutations have been identified. Although genotype–phenotype correlations are uncommon in neurofibromatosis type 1, three well-established correlations have been identified. Individuals with 1.4 Mb deletions that encompass the entire *NF1* gene typically show facial dysmorphism, reduced intellectual abilities and an increased incidence of cancer.[11] In addition, ~1% of people with neurofibromatosis type 1 have mutations that affect codon 1809 and typically present with café-au-lait macules (CALMs), short stature and pulmonic stenosis, but lack externally visible plexiform or dermal neurofibromas. In addition, another mutation has been associated with an absence of neurofibromas.[12,13]

The precise mechanisms underlying the development of the clinical manifestations of neurofibromatosis type 1 can vary, such that some manifestations result from haploinsufficiency of *NF1*, whereas others require biallelic *NF1* inactivation or the addition of modifying factors, such as hormones or other genetic alterations. For example, biallelic *NF1* inactivation is required for the development of CALMs and neurofibromas, but cooperating genetic alterations, such as *TP53* mutation, are required for the formation of malignant peripheral nerve sheath tumors (MPNSTs).

With the discovery of *NF1* in 1990, it became possible to envision a future in which treatments might emanate from a deeper understanding of the function of neurofibromin.[14,15] Neurofibromin is expressed in many cell types, including neurons, glial cells, immune cells, endothelial cells and in cells of the adrenal medulla, but probably functions differently in distinct cell types. Close examination of the predicted amino acid sequence of neurofibromin revealed that a small 300-residue domain of neurofibromin was structurally like a family of proteins that function as negative regulators of the RAS proto-oncogene. These proteins, termed GTPase-activating proteins, inactivate RAS by accelerating the conversion of active GTP-bound RAS to the inactive GDP-bound form. In this manner, loss of neurofibromin expression, as seen in tumors associated with neurofibromatosis type 1, is predicted to lead to increased cell growth and survival through hyperactivation of RAS.[16,17] RAS then transmits its growth-promoting signal through the AKT–mechanistic target of rapamycin (mTOR) and MEK–extracellular signal-regulated kinase (ERK) effector pathways.[18,19] This pathway is outlined in **Figure 2.3** below.

Figure 2.3- Neurofibromin is a GTPase-activating protein of Ras (Ras-GAP). It downregulates the Ras signaling pathway by promoting the hydrolysis of the active form of Ras (GTP-bound Ras) to an inactive form of Ras (GDP-bound Ras) by increasing the intrinsic GTPase activity of Ras. With mutated NF1, this inactivation is dysfunctional, leading to increased downstream effect, such as increased cell growth and cell survival, leading to the cancer phenotypes described earlier.



Selumetinib and Repression of *Map2k1/2k2*

NF1 tumors can be characterized by elevated Ras and MAPK (mitogen-activated protein kinase) signaling. But therapeutic agents which target the Ras pathway have provided limited responses. The genetic cause of NF1 is mostly due to constitutional heterozygous loss of function mutations of tumor suppressor gene NF1. There is somatic inactivation of wild type allele in Schwann cells of neurofibromas (according to Knudson 2-hits hypothesis).[17] This explains the hyperactivation of RAS MAPK pathway in NF1.[20] In this condition, selumetinib, an oral inhibitor of MEK (MAPK Kinases)1 and MEK (MAPK Kinases)2 showed response in children with NF1, on the basis of promising activity in adults with several types of advanced-stage cancers.[21]

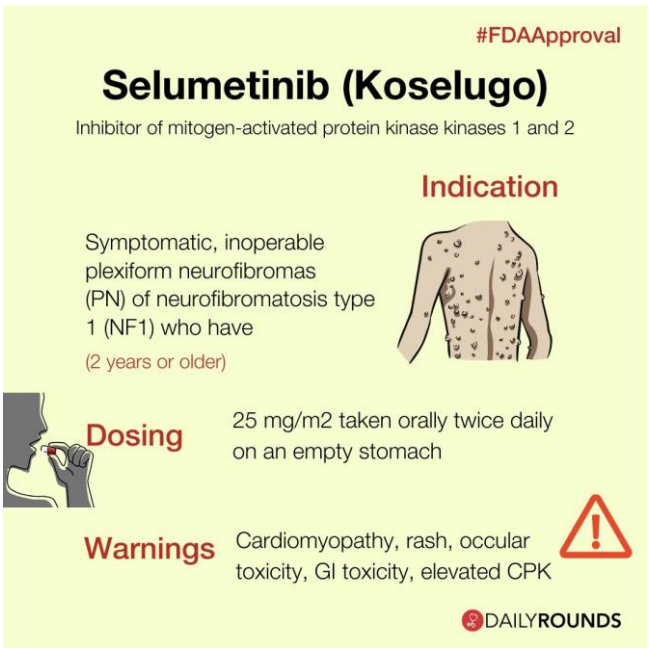
Three major subfamilies of the MAPK pathways have been identified; JNK/SAPK's (c-Jun N-terminal kinase/Stress activated protein kinase), p38 MAPK and MEK/ERK (extracellular signal-related kinase). MEK contains two consensus kinase motifs, which are involved in the phosphorylation of serine/threonine and tyrosine residues. Two homologues exist, MEK1 and MEK2, which have only one known substrate, ERK1/2, which has multiple downstream effectors involved in a number of cellular functions including transcription, cell cycle progression and cell motility. Once activated, ERK, along with RAF (rapidly accelerated fibrosarcoma) and MEK, migrate to the nucleus where they activate cyclin D1 and down regulate p27 thus driving cell proliferation. Activation may also lead to inhibition of apoptosis or activation of anti-apoptotic proteins.[22] MEK1/2 were the targets we chosen to repress.

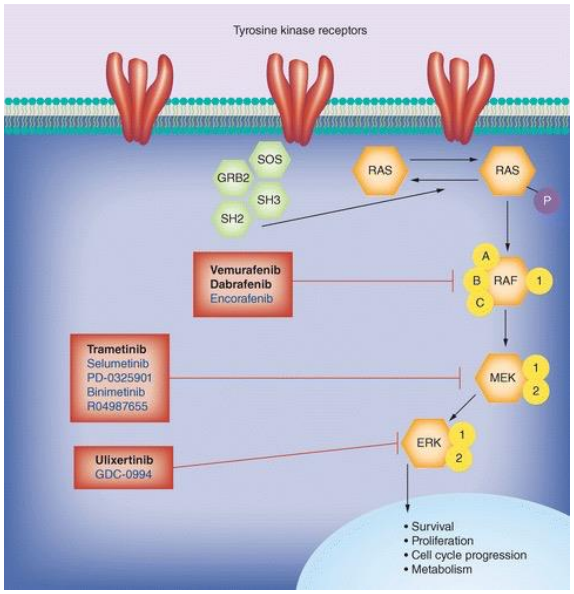
MEK1/2 proteins are upstream regulators of the ERK pathway. Both MEK and ERK are critical components of the Ras-regulated RAF-MEK-ERK pathway, which is often activated in different types of cancers. Oral dosing of selumetinib inhibited ERK phosphorylation, and reduced neurofibroma numbers, volume, and proliferation in animal models.[23]

The major limitation of this molecule regarding its mechanism of action is dose dependent effect of MEK inhibition in growth of neurofibroma. It is a cytotoxic agent. In the phase I study by Dombi et al., slow tumor regrowth was observed in several patients in whom dose reduction or interruption were done due to any toxicity. Prolonged use of selumetinib may be required to get response in progressive tumors.[23] In animal models, intermittent dosing was found to be effective which suggested that limited inhibition of

MEK is also sufficient to shrink tumors.[24] Currently, an intermittent dosing strategy is being evaluated in some ongoing trials.[22]

Figure 2.4- Selumetinib (a Map2k1/Map2k2 repressor) has been shown to reduce plexiform lesions in NF1 and improve patient quality of life.





The success of selumetinib (a MEK inhibitor) is the proof-of-concept we relied on in designing our approach: Using CRISPR repression to reduce the expression of the *Map2k1* and *Map2k2* genes, a part of the RAS-ERK pathway downstream of neurofibromin. Our approach offers several advantages over selumetinib, in that it is a genetic therapy where the cargo can be expressed by a Schwann cell specific promoter, preventing expression of the therapeutic systemically, theoretically limiting the side-effect profile of such a therapy.

CRISPR Repression of *Map2k1*/*Map2k2*

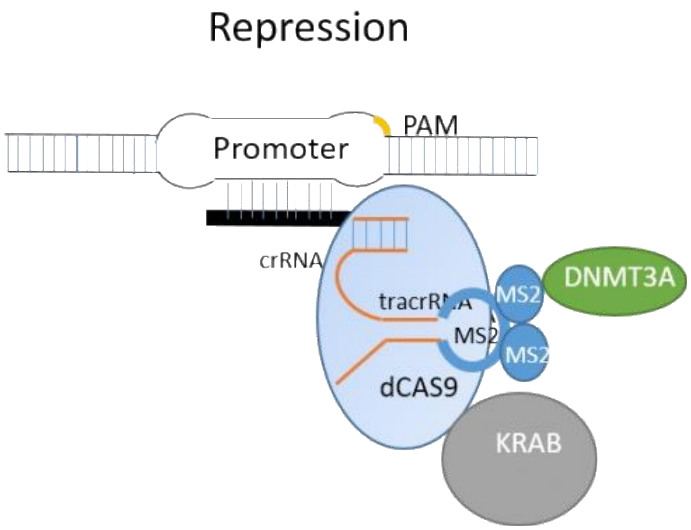
In order to repress the expression of *Map2k1*/*Map2k2*, as a way of creating an epigenetic therapy that mimics the success of selumetinib, we employed a CRISPR repression system termed CRISPR-Off.[25] CRISPR-Off is a programmable epigenetic memory writer consisting of a single dead Cas9 fusion protein that establishes DNA methylation and repressive histone modifications. Transient CRISPR-Off expression initiates highly specific DNA methylation and gene repression that has been shown to be maintained through cell division at some genetic loci.[26]

CRISPR repression can be achieved by fusing a Krüppel-associated box (KRAB) domain to a deactivated Cas9 (dCas9) protein, which is then targeted to a specific genomic locus using guide RNAs. The KRAB domain interacts with endogenous co-repressor proteins, such as KAP1, which recruit chromatin-modifying enzymes, including DNA methyltransferase 3A and lysine-specific histone demethylase 1A (LSD1), to the targeted genomic locus. DNA methyltransferase 3A then catalyzes the addition of a methyl group to cytosine residues in the DNA, leading to transcriptional repression by preventing the binding of transcription factors and other regulatory proteins.[27]

The KRAB-dCas9 system can be further enhanced by fusing additional effector domains, such as DNMT3A/L. The DNMT3A/L domain is a catalytically active DNA methyltransferase domain that can also methylate cytosine residues in the targeted genomic locus. By fusing the DNMT3A/L domain to KRAB-dCas9, transcriptional repression can be enhanced, leading to more efficient gene silencing. Overall, the

KRAB-dCas9 system with DNMT3A/L effector domain is a powerful tool for achieving targeted gene repression, and has potential applications in gene therapy and genetic research.[28,29]

Figure 2.5- CRISPROff can repress the expression of endogenous genes by fusing dCas9 to KRAB/DNMT3A/L domains and designing a guide RNA to target promoters



MATERIALS AND METHODS

CRISPR/Cas9 Guide Design

CRISPR guide RNAs were designed for a Cas9 system. The bioinformatics tool ChopChop was utilized to identify all the potential guides 400bp upstream of the promoter for both the *Map2k1* and *Map2k2* genes in the mouse genome.[30] **Figures 2.6 and 2.7** show the output of the ChopChop analysis for *Map2k1* and *Map2k2* respectively. 10 guides were chosen for screening for both genes. Although the ChopChop algorithm ranks guides based off potential off-target affects and predicted activity levels, guides were manually chosen to include guides in different regions of the promoter to account for any variation in activity due to steric blocking by other transcription factors.

Following identification of guide sequences *in silico*, each guide was cloned into a vector expressing the sgRNA. Oligonucleotides were ordered from IDT containing the 20bp spacer designed from ChopChop, with overhangs enabling sticky-end ligation into the vector lenti-sgRNA puro (Addgene #10499) digested with *BsmBI*.[31] The lenti-sgRNA puro vector backbone was digested with *BsmBI*, gel purified using QIAquick Gel Extraction Kit. Oligonucleotides were phosphorylated and annealed in T4 Ligation Buffer using T4 PNK (NEB). These annealed, and phosphorylated oligonucleotides were then diluted at 1:200 before being ligated into the *BsmBI* digested lenti-sgRNA puro vector using T4 DNA Ligase (NEB). Successful insertion of the oligonucleotides containing the sgRNA sequence was confirmed via Sanger sequencing.

Cell Transfection

To screen the guides for the strongest potential repressor for each gene, Neuro2A cells were transfected using Lipofectamine 3000, with subsequent qPCR for relative gene mRNA levels. Neuro2A cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) media supplemented with 10% fetal bovine serum (FBS) and Penicillin/Streptomycin at 37C under 5% CO2. A total of 1 million cells at 60-70% confluency was transfected using Lipofectamine 3000 following the manufacturer's instructions. Transfections were performed in 12-well plates using 500ng of the sgRNA expressing vector, and 500ng of the CRISPR-Off vectors. Cells were harvested for RNA extraction at day 3 post lipofection using a Qiagen's RNeasy kit for the initial guide screen, and at days 3, 7, 14, and 21 for the long-term repression characterization study.

Fluorescence-activated cell sorting

FACS (Fluorescence-activated cell sorting) is a technique used to separate and isolate specific populations of cells based on their physical and fluorescent characteristics. Because the CRISPR-Off vector contains a blue fluorescent protein tag, cells successfully transfected will emit blue fluorescence. Two screens were conducted, the first without FACS sorting for successfully transfected cells. Based off the strongest repressors from this screen, guides were rescreened, and FACS sorting was conducted, as well as dual-guide screens in case single guides did not provide adequate repression.

The FACS sorting process involves several steps. First, the cells are suspended in a liquid stream and passed through a flow cell, where they are illuminated with a

laser. As the cells pass through the flow cell, they emit fluorescent light, which is detected by a series of photomultiplier tubes. The data from the photomultiplier tubes is processed by a computer, which analyzes the amount of blue fluorescence emitted by each cell and determines which cells should be sorted. Cells that emit blue fluorescence were directly captured in 200ul of DNA/RNA Shield (Zymo) for RNA extraction using a Qiagen RNAeasy Micro Kit for the guide screen. For the long term repression characterization study, cells were captured in 200ul of PBS supplemented with 1% bovine serum albumin (BSA), and subsequently cultured in DMEM with 10% FBS.

Quantitative PCR (qPCR)

RNA was extracted from harvested cells using Qiagen RNAeasy Micro Kit (Qiagen). RNA was then reverse transcribed using the AB high-capacity cDNA synthesis kit (ThermoFisher) according to the manufacturer's instructions. Primers for quantitative PCR (qPCR) were designed using Primer3. qPCR was performed in triplicate using PowerUp SYBR Green Master Mix (ThermoFisher) with the CFX384 Real-Time System C1000 Touch system (Bio-Rad). Gene expression analysis was performed with gene specific primers (**Table 2.1**) using three biological replicates. Relative target gene expression was calculated as the difference between the target gene and the GAPDH reference gene (dCq = Cq[target] − Cq[GAPDH]). Gene expression results are indicated as fold change to a reference sample (CRISPR-Off without targeting gRNA (NG)), using the ddCq method. Results from the guide screen are shown in **Figure 2.8**.

Table 2.1- qPCR Primers used to determine relative expression of Map2k1 and Map2k2 following transfection with CRISPR-Off and a guide expressing vector

Primer Name	Primer Sequence (5'-3')
Map2k1_qPCR_1F	AAGGTCTCCCACAAGCCATCTG
Map2k1_qPCR_1R	AGTTGCACTCGTGCAGTACCTG
Map2k2_qPCR_1F	GCCAAGCGGATTCCTGAAGACA
Map2k2_qPCR_1R	GCGAGAGTTCACCAGGATGTTG
GAPDH_1F	CATCACTGCCACCCAGAAGACTG
GAPDH_1R	ATGCCAGTGAGCTTCCCGTTCAG

Long-term Repression Characterization

The next experiment that was conducted involved characterizing the long-term repression capabilities of both *Map2k1* Guide 7 and *Map2k2* Guide 1. Neuro2A cells were lipofected with 500ng of CRISPR-Off, as well as 500ng of Guide 1 and Guide 7. Cells were FACS sorted for successfully transfected cells at day 3, and grown for three weeks. Cells were passaged every four days, and RNA extracted from a portion of the cells at D3, D7, D14, and D21. qPCR was repeated as described above. Relative target gene expression was calculated as the difference between the target gene and the GAPDH reference gene (dCq = Cq[target] − Cq[GAPDH]). Gene expression results are indicated as fold change to a reference sample (CRISPR-Off without targeting gRNA (NG)) using the ddCq method. Results from this long-term characterization study screen are shown in **Figure 2.9.**

DISCUSSION

This chapter of the thesis presents preliminary work in designing an epigenetic therapy based around the repression of *Map2k1* and *Map2k2* for the treatment of NF1. First, an array of dCas9 based sgRNAs were designed and screened in Neuro2A cells to identify the strongest repressor of *Mapk21* and *Map2k2*. As shown in **Figures 2.6** and **2.7**.

The results of the guide screen conducted in Neuro2A cells are shown in **Figure 2.8**. *Map2k1* guide 7 and *Map2k2* guide 1 were identified as the strongest repressors. In the screen conducted after FACS sorting, Guide 7 for *Map2k1* caused nearly 70% repression in *Map2k1* expression, and Guide 1 for *Map2k2* caused nearly 90% repression in *Map2k2* expression, as shown in **figure 2.8.** The results for guide 7 and guide 1 are highlighted with a red arrow. Interestingly, a dual guide approach yielded similar repression capabilities for *Map2k2*, and marginally improved repression for *Map2k1*. It still needs to be determined if these repression levels are adequate in reducing the growth of NF1 afflicted cells to reach a therapeutically relevant level.

Once guides 7 (*Map2k1)* and guide 1 (*Map2k2)* were identified, they were used to characterize the long-term repression of the CRISPR-Off system for these specific genes. As shown in **Figure 2.9.** At day 3, baseline repression for *Map2k1* was at about 65%, and over the course of 3 weeks, plateaued at around 60%. *Map2k2* had a stronger baseline repression of about 85% at day 3, and climbed incrementally over the course of three weeks to nearly 75% repression. This finding corroborates other studies in showing that CRISPR-Off is capable of repressing genes across multiple

Figure 2.6- Map2k1 Cas9 guide design for CRISPR repression

Rank	Name	Guide w/ 3' NGG PAM	Distance to TSS (bp)
1	Map2k1_1_Mus	GATGACGCGGAAATACGTCACGG	442
4	Map2k1_2_Mus	TGTCGGCGAAGCGACGGGGGCGG	393
6	Map2k1_3_Mus	CGCGCAGCTGGTTCTCCGCGTGG	126
10	Map2k1_4_Mus	GAGGCGCGGCGCATTGGCTTGGG	465
15	Map2k1_5_Mus	ACCGGTTCCTCGGGCTCGGACGG	181
17	Map2k1_6_Mus	GACGTCTGTGCGCGGCGTCTCGG	277
20	Map2k1_7_Mus	CAGGGAGCCGTCCGAGCCCGAGG	188
23	Map2k1_8_Mus	GGCCGCACTTTCTCCAAGCTGGG	237
24	Map2k1_9_Mus	AGCTGGTTCTCCGCGTGGGTTGG	121
28	Map2k1_10_Mus	AACCAGCTGCGCGCGCTCGCCGG	136

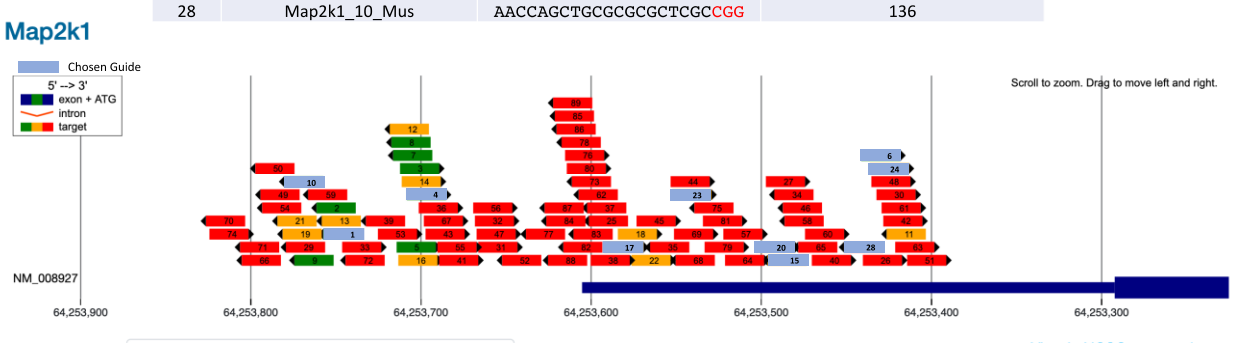


Figure 2.7- Map2k2 Cas9 guide design for CRISPR repression

Rank	Name	Guide w/ 3' NGG PAM	Distance to TSS (bp)
1	Map2k2_1_Mus	GGGCGAGCAGACCGACGCCGAGG	120
2	Map2k2_2_Mus	GACGGGCCCTCGGCGATGGTAGG	-42
3	Map2k2_3_Mus	CAGCCGGGGCCCATAGGGGGCGG	28
4	Map2k2_4_Mus	CCCCCTATGGGCCCCGGCTGAGG	25
5	Map2k2_5_Mus	CGCCTCAGCCGGGGCCCATAGGG	23
6	Map2k2_6_Mus	CGCGTCGCGCCGCCCCCTATGGG	37
7	Map2k2_7_Mus	TCTGTACGATCAGCTTCCCGGGG	345
8	Map2k2_8_Mus	TTACACACGGCAGGACCGCGAGG	401
9	Map2k2_9_Mus	TATCAACCCTACCATCGCCGAGG	35
10	Map2k2_10_Mus	CGAGGTCCCGCCCTCAACCGAGG	237

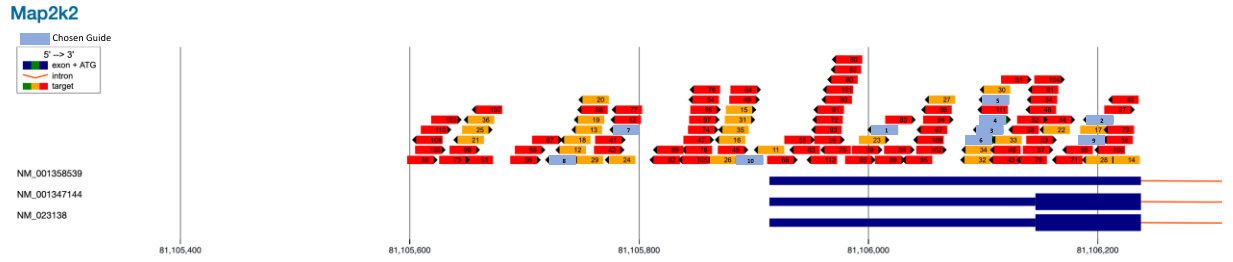
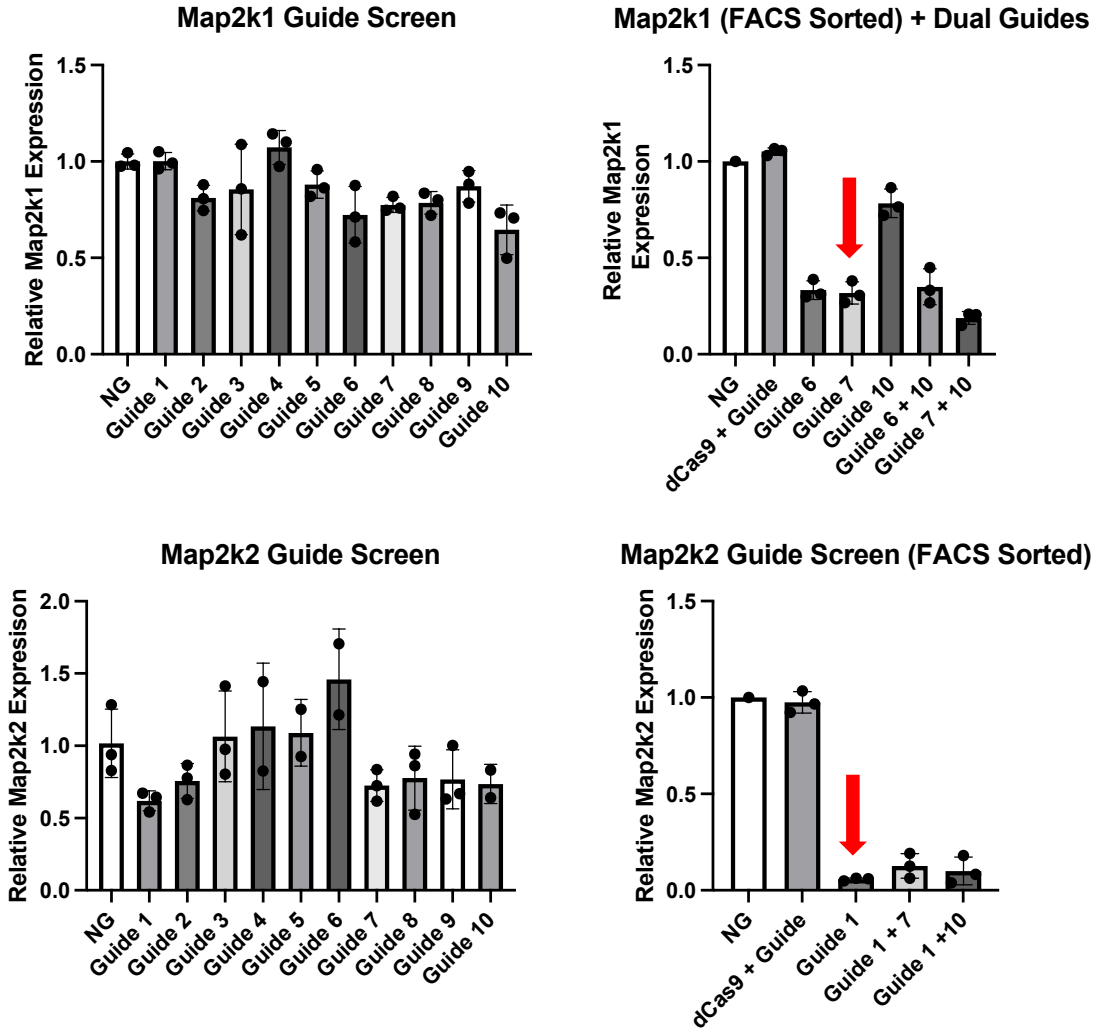


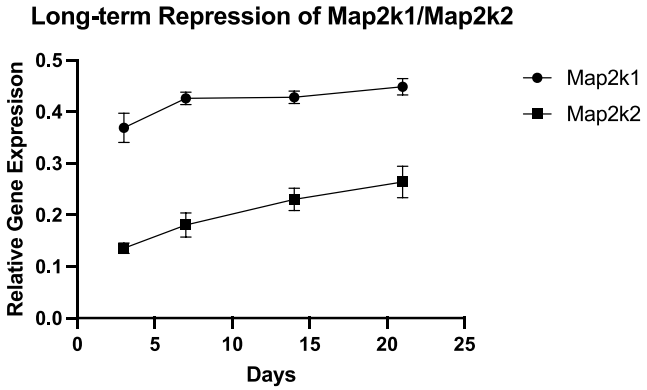
Figure 2.8- Results of the guide screen for *Map2k1* and *Map2k2*. The left graphs include all 10 guides for both *Map2k1* and *Map2k2*, done without FACS sorting. The left graphs used the strongest repressors from the right screen, but with repeated with FACS sorting. The strongest repressors (Guide 7 for *Map2k1* and Guide 1 for *Map2k2*) are highlighted with a red arrow.



generations.[26] It would be insightful to extend the length of this study to determine if repression is reverted on a longer time scale, providing insight into dosing frequency for persistent repression.

The next step in advancing the therapy is to clone a single vector that houses the entire CRISPR-Off platform as well as a guide for *Map2k1* and *Map2k2*. By using a Schwann cell specific promoter, the cargo can be selectively expressed in Schwann

Figure 2.9- Characterizing the Long-term Repression of CRISPROff for Map2k1/Map2k2



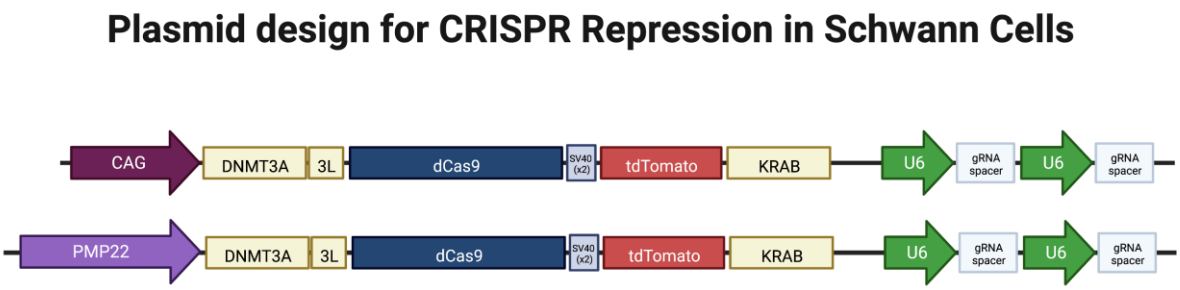
cells, limiting the side effect profile of the *Map2k1/Map2k2* repression. A schematic showing this proposed vector design is shown in **Figure 2.10**.

A serious limitation of this work is the cell type used for this preliminary work. We used Neuro2A cells for this study, as they were readily available and the transfection protocols well established. The next step is to test these tools in Schwann cells, a more clinically relevant cell line to study NF1. Additionally, off-target effects of the *Map2k1/2k2* guides need to be assayed via Chip-seq and RNA-seq.

Chip-seq and RNA-seq are two techniques that can be used to measure off-target effects of CRISPR gene editing. Chip-seq (chromatin immunoprecipitation sequencing) is a method used to identify genomic regions where a particular protein of interest is binding. In the context of CRISPR-Off-target effects, Chip-seq can be used to identify genomic regions where the CRISPR complex is binding, indicating potential off-target effects. By comparing Chip-seq data from cells that have been treated with CRISPR to cells that have not been treated, researchers can identify potential off-target sites. RNA-seq (RNA sequencing) is a technique used to identify the types and amounts of RNA molecules present in a sample. In the context of CRISPR-Off-target effects,

RNA-seq can be used to identify changes in gene expression that may be caused by off-target effects. By comparing RNA-seq data from cells that have been treated with CRISPR to cells that have not been treated, researchers can identify changes in gene expression that may be caused by off-target effects. Both Chip-seq and RNA-seq can be used in combination to identify potential off-target effects of CRISPR gene editing. By identifying potential off-target sites using Chip-seq and then examining changes in gene expression using RNA-seq, researchers can gain a better understanding of the potential impact of off-target effects on cellular function.

Figure 2.10- Schematic showing a single vector that houses the entire CRISPROff platform as well as a guide for Map2k1 and a guide for Map2k2. By using a Schwann cell specific promoter, such as P0, the cargo can be selectively expressed in Schwann cells.



Finally, this work nearly completes Aim 1 of the larger project in collaboration with the Battelle group. Once a single vector housing guides 1,7 and CRISPR-Off with a Schwann cell specific promoter (**Figure 2.10**), is created, the Battelle group will test the repression capabilities in their polymer nanoparticles. The Battelle team synthesized, purified, and characterized six polymer nanoparticle (PNP) libraries (each containing 96 different PNP compositions for a total of 576 PNPs. Following their production and physicochemical characterization, 508 of these PNPs proceeded forward to *in vitro* and

in vivo screening processes. *In vitro* screening assays for payload loading efficiency, cytotoxicity, transfection efficiency, and immunogenicity were completed. The data produced by these synthesis, characterization, and assessment efforts have been shared with Battelle's data science team, who have begun applying the data to Battelle's machine learning model. Ultimately, the goal is to develop novel PNPs that house a single vector expressing CRISPR-Off, Guide 7 and Guide 1 behind a Schwann cell specific promoter, with efficient delivery to Schwann cells.

There are some larger questions that have yet to be resolved when considering the efficacy of our proposed *Map2k1*/*Map2k2* repression therapy. First, what efficiency is required in delivery to prevent the progression of neurofibromas and plexiform neurofibromas. Whether all manifestations of neurofibromatosis type 1 will respond to the same approach to treatment is unknown. Although strong evidence supports that lack of neurofibromin activity in Schwann cells leads to increased RAS signaling, other mechanisms might account for the other manifestations of this disorder, such as cognitive deficits. In addition, the optimal time to initiate treatment in individuals is uncertain; whether plexiform neurofibromas should be treated at diagnosis, which would require the use of toxic treatments in young children, or whether treatment should be delayed until they are symptomatic, is a topic of discussion. Whether treatment at a single point in time permanently can prevent the growth of tumors or whether there is a need for continued treatment is unknown, as is the extent of treatment-related toxicity accepted by patients. It is likely that the acceptability of adverse effects of drug treatment is different when one is treating a life-threatening malignancy, compression of

the spinal cord, disfigurement, or moderate cosmetic impairment. Finally, how to accurately identify individuals whose tumors require treatment requires establishment.

Whether any single therapy will be sufficient to treat any one, never mind all, of the manifestations of neurofibromatosis type 1 remains unclear. Given the partial responses seen so far by targeting the RAS pathway, combinations of therapeutic agents might be needed, both to target different disease mechanisms and to avoid the development of treatment resistance. Similarly, it is likely that different molecular subsets of tumors exist, each with their own individual drug sensitivities.

REFERENCES

1. Ly, K. I. & Blakeley, J. O. The Diagnosis and Management of Neurofibromatosis Type 1. *Med. Clin. North Am.* **103**, 1035–1054 (2019).
2. Tamura, R. Current Understanding of Neurofibromatosis Type 1, 2, and Schwannomatosis. *Int. J. Mol. Sci.* **22**, 5850 (2021).
3. Gutmann, D. H., Wood, D. L. & Collins, F. S. Identification of the neurofibromatosis type 1 gene product. *Proc Natl Acad Sci USA* **88**, (1991).
4. Uusitalo, E. Incidence and mortality of neurofibromatosis: a total population study in Finland. *J Invest Dermatol* **135**, (2015).
5. Evans, D. G. Mortality in neurofibromatosis 1: in North West England: an assessment of actuarial survival in a region of the UK since 1989. *Eur J Hum Genet* **19**, (2011).
6. Rasmussen, S. A., Yang, Q. & Friedman, J. M. Mortality in neurofibromatosis 1: an analysis using U.S. death certificates. *Am J Hum Genet* **68**, (2001).
7. Masocco, M. & Orphanet, J. Mortality associated with neurofibromatosis type 1: a study based on Italian death certificates. *Rare Dis* **6**, (2011).
8. Zoller, M., Rembeck, B., Akesson, H. O. & Angervall, L. Life expectancy, mortality and prognostic factors in neurofibromatosis type 1. A twelve-year follow-up of an epidemiological study in Goteborg, Sweden. *Acta Derm Venereol* **75**, (1995).
9. Duong, T. A. Mortality associated with neurofibromatosis 1: a cohort study of 1895 patients in 1980–2006 in France. *Orphanet J Rare Dis* **6**, (2011).
10. Messiaen, L. M. Exhaustive mutation analysis of the NF1 gene allows identification of 95% of mutations and reveals a high frequency of unusual splicing defects. *Hum Mutat* **15**, (2000).
11. Pasmant, E. NF1 microdeletions in neurofibromatosis type 1: from genotype to phenotype. *Hum Mutat* **31**, (2010).
12. Rojnueangnit, K. High incidence of Noonan syndrome features including short stature and pulmonic stenosis in patients carrying NF1 missense mutations affecting p. Arg1809: genotype-phenotype correlation. *Hum Mutat* **36**, (2015).
13. Upadhyaya, M. An absence of cutaneous neurofibromas associated with a 3-bp inframe deletion in exon 17 of the NF1 gene (c.2970–2972 delAAT): evidence of a clinically significant NF1 genotype–phenotype correlation. *Am J Hum Genet* **80**, (2007).
14. Viskochil, D. Deletions and a translocation interrupt a cloned gene at the neurofibromatosis type 1 locus. *Cell* **62**, (1990).
15. Wallace, M. R. Type 1 neurofibromatosis gene: identification of a large transcript disrupted in three NF1 patients. *Science* **249**, (1990).
16. DeClue, J. E. Epidermal growth factor receptor expression in neurofibromatosis type 1-related tumors and NF1 animal models. *J Clin Invest* **105**, (2000).
17. Basu, T. N. Aberrant regulation of ras proteins in malignant tumour cells from type 1 neurofibromatosis patients. *Nature* **356**, (1992).
18. Dasgupta, B., Yi, Y., Chen, D. Y., Weber, J. D. & Gutmann, D. H. Proteomic analysis reveals hyperactivation of the mammalian target of rapamycin pathway in neurofibromatosis 1-associated human and mouse brain tumors. *Cancer Res* **65**, (2005).

19. Johannessen, C. M. The NF1 tumor suppressor critically regulates TSC2 and mTOR. *Proc Natl Acad Sci USA* **102**, (2005).
20. Gutmann, D. H., Blakeley, J. O., Korf, B. R. & Packer, R. J. Optimizing biologically targeted clinical trials for neurofibromatosis. *Expert Opin Investig Drugs* **22**, (2013).
21. Hutchinson, L. Selumetinib MEKing differences in NF1. *Nat. Rev. Clin. Oncol.* **14**, 140–140 (2017).
22. Holt, S. V. *et al.* The MEK1/2 inhibitor, selumetinib (AZD6244; ARRY-142886), enhances anti-tumour efficacy when combined with conventional chemotherapeutic agents in human tumour xenograft models. *Br. J. Cancer* **106**, 858–866 (2012).
23. Dombi, E. Activity of selumetinib in neurofibromatosis type 1-related plexiform neurofibromas. *N Engl J Med* **375**, (2017).
24. Beert, E. Atypical neurofibromas in neurofibromatosis type 1 are premalignant tumors. *Genes. Chromosomes Cancer* **50**, (2011).
25. Carlson-Stevermer, J. *et al.* CRISPRoff enables spatio-temporal control of CRISPR editing. *Nat. Commun.* **11**, 5041 (2020).
26. Nuñez, J. K. *et al.* Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. *Cell* **184**, 2503-2519.e17 (2021).
27. Rivenbark, A. G. *et al.* Epigenetic reprogramming of cancer cells via targeted DNA methylation. *Epigenetics* **7**, 350–360 (2012).
28. Groner, A. C. *et al.* KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS Genet.* **6**, e1000869 (2010).
29. Siddique, A. N. *et al.* Targeted methylation and gene silencing of VEGF-A in human cells by using a designed Dnmt3a-Dnmt3L single-chain fusion protein with increased DNA methylation activity. *J. Mol. Biol.* **425**, 479–491 (2013).
30. Labun, K. *et al.* CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* **47**, W171–W174 (2019).
31. Stringer, B. W. *et al.* A reference collection of patient-derived cell line and xenograft models of proneural, classical and mesenchymal glioblastoma. *Sci. Rep.* **9**, 4902 (2019).

Chapter 3- Developing Epigenetic Therapies for SYNGAP1-related ID

ABSTRACT

SYNGAP1-related ID (ID) is a neurodevelopmental disorder (NDD) characterized by moderate to severe ID starting early in life, as well as recurrent seizures, hyperactivity, and an Autism Spectrum Disorder diagnosis. The disease is caused by *de novo*, loss-of-function mutations in the *SYNGAP1* gene found on chromosome 6. Currently, there are no disease-modifying treatments for SYNGAP1-related ID nor any NDD with a known genetic haploinsufficiency, contributing to a significant biomedical unmet need. We hypothesize that single guide RNA (sgRNA)-targeted CRISPR activation *in vitro* and *in vivo* will demonstrate disease modifying levels of transcribed RNA and protein expression of SYNGAP1 and will result in significant restoration of molecular, functional, neurophysiological and neurohistological phenotypes associated with healthy brain function in patient-derived cells, and SYNGAP1 mutant mouse models with minimal to no toxicity or off-target effects. Our preliminary CRISPR activation guide screen has identified a promising guide that has shown to activate *Syngap1* in a neuronal cell line and the brain of a wild-type mouse.

INTRODUCTION

SYNGAP1-related Intellectual Disability

Abnormal structure or function of neuronal synapses is a frequent underlying mechanism of intellectual disability and autism spectrum disorder.[1] Many genes encode proteins that regulate synaptic structure and function. One known cause of ID is mutations in *SYNGAP1*. Located on chromosome 6p21.3, *SYNGAP1* encodes the cytosolic protein SYNGAP1 (SYNaptic GTPase Activating Protein 1).[2] SYNGAP1 is a complex protein that is an abundant component of the postsynaptic density (PSD) of excitatory glutamatergic neurons, where it is present as a part of the N-methyl-D-aspartate receptor (NMDAR) complex. SYNGAP1 is a negative regulator of small GTPases (e.g., Ras, Rap) and regulates synaptic function by controlling α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor (AMPAR) trafficking to the postsynaptic membrane. Therefore, SYNGAP1 is essential for synaptic development, structure, function, and plasticity.[3]

De novo mutations in *SYNGAP1* are a common cause of non-syndromic ID (i.e., ID without any concomitant physical abnormalities), ASD, and epilepsy.[4,5] All mutations in *SYNGAP1* are loss-of-function and cause haploinsufficiency of SYNGAP1, leading to a defined phenotype, termed SYNGAP1-related ID , characterized by ID and epilepsy. Mutations in *SYNGAP1* account for 2%–8% cases of sporadic ID and >85% of individuals having this mutation also suffer from epilepsy.[4–8] The frequency of *SYNGAP1* mutations is similar to that of fragile X syndrome, the most common inherited cause of ID. *SYNGAP1* mutations have also been implicated in schizophrenia.[6,9]

SYNGAP1-related ID Clinical Features

To date, more than 200 patients with *SYNGAP1* mutations have been reported.[10] Alterations in SYNGAP1 protein arise from *de novo* mutations which lead to haploinsufficiency of SYNGAP1. In approximately 89% of cases, the mutation results in a pathogenic SYNGAP1 variant which is a truncated version of the protein due to frameshift or nonsense mutations. Pathogenic variants also arise due to missense mutations and splice site mutations (about 30% of cases)[7,11–13]. In 11% of cases, 6p21.3 microdeletions are the cause of SYNGAP1-related ID.[13] Balanced translocations have also been reported as a cause of *SYNGAP1* haploinsufficiency.[4,14]

SYNGAP1 mutations lead to developmental and epileptic encephalopathy (DEE) or developmental encephalopathy with or without epilepsy.[15] In developmental encephalopathy, developmental delay arises as a direct consequence of the genetic defect; while in epileptic encephalopathy, the epileptiform activity interferes with development and directly causes cognitive slowing and regression; in DEE, both factors play a role. Among different study cohorts, estimates of *SYNGAP1*-DEE prevalence vary greatly from 35% to 98%.[4,12,13]

More than 80% of individuals with *SYNGAP1* mutations have generalized epilepsy, with focal seizures occurring only in a minority of cases.[10,13] The age of seizure onset ranges from 3 months to 7 years (most often at 2–3 years) and developmental delays typically precede seizure onset.[13] Developmental plateau or regression accompanies seizure onset in many patients, consistent with a diagnosis of DEE. Patients can manifest multiple seizure types, with absence seizures being the

most common type encountered; these include eyelid myoclonia with absences, typical absences, atypical absences, and myoclonic absences. Other seizure types such as myoclonic, atonic, myoclonic-atonic, and tonic-clonic are also observed.[4,13,16] Seizures are quite frequent, ranging from 10 to 100 per day, but brief, each lasting a few seconds.[14] The seizures may respond to anti-epileptic drugs such as valproic acid, lamotrigine, topiramate, or clobazam, but epilepsy is pharmacoresistant in approximately 50% cases.[4,7,10]

All SYNGAP1-related ID patients suffer from ID, which is typically moderate to severe. [4,7,13] Global developmental delays become manifest in the first or second year of life.[4] Motor delays include delayed age of sitting unaided (mean 12 months) and walking unaided (onset between 2 and 3 years).[7,12] Language is severely impaired with delays in both expressive and receptive speech development.[12,13] One third of individuals >5 years old remain non-verbal, and in verbal patients language ranges from single words to brief sentences.[10] It has been observed that milder phenotypes occur more often in patients with mutations in exons 1–4 compared to more severe phenotypes when mutations involve exons 8–15.[13] About half of patients with *SYNGAP1* mutations are autistic.[17] No association has been found between ASD and the severity of ID, or between ASD and the location of the mutation on the *SYNGAP1* gene. Up to three-quarters of SYNGAP1-related ID patients suffer from severe behavioral problems including hyperexcitability, aggression, oppositional behavior, tantrums, and self-injury. Sleep difficulties are reported in as many as 60% of patients, with problems in both initiating and maintaining sleep; these are managed with

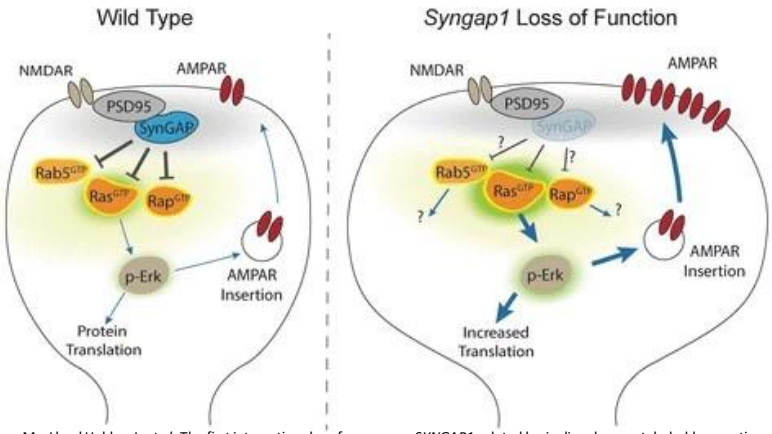
melatonin, clonidine, or trazodone. Eating difficulties, such as oral aversion and oral hypersensitivity, are also common. [4,12,13]

Function of the SYNGAP1 Protein

The SYNGAP1 protein is a Ras GTPase activating protein that interacts with the proteins in the PSD of excitatory glutamatergic synapses.[18] The PSD is a protein complex that lines postsynaptic membranes and consists of several proteins, including scaffold proteins (PSD-95), Ca^{2+}/calmodulin dependent protein kinase (CaMKII), and synaptic GTPases like proteins of the Ras family and their regulators, including SYNGAP1. The PSD has an important role in control of receptor clustering and function, thereby modulating neurotransmission. SYNGAP1 is present in the PSD as part of a complex with NMDAR subunits and PSD-95.[3,18]

SYNGAP1 is a GTPase activating protein (GAP) that negatively regulates the activity of the small G proteins, such as Ras and Rap.[18] SYNGAP1 is an important component of NMDAR-dependent activation of Ras-dependent signalling pathways, exemplified by SYNGAP1 regulation of ERK/MAPK signalling.[19] *In-vitro* studies have demonstrated that SYNGAP1 negatively regulates the activity of Rap even more potently than Ras, but as yet the functional role of the RapGAP action has not been studied.[20] Ras- and Rap-GAP activity is subject to regulation by 3 different kinases: CaMKII, cyclin-dependent kinase-like 5 (CDK5), and Polo-like kinase 2 (PLK2). CaMKII activation causes rapid phosphorylation of SYNGAP1 and increases both its Ras-GAP and Rap-GAP activity.[21,22] Ras and Rap are independent regulators of AMPAR trafficking to the postsynaptic membrane, with Ras increasing synaptic delivery of

Figure 3.1 - Schematic of signaling pathways regulated by the *Syngap1* gene. SYNGAP protein inhibits the activation of various small GTPases. In dendritic spines, SYNGAP suppresses Ras/ Erk activity and limits growth-related processes including protein translation and AMPAR receptor exocytosis. Reduced SYNGAP protein levels causes elevated basal Ras/Erk signaling. This results in increased AMPAR surface incorporation thought to contribute to excessive excitation in neural circuits.



Weldon, M., Kilinc, M., Lloyd Holder, J. *et al.* The first international conference on *SYNGAP1*-related brain disorders: a stakeholder meeting of families, researchers, clinicians, and regulators. *J Neurodevelop Disord* **10**, 6 (2018). https://doi.org/10.1186/s11689-018-9225-1

AMPARs and Rap mediating removal of synaptic AMPARs.[23] This differential activity of SYNGAP1 on these two small GTPases with opposite functions confirms its role as an important regulator of post-synaptic membrane composition and synaptic plasticity.

Animal models with heterozygous *Syngap1* knockout (*Syngap1* $^{+/-}$ genotype) have helped to elucidate the role of SYNGAP1 in neuronal development and synaptic pathology due to *SYNGAP1* haploinsufficiency. *Syngap1*$^{+/-}$ mice show phenotypes similar to those observed in SYNGAP1-related ID patients, including deficits in behavior (hyperactivity, anxiety, and hyper-arousal), social isolation, lack of social memory cognition and learning, impaired working and reference memory, slow spatial learning, and widespread cortical excitability as manifest by frequent epileptiform discharges.[9,24,25] Zebrafish models with reduced *SYNGAP1* expression show spontaneous seizure-like behavior.[26] Mice with homozygous *Syngap1* knock-out (*Syngap1*$^{-/-}$ genotype) have significantly smaller brains, a failure-to-thrive phenotype,

impaired motor skills, and die within 1 or 2 days of birth. These observations indicate that SYNGAP1 is crucial for early brain development and neuron proliferation.

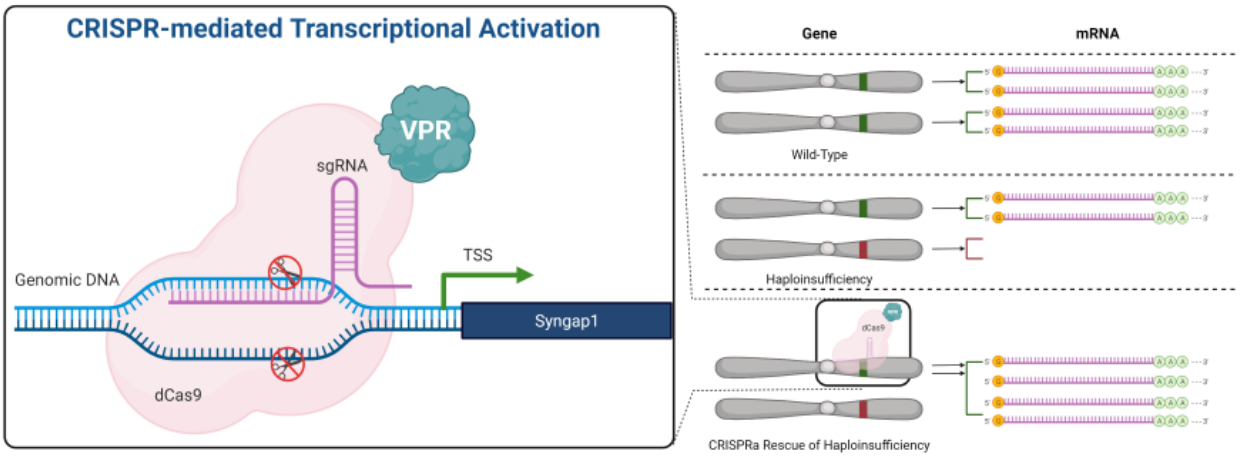
CRISPR activation

To rescue the *SYNGAP1* haploinsufficiency, we employed a CRISPR activation approach using a modified Cas9 protein conjugated to the transcriptional activators VPR. CRISPR activation (CRISPRa) is a epigenome editing technique that uses the CRISPR-Cas system to selectively activate genes within a cell. This is accomplished by fusing a transcriptional activator to a catalytically dead Cas9 (dCas9) protein, which can bind to specific DNA sequences and recruit transcriptional machinery to activate gene expression.[27] CRISPRa has numerous potential applications in both basic research and gene therapy, allowing for targeted activation of genes involved in disease pathways or other cellular processes.[28] However, the large size of most Cas endonucleases (950-1400 amino acids) can restrict applications. The recent discovery of miniature Cas endonucleases (400-800 amino acids), such as Cas12f, provides the potential to overcome this limitation.[29] Therefore, we designed guides using both Cas9 and Cas12f, and empirically tested them at the *Syngap1* promoter region to determine the strongest potential activator.

Transactivation domains in an epigenome editing toolbox (a set of components for expression of sgRNA, dCas9 and effector domains) need to be modular, in the sense that they retain their function in fusions with different DNA-binding moieties of such gene activation systems. Several potent activators have met the criteria and are commonly used for artificial transcriptional regulation. The first and the simplest

activation domain is the virion protein 16 (VP16) from the herpes simplex virus type 1 transactivation domain, which has been successfully used for transcriptional activation in mammalian cells.[30–32] The VP64 domain is a tetramer of the minimal activator VP16 (four VP16 linked with Gly-Ser linkers), which was shown to be a stronger activator than a single VP16 and has been routinely used for artificial transcriptional activation.[33] Another commonly used transcriptional activation domain is the p65, the larger subunit of the NF-kappa B transcription factor, known to induce strong gene transactivation in mammalian cells.[34] The Rta domain is another transcriptional activator in this portfolio, encoded by Kaposi's sarcoma associated herpesvirus/human herpesvirus 8 (KSHV/HHV-8). It activates the expression of viral genes in the lytic cycle.[35,36] Finally, the VPR transactivator is a chimeric unit composed of the activation domains of VP64, p65 and Rta. It was recently developed for more robust gene activation by using several activation units, exhibiting a strong synergistic effect.[37] For this study, we utilized the VPR transactivator.

Figure 3.2 - CRISPR Activation can be used to rescue the Syngap1 haploinsufficiency by using a dCas9 protein fused to VPR with guide RNAs binding upstream of the transcriptional start site of the endogenous *Syngap1* gene. (Created with BioRender.com)



MATERIALS AND METHODS

CRISRP/Cas9 Guide Design

CRISPR guide RNAs were designed for both a Cas9 system as well as a Cas12f system. The bioinformatics tool ChopChop was utilized to identify all the potential guides 400bp upstream of the Syngap1 promoter.[38] **Figures 3.3 and 3.4** show a table identifying the guide sequence, distance from the transcriptional start site, as well as a schematic of all the potential guides in the Syngap1 region. 8-10 guides were chosen for each Cas system, which are identified in blue on the genomic schematic. Although the ChopChop algorithm ranks guides based off potential off-target affects and predicted activity levels, guides were manually chosen to include guides in different regions of the promoter to account for any variation in activity due to steric blocking by other transcription factors. Cas9 guides were chosen based off a 3' NGG PAM sequence, while Cas12f guides utilized a 5' TTTR PAM sequence. Due to the more restrictive nature of the Cas12f PAM sequence, only 8 guides were possible in the promoter region of the *Syngap1* gene, so all these guides were chosen for the subsequent guide screen.

Following identification of guide sequences *in silico*, guides was cloned into a vector expressing the sgRNA. Oligonucleotides were ordered from IDT containing the 20bp spacer designed from ChopChop, with overhangs enabling sticky-end ligation into the vector lenti-sgRNA puro (Addgene #10499) digested with *BsmBI*.[39] The lenti-sgRNA puro vector backbone was digested with *BsmBI*, gel purified using QIAquick Gel Extraction Kit. Oligonucleotides were phosphorylated and annealed in T4 Ligation Buffer using T4 PNK (NEB). These annealed, and phosphorylated oligonucleotides were then

diluted at 1:200 before being ligated into the *BsmBI* digested lenti-sgRNA puro vector using T4 DNA Ligase (NEB). Successful insertion of the oligonucleotides containing the sgRNA sequence was confirmed via Sanger sequencing.

Figure3.3 - Syngap1 Cas9 Guide Design for CRISPR Activation

Rank	Name	Guide w/ 3' NGG PAM	Distance to TSS (bp)
1	Syngap1_1_mus	CAGCTTCTACGGGACTTCCCTGG	125
2	Syngap1_2_mus	GAGACCCAATTTGTTGAGTATGG	215
3	Syngap1_3_mus	CACCTCCATACTCAACAAATTGG	210
4	Syngap1_4_mus	ACCCGGAAAAGGAGAGAACGAGG	83
5	Syngap1_5_mus	TTTTCCGGGTCTCCCTGGGTAGG	70
6	Syngap1_6_mus	GGTCTTGGGTCAGCTTCTACGGG	135
7	Syngap1_7_mus	TTTCCGGGTCTCCCTGGGTAGGG	69
8	Syngap1_8_mus	TTCCGGGTCTCCCTGGGTAGGGG	68
9	Syngap1_9_mus	GGGGTCTGTCTGAGGGTCTTGGG	149
10	Syngap1_10_mus	ACCCAATTTGTTGAGTATGGAGG	212

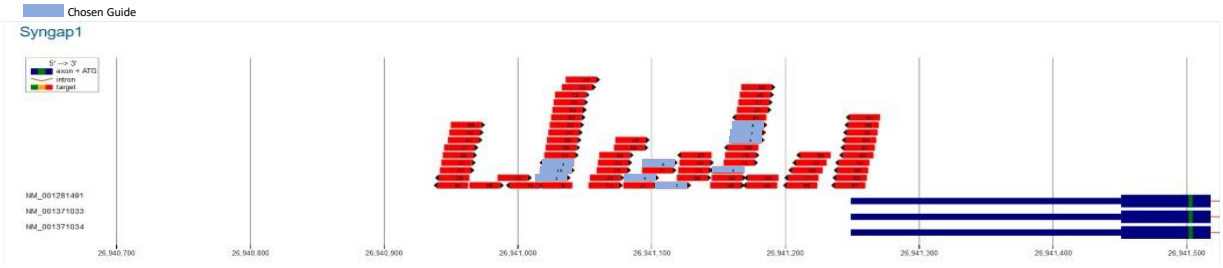
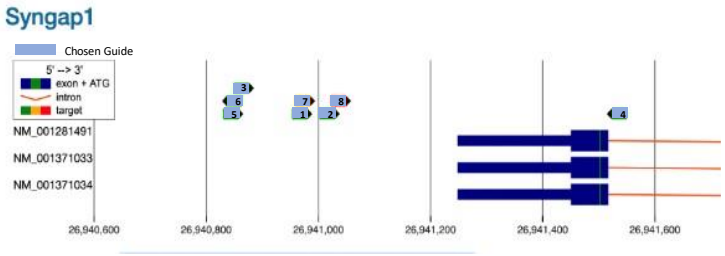


Figure 3.4 - Syngap1 Cas12f Guide Design for CRISPR Activation

Rank	Name	Guide w/ 5' TTTR PAM	Distance to TSS (bp)
1	Syngap1_11_mus	TTTAATTTAGGGGGTGGGGTCTGAGCT	271
2	Syngap1_12_mus	TTTGGATCTGTGAGACCCAATTTGTTG	222
3	Syngap1_13_mus	TTTGGGTCTCCCAAACTACAAGTCCAG	375
4	Syngap1_14_mus	TTTGGCCCTCCCCTCCCCTTACCACCA	-272
5	Syngap1_15_mus	TTTGGGTCTAGGTCTCTGTTTTGGGTC	394
6	Syngap1_16_mus	TTTGGGAGACCCAAAACAGAGACCTAG	387
7	Syngap1_17_mus	TTTAGGGGGTGGGGTCTGAGCTCAGAG	266
8	Syngap1_18_mus	TTTGTTGAGTATGGAGGTGGGGGGGGG	202



Split dCas9-VPR

A split dCas9-VPR system containing two plasmids was obtained from the Fink Lab at UC Davis. The N-terminal region of the dCas9 protein is packaged in one vector, and the C-terminal region and the VPR activator is packaged in another. A split dCas9 system using inteins involves dividing the dCas9 protein into two separate fragments, which are then fused to two parts of an intein. Inteins are protein sequences that are capable of self-splicing and can join two separate protein fragments together. When the two dCas9-intein fragments are brought together, the inteins catalyze their fusion, creating a functional dCas9 protein. A split system is useful for downstream AAV applications, where a 4.7kb packaging limit requires the use of two separate vectors to deliver the total CRISPR activation cargo. A schematic showing the split dCas9 system for AAV packaging is shown in **Figure 3.4**.

Figure 3.4 - A dual AAV system that can effectively deliver the dCas9 -effector fusion domain and the targeting sgRNA by splitting the dCas9 and effectors across two AAV vectors. This dual system has demonstrated efficient transduction and expression of full - length dCas9 protein that is biologically active

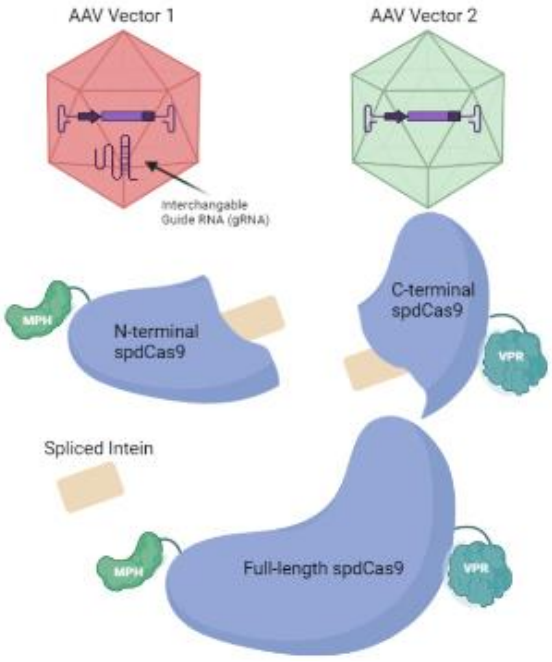


Figure courtesy of Julian Halmai from the Fink lab at UC Davis

Cell Transfection

To screen the guides for the strongest potential activator for *Syngap1*, Neuro2A cells were transfected using Lipofectamine 3000, with subsequent RT-qPCR for relative gene mRNA levels. Neuro2A cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) media supplemented with 10% fetal bovine serum (FBS) and Penicillin/Streptomycin at 37C under 5% CO2. A total of 1 million cells at 60-70% confluency was transfected using Lipofectamine 3000 following the manufacturer's instructions. Transfections were performed in 12-well plates using 500ng of the sgRNA expressing vector, and 500ng of each plasmid of the split dCas9-VPR system obtained from the Fink Lab. Cells were harvested for RNA extraction at day 3 post lipofection using a Qiagen's RNeasy kit.

The dCas9 screen was conducted without FACS sorting for successfully transfected cells. The dCas12f screen was conducted with FACS sorting, with one promising dCas9 guide from the initial screen as well. The results of both screens are shown in **Figure 3.5**.

Quantitative PCR (qPCR)

RNA was extracted from harvested cells using Qiagen RNAeasy Micro Kit (Qiagen). RNA was reverse transcribed using the AB high-capacity cDNA synthesis kit (ThermoFisher) according to the manufacturer's instructions. Primers for quantitative PCR (qPCR) were designed using Primer3. RT-qPCR was performed in triplicate using PowerUp SYBR Green Master Mix (ThermoFisher) with the CFX384 Real-Time System C1000 Touch system (Bio-Rad). Gene expression analysis was performed with

Syngap1 gene specific primers (**Table 3.1**) using three biological replicates and three technical replicates. Relative target gene expression was calculated as the difference between the target gene and the GAPDH reference gene (dCq = Cq[target] − Cq[GAPDH]). Gene expression results are indicated as fold change to a reference sample (dCas9 combination without targeting gRNA), using the ddCq method.

Table 3.1 - qPCR Primers used to determine relative expression of *Syngap1*

Primer Name	Primer Sequence (5'-3')
Syngap1_qPCR_1F	CGAAGTGCTGACCATGAC
Syngap1_qPCR_1R	CGCCTGTTGTCCTTGTTG
GAPDH_1F	CATCACTGCCACCCAGAAGACTG
GAPDH_1R	ATGCCAGTGAGCTTCCCGTTCAG

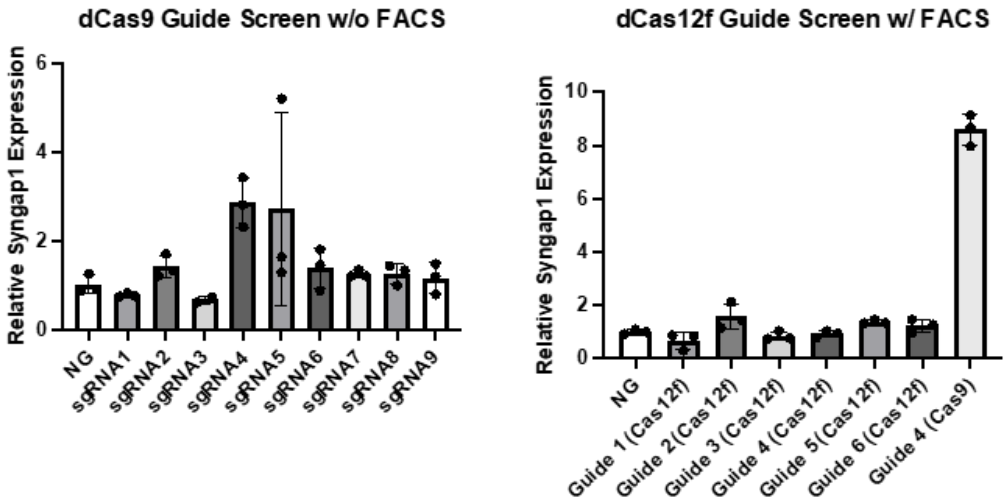
AAV Treatment in Wild-Type Mice

As shown in **Figure 3.4**, the sgRNA expressing plasmid and split dCas9 system were packaged into an AAV9 vector for transduction into wild-type mouse. These AAV vectors were injected into the brain of wildtype mice (n=3). Non-effector treated mice (NT) were used as a negative control (n=3). Nine days following administration, brains were isolated for RT-qPCR. This work was done by Julian Halmai from the Fink lab.

DISCUSSSION

In the efforts towards creating an epigenetic therapy for the *Syngap1* haploinsufficiency, this project was successful in identifying a promising single guide RNA that has shown to increase the expression of *Syngap1 in vitro* in Neuro2A cells (**Figure 3.5**). Without FACS sorting, the strongest activator was dCas9 was guide 4, with a 2.8 increased relative *Syngap1* expression. In the second screen, which was FACS sorted for successfully transfected cells, guide 4 showed about an 8 fold relative increase in *Syngap1* expression. Comparatively, none of the dCas12f showed promise as a strong activator, so dCas9 Guide 4 was chosen for subsequent *in vivo* experiments.

Figure 3.5 - *Syngap1* Guide Screen for CRISPR Activation dCas9 guides were screened without FACS sorting for successfully transfected cells. (Left) dCas12f guides were screened with FACS sorting, along with dCas9 Guide 4. (Right)

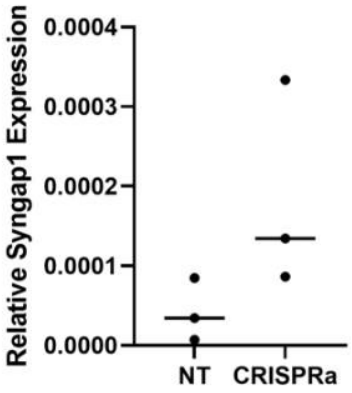


One serious limitation of this initial screen is the cell type used to screen the guides. Neuro2A cells, although an easy to work with cell line, is not a clinically relevant cell line for *Syngap1*. Ideally, showing increased *Syngap1* activation in a *Syngap1*

haploinsufficient cell line would be the logical next step in the project. Additionally, another limitation of the methodology used here is that the qPCR analysis was done on batch cells. Therefore, we are unsure if the measured increased *Syngap1* expression is due to significant activation in a limited group of cells, or moderate activation in a large number of cells. With an *in situ* hybridization assay, such as RNAscope, we can determine if CRISPR activation with our split dCas9-VPR system is due to a small effect in many cells or a large effect in a few cells, and measure biodistribution on a per-cell basis, rather than on bulk tissue. Additionally, we can measure rescue of SYNGAP1 protein levels using Western blot.

After packaging dCas9 guide 4 into an AAV vector, and delivering it to wild-type mice with the split dCas9-VPR AAV vector, we were able to see a trend towards increased *Syngap1* expression. (**Figure 3.6**). Mice treated with AAV-CRISPRa showed a trend toward increased transcript level when compared to Non-treated (NT) controls, however, the difference wasn't statistically significant. Because wild-type mice are not *Syngap1* deficient, it may be more difficult to measure any significant activation. The

Figure3.6 - Relative Syngap1 Expression in WT Mice after treatment with AAV packaged dCas9 + Guide 4



next step would be to deliver this therapy to a *Syngap1* haploinsufficient mouse, or increase the sample size of this initial *in vivo* study to determine if this increased trend is statistically significant.

Another follow-up experiment is to characterize the temporal extent of the activation. With our preliminary guide screen, cells were harvested for qPCR analysis at day 3 in cell lines, and day 9 from mouse tissue. By extending this timeline, and sampling cells/harvesting mice at various timepoints, we can measure if the CRISPR activating effect persists. CRISPR activation has been found to work long-term in other studies for different loci, but it still has to be determined if such an effect is possible at the Syngap1 promoter.[27]

Because neurons are post-mitotic cells, and have lost the ability to divide and replicate themselves in the brain, a big determinant for the persistence of any CRISPR activation based therapy is the duration of AAV expression in these cells. The duration of AAV expression in the brain can depend on several factors, including the serotype of the AAV, the specific promoter used to drive gene expression, and the target cell type.[40] In general, AAV is known to be a relatively stable vector, and it can maintain gene expression in the brain for several months to several years after injection.[41] However, the exact duration of AAV expression can vary widely depending on the specific experimental conditions. Some studies have reported stable AAV expression in the brain for up to five years or more, while others have observed a gradual decline in expression over a period of several months.[42] Overall, the duration of AAV expression in the brain is an important consideration for designing and interpreting experiments using this vector. Researchers should carefully select the appropriate AAV serotype and

promoter for their specific experimental needs and monitor AAV expression over time to ensure that it remains stable and consistent. In our initial study, we used AAV9. One study that specifically looked at the persistence of AAV9 in the brain found that the vector could drive long-term expression of a fluorescent protein for up to 6 months after injection in mice. However, the study also found that there was some variability in the persistence of AAV9 expression across different brain regions, with higher levels of expression in the cortex compared to the hippocampus.[42] Newer AAV capsids have since been developed, such as AAV-B10, which have the ability to efficiently transduce human cell models, and the brains of mice and nonhuman primates following intravenous tail vein injection.[43]

The results outlined in this project is preliminary work in a large collaboration between the Fink, Silverman, and Segal labs. Future work in the pipeline includes the testing this therapy in a *Syngap1* deficient mouse, which will be done by the Silverman lab. The experimental design for future experiment include injecting (IV tail-vein) 5-week-old mice (n= 15-20 aged-matched male and female mice per group) with 1×10^{11} vg/kg of the mouse-targeted AAV-B10/CRISPRa. 21 days post-injection, rescue of motor skills and motor coordination will be assessed by rotarod for balance coordination, and the Digigait for temporal and spatial parameters of gait. To examine rescue of cognitive impairments relevant to SYNGAP1, treated mice will undergo a cognitive battery including a Y-maze assay, novel object recognition and the touchscreen assay. To examine rescue of behavioral seizures and EEG spiking using continuous non-tethered recordings, treated mice will be implanted with a wireless device designed to collect EEG data in free-moving animals in their home cage. We will

test for behavioral seizure susceptibility following convulsant administration. To examine sleep rescue, quality by sleep stages and quantity by time spent will be recorded using continuous non-tethered recording. EEG acquired from treated mice will also be analyzed for sleep including frequency and duration of all sleep stages and sleep spindle production. At the conclusion of this functional assessment, brain tissue will be isolated for either molecular (n= 3-4 male/female mice per group) or histological (n = 3-4 male/female mice per group) assessment to determine molecular mechanism of action for functional reactivation of SYNGAP1.

The level of neurons that increase expression of *Syngap1* will be dependent on the regional distribution and transduction efficiency of the AAV-CRISPRa. It is possible that due to incomplete transduction efficiency in the brain we will have to perform additional immunofluorescent co-localization studies to look at neurons that were transduced with AAV prior to assessing Syngap1 transcript with RNAscope. In this analysis we will be able to only look at “treated” cells that have received AAV-CRISPRa to assess optimal gene regulation via RNAscope.

REFERENCES

1. Zoghbi, H. Y. & Bear, M. F. Synaptic Dysfunction in Neurodevelopmental Disorders Associated with Autism and Intellectual Disabilities. *Cold Spring Harb. Perspect. Biol.* **4**, a009886 (2012).
2. Proteomic analysis of NMDA receptor–adhesion protein signaling complexes | Nature Neuroscience. https://www.nature.com/articles/nn0700_661.
3. Sheng, M. & Kim, E. The Postsynaptic Organization of Synapses. *Cold Spring Harb. Perspect. Biol.* **3**, a005678 (2011).
4. Berryer, M. H. *et al.* Mutations in SYNGAP1 Cause Intellectual Disability, Autism, and a Specific Form of Epilepsy by Inducing Haploinsufficiency. *Hum. Mutat.* **34**, 385–394 (2013).
5. Hamdan, F. F. *et al.* Mutations in SYNGAP1 in Autosomal Nonsyndromic Mental Retardation. *N. Engl. J. Med.* **360**, 599–605 (2009).
6. Fitzgerald, T. W. *et al.* Large-scale discovery of novel genetic causes of developmental disorders. *Nature* **519**, 223–228 (2015).
7. Hamdan, F. F. *et al.* De Novo SYNGAP1 Mutations in Nonsyndromic Intellectual Disability and Autism. *Biol. Psychiatry* **69**, 898–901 (2011).
8. Rauch, A. *et al.* Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *The Lancet* **380**, 1674–1682 (2012).
9. Guo, X. *et al.* Reduced Expression of the NMDA Receptor-Interacting Protein SynGAP Causes Behavioral Abnormalities that Model Symptoms of Schizophrenia. *Neuropsychopharmacology* **34**, 1659–1672 (2009).
10. Weldon, M., Kilinc, M., Lloyd Holder, J. & Rumbaugh, G. The first international conference on SYNGAP1-related brain disorders: a stakeholder meeting of families, researchers, clinicians, and regulators. *J. Neurodev. Disord.* **10**, 6 (2018).
11. O'Roak, B. J. *et al.* Recurrent de novo mutations implicate novel genes underlying simplex autism risk. *Nat. Commun.* **5**, 5595 (2014).
12. Parker, M. J. *et al.* De novo, heterozygous, loss-of-function mutations in SYNGAP1 cause a syndromic form of intellectual disability. *Am. J. Med. Genet. A.* **167**, 2231–2237 (2015).
13. Vlaskamp, D. R. M. *et al.* SYNGAP1 encephalopathy: A distinctive generalized developmental and epileptic encephalopathy. *Neurology* **92**, e96–e107 (2019).
14. Writzl, K. & Knegt, A. C. 6p21.3 microdeletion involving the SYNGAP1 gene in a patient with intellectual disability, seizures, and severe speech impairment. *Am. J. Med. Genet. A.* **161**, 1682–1685 (2013).
15. Scheffer, I. E. *et al.* ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia* **58**, 512–521 (2017).
16. von Stülpnagel, C. *et al.* Chewing induced reflex seizures ("eating epilepsy") and eye closure sensitivity as a common feature in pediatric patients with SYNGAP1 mutations: Review of literature and report of 8 cases. *Seizure* **65**, 131–137 (2019).
17. Wright, D., Kenny, A., Eley, S., McKechanie, A. G. & Stanfield, A. C. Clinical and behavioural features of SYNGAP1-related ID : a parent and caregiver description. *J. Neurodev. Disord.* **14**, 34 (2022).

18. Kim, J. H., Liao, D., Lau, L.-F. & Huganir, R. L. SynGAP: a Synaptic RasGAP that Associates with the PSD-95/SAP90 Protein Family. *Neuron* **20**, 683–691 (1998).
19. Komiyama, N. H. *et al.* SynGAP Regulates ERK/MAPK Signaling, Synaptic Plasticity, and Learning in the Complex with Postsynaptic Density 95 and NMDA Receptor. *J. Neurosci.* **22**, 9721–9732 (2002).
20. Krapivinsky, G., Medina, I., Krapivinsky, L., Gapon, S. & Clapham, D. E. SynGAP-MUPP1-CaMKII Synaptic Complexes Regulate p38 MAP Kinase Activity and NMDA Receptor- Dependent Synaptic AMPA Receptor Potentiation. *Neuron* **43**, 563–574 (2004).
21. Oh, J. S., Manzerra, P. & Kennedy, M. B. Regulation of the Neuron-specific Ras GTPase-activating Protein, synGAP, by Ca2+/Calmodulin-dependent Protein Kinase II*. *J. Biol. Chem.* **279**, 17980–17988 (2004).
22. Walkup, W. G. *et al.* Phosphorylation of Synaptic GTPase-activating Protein (synGAP) by Ca2+/Calmodulin-dependent Protein Kinase II (CaMKII) and Cyclin-dependent Kinase 5 (CDK5) Alters the Ratio of Its GAP Activity toward Ras and Rap GTPases*. *J. Biol. Chem.* **290**, 4908–4927 (2015).
23. Zhu, J. J., Qin, Y., Zhao, M., Van Aelst, L. & Malinow, R. Ras and Rap Control AMPA Receptor Trafficking during Synaptic Plasticity. *Cell* **110**, 443–455 (2002).
24. Kilinc, M. *et al.* Species-conserved SYNGAP1 phenotypes associated with neurodevelopmental disorders. *Mol. Cell. Neurosci.* **91**, 140–150 (2018).
25. Ozkan, E. D. *et al.* Reduced Cognition in Syngap1 Mutants Is Caused by Isolated Damage within Developing Forebrain Excitatory Neurons. *Neuron* **82**, 1317–1333 (2014).
26. Kozol, R. A. *et al.* Two knockdown models of the autism genes SYNGAP1 and SHANK3 in zebrafish produce similar behavioral phenotypes associated with embryonic disruptions of brain morphogenesis. *Hum. Mol. Genet.* **24**, 4006–4023 (2015).
27. Jiang, L. *et al.* CRISPR activation of endogenous genes reprograms fibroblasts into cardiovascular progenitor cells for myocardial infarction therapy. *Mol. Ther.* **30**, 54–74 (2022).
28. Maeder, M. L. *et al.* CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* **10**, 977–979 (2013).
29. Nguyen, G. T., Dhingra, Y. & Sashital, D. G. Miniature CRISPR-Cas12 endonucleases - Programmed DNA targeting in a smaller package. *Curr. Opin. Struct. Biol.* **77**, 102466 (2022).
30. Critical Structural Elements of the VP16 Transcriptional Activation Domain | Science. https://www.science.org/doi/10.1126/science.1846049.
31. Sadowski: GAL4-VP16 is an unusually potent transcriptiona... - Google Scholar. https://scholar.google.com/scholar_lookup?title=GAL4-VP16%20is%20an%20unusually%20potent%20transcriptional%20activator&publication_year=1988&author=I.%20Sadowski&author=J.%20Ma&author=S.%20Triezenberg&author=M.%20Ptashne.
32. Seipel, K., Georgiev, O. & Schaffner, W. Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. *EMBO J.* **11**, 4961–4968 (1992).

33. Beerli, R. R., Segal, D. J., Dreier, B. & Barbas III, C. F. Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14628–14633 (1998).
34. Schmitz, M. l. & Baeuerle, P. a. The p65 subunit is responsible for the strong transcription activating potential of NF-kappa B. *EMBO J.* **10**, 3805–3817 (1991).
35. Kaposi's Sarcoma-Associated Herpesvirus Open Reading Frame 50/Rta Protein Activates the Entire Viral Lytic Cycle in the HH-B2 Primary Effusion Lymphoma Cell Line | Journal of Virology. https://journals.asm.org/doi/full/10.1128/JVI.74.13.6207-6212.2000.
36. Sun, R. *et al.* A viral gene that activates lytic cycle expression of Kaposi's sarcoma-associated herpesvirus. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 10866–10871 (1998).
37. Chavez, A. *et al.* Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* **12**, 326–328 (2015).
38. Labun, K. *et al.* CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* **47**, W171–W174 (2019).
39. Stringer, B. W. *et al.* A reference collection of patient-derived cell line and xenograft models of proneural, classical and mesenchymal glioblastoma. *Sci. Rep.* **9**, 4902 (2019).
40. Naso, M. F., Tomkowicz, B., Perry, W. L. & Strohl, W. R. Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. *BioDrugs Clin. Immunother. Biopharm. Gene Ther.* **31**, 317–334 (2017).
41. Fischell, J. M. & Fishman, P. S. A Multifaceted Approach to Optimizing AAV Delivery to the Brain for the Treatment of Neurodegenerative Diseases. *Front. Neurosci.* **15**, 747726 (2021).
42. Hanlon, K. S. *et al.* Selection of an Efficient AAV Vector for Robust CNS Transgene Expression. *Mol. Ther. - Methods Clin. Dev.* **15**, 320–332 (2019).
43. Goertsen, D. *et al.* AAV capsid variants with brain-wide transgene expression and decreased liver targeting after intravenous delivery in mouse and marmoset. *Nat. Neurosci.* **25**, 106–115 (2022).

Chapter 4- Rescuing the 22q11.2 Deletion Syndrome (22q DS) Haploinsufficiency

ABSTRACT

22q11.2 deletion syndrome (22q DS) is a disorder characterized by a spontaneous chromosomal microdeletion on the order of three million base pairs, containing approximately forty protein coding genes. The disorder occurs in about 1/4000 live births and is considered the second most common cause of intellectual disability in the world. Patients with the deletion exhibit wide ranging symptoms, including neurological and behavioral symptoms such as attention deficit hyperactive disorder, anxiety, and schizophrenia. Here, we describe a strategy to upregulate a subset of these forty protein coding genes to rescue the haploinsufficiency using a dead Cas12a (dCas12a) protein fused to transcriptional activators. We will be using dCas12a instead of the more widely used dCas9 because Cas12a is more suited for multiplex activation of multiple genes due to being a smaller protein with a smaller overall guide length. To upregulate multiple genes, we will employ a multiplex approach designing a single plasmid that contains up to 25 guide RNAs. The prevailing theory regarding the gene-phenotype relationship with 22q is that there is a wide variety of complex interactions with many shared molecular targets between the genes involved in the deletion. A multiplexed dCas12a activation approach could therefore be both a critical tool for mapping the genetic causes of the 22Q DS phenotypes, as well as for developing multi-gene therapies.

INTRODUCTION

22q DS is a chromosomal microdeletion disorder, leading to the deletion of 3.5 million base pairs on one of the 22nd chromosomes.[1] The deletion occurs during meiosis, due to the region being flanked by two homologous low-copy repeat regions. 22q DS is considered a haploinsufficiency, because only one of the 22nd chromosomes is affected, leading to about half as much mRNA produced for about forty protein coding genes in that region.[1] Unsurprisingly, the deletion of 40 protein coding genes can lead to a wide range of symptoms, as 22q DS affects many organ systems in patients. When patients are born, the most worrying symptoms are severe cardiac deficiencies, and craniofacial abnormalities. Thankfully, surgical advancements have addressed these symptoms. However, between the ages of 3-5, patients experience severe neurological and cognitive symptoms, such as attention deficit disorder, anxiety, autism spectrum disorders, and even later in life, schizophrenia. Unfortunately, these neurological and cognitive symptoms are mostly unaddressed in the clinic.[1–3]

Our main hypothesis is that the haploinsufficiency caused by the 22q11.2 deletion syndrome can be rescued through upregulation of the genes on the non-deleted chromosome. In order to test this hypothesis, we will use DNAse-dead Cas12a (dCas12a) for transcriptional upregulation of a subset of deleted genes associated with the neurological and cognitive phenotype. dCas12a is a powerful tool that has the potential to transcriptionally upregulate specific endogenous genes in the genome.[4,5] The basic premise of using a DNAse-dead Cas protein is that the nuclease region of the protein is mutated to prevent it from introducing double stranded breaks in the genome. Instead, the dCas is conjugated to transcriptional activators, such as VPR, consisting of

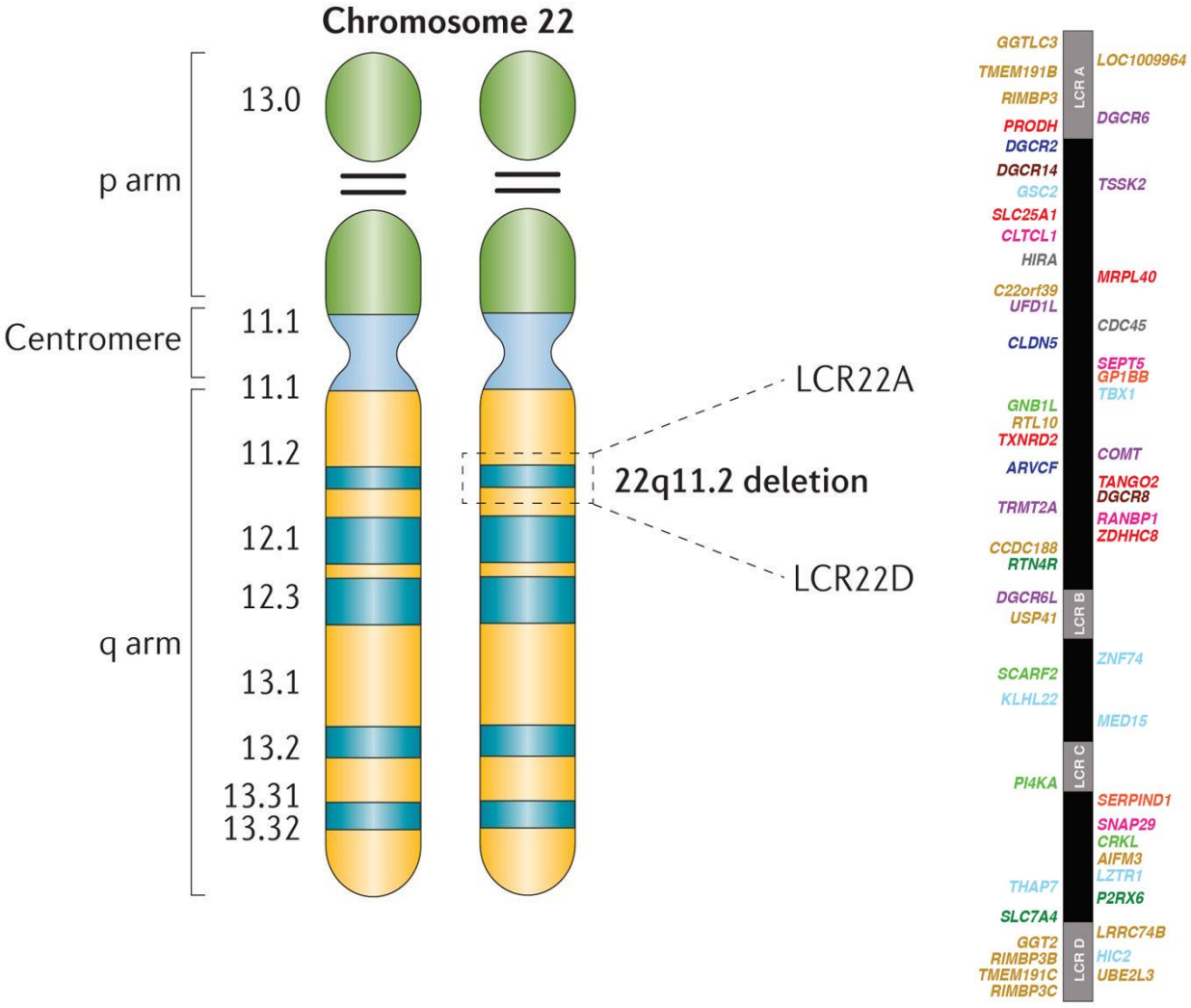
VP64, Rta, and p65, to recruit these factors to the specific regions of the genome. The ability of dCas to complex with guide crRNAs is still intact. Therefore, by designing crRNA guides that bind to promoter or enhancer elements of the target gene, the dCas-VPR system can recruit transcriptional regulatory elements to virtually any gene in the genome, driving overexpression of endogenous genes. The scientific premise for our study is based on a previous demonstration that dCas9-VPR was able to rescue the haploinsufficiency for the *Sim1* model of obesity.[6] That study designed sgRNAs to target the promoter and enhancer of the *Sim1* gene, increasing mRNA expression about 4-fold in vitro in mouse Neuro2A (N2A) cell lines, and reverting *Sim1* haploinsufficient mice to wild types levels of mRNA expression *in vivo* using AAV-delivered dCas9 and sgRNA. Importantly, the study was able to phenotypically reverse the obesity phenotype through AAV delivered dCas9-guide RNAs. The study provides strong rationale for our approach. However, that study only required upregulation of a single gene. 22q DS is a multigene disorder that requires multiple genes to be upregulated to rescue the haploinsufficiency. Therefore, we decided to pursue dCas12a instead of dCas9. Unlike Cas9, Cas12a is a smaller protein with a smaller guide length, important for considering AAV as a means of delivery which has a packaging limit of about 4.7 kilobases.[7] Additionally, Cas12a functions as both an RNase and a DNase able to process and mature its own crRNA. CRISPR arrays for Cas12a are transcribed as long RNA transcripts with 30-nt spacer regions (responsible for DNA recognition) and 36-nt direct repeats (DRs). With this long array, and the nuclease's ability to process this array, the Cas12a system has been shown to target up to 25 genes through efficient multiplexing.[8]

This ability to multiplex makes dCas12a an ideal Cas system to upregulate the multiple genes involved in 22q DS.

22q11.2 Deletion Syndrome

22q11.2 DS is common and is the most frequent chromosomal microdeletion syndrome. The prevalence of this disorder has been estimated to range from 1 per 3,000 to 1 per 6,000 live births, based on the diagnosis of infants with major birth defects and a few population screening studies conducted between the early 1990s and early 2000s using FISH technology.[9–13] Today, most (90–95%) newly identified patients with 22q11.2DS are found to have *de novo* deletions — that is, neither parent has the 22q11.2 deletion.[14] However, owing to improved survival and thus higher reproductive fitness of individuals with 22q11.2 DS, the prevalence, especially of the inherited types, is expected to increase.[15,16] As 22q11.2DS is a haploinsufficient disorder, approximately half of the children of individuals with 22q11.2DS will have the deletion. Similarly, smaller, atypical nested deletions between the low copy repeats on chromosome 22 (LCR22B–LCR22D and LCR22C–LCR22D deletions)— not typically detected by clinically available FISH probes and, therefore, not included in the population studies from the 1990s — are often familial and have reduced penetrance and/or a milder expression; thus, these patients are more likely to reproduce. As a consequence, the proportion of patients ascertained at The Children's Hospital of Philadelphia with a LCR22B–LCR22D deletion inherited from an affected parent is higher (60%) than it is for patients with the typical LCR22A–LCR22D deletion (~10%).[17]

Figure4.1 - 22q11.2 DS is caused by a chromosomal microdeletion on the order of about 3 million base pairs, affecting up to 40 different protein coding genes



In the general population, 22q11.2 deletion is among the most common detectable cause of several conditions. A large proportion of patients with congenital heart disease (CHD) have 22q11.2 deletions: 52% of those with interrupted aortic arch type B, 34% of those with truncus arteriosus, 16% of those with tetralogy of Fallot and ~5–10% of those with ventricular septal defects.[18,19] Other conditions associated with 22q11.2 deletions include velopharyngeal insufficiency (12.5–30% of patients), cleft palate (10% of patients; cleft lip with or without cleft palate in 1–2% of patients),

developmental disabilities (2–3% of patients) and schizophrenia (0.5–1% of patients).[20–25] By contrast, the prevalence of the 22q11.2 deletion in other conditions with overlapping symptoms (for example, hypoparathyroidism) is as yet unknown.

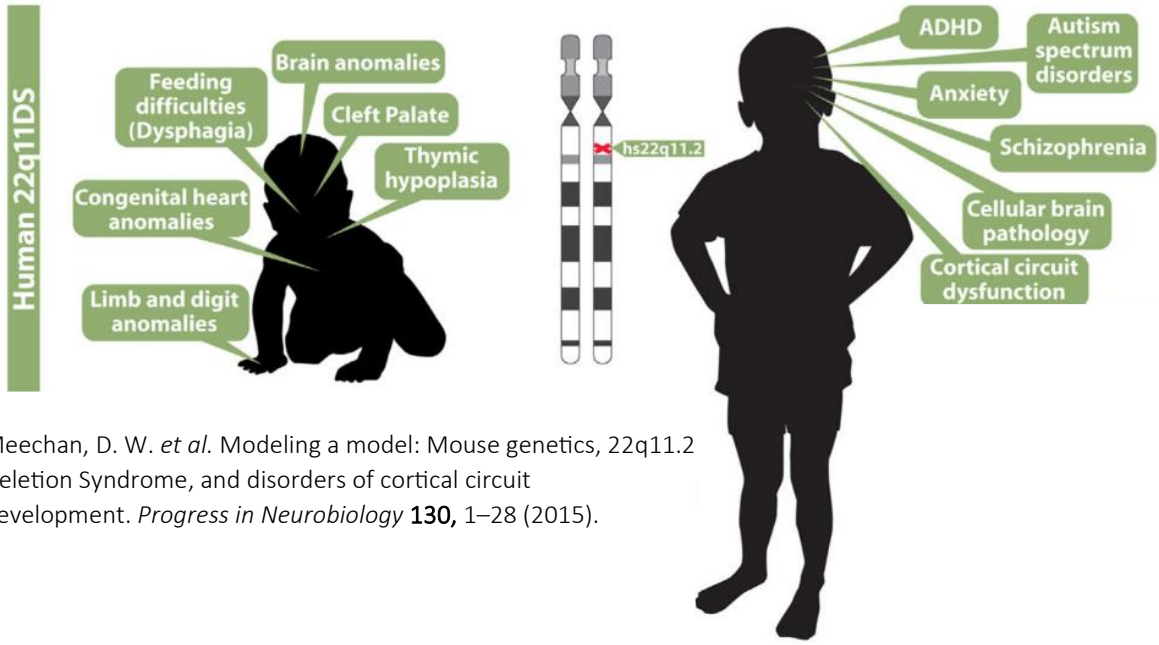
Premature mortality, although lower than suggested by early reports of DiGeorge syndrome, remains profound at all ages. About 4% of all infants with 22q11.2DS succumb, with mortality figures exceeding those for infants with similar malformations.[16,26] Cardiac defects, hypocalcaemia and airway malacia (in which cartilage defects lead to collapsibility of the airway) are risk factors for early death, with median age at death of 3–4 months.[16,26] In adults, premature death occurs at median age in the 40s; causes are multiple, including sudden unexplained death, but are not necessarily related to cardiac defects or psychotic illness.[27] Larger studies are needed to more fully clarify mortality risks prospectively at all ages.[27]

Clinical Manifestations

22q11.2 DS manifests itself differently at various ages. In infancy or childhood, typical symptoms include some combination of congenital heart defects, chronic infection, nasal regurgitation, hypernasal speech, hypocalcaemia, feeding difficulties, developmental and language delays, behavioral differences and learning disabilities.[–31] Renal abnormalities, laryngo-tracheo-esophageal abnormalities, hypothyroidism, intrauterine growth retardation, short stature, skeletal differences such as vertebral anomalies, club feet, polydactyly and scoliosis, thrombocytopenia, hearing loss, microcephaly, idiopathic seizures and hypotonia are less frequent.

In adolescence and adulthood, behavioral abnormalities, in many instances indicative of emerging psychiatric illness, can lead to the diagnosis.[30] The presence of subtle but characteristic facial features can assist with identification at any age. However, the opportunity for diagnosis can be missed when typical craniofacial and other typical congenital features such as cardiac or palatal defects are absent. This can be the case even when presenting with other high-prevalence conditions associated with 22q11.2DS, such as hypocalcaemia or psychotic illness. Major medical centers when less experienced with caring for patients with 22q11.2DS are not exempt from overlooking the 22q11.2DS diagnosis in part owing to the broad phenotypic variability.[28] Some adults are only diagnosed following the birth of an affected child.

Figure4.2 - 22q11.2 Deletion syndrome affects many organ systems. The neurological/behavioral symptoms arise around 3 -5 years of age, and are the most clinically unaddressed parts of the disorder.



Meechan, D. W. *et al.* Modeling a model: Mouse genetics, 22q11.2 Deletion Syndrome, and disorders of cortical circuit development. *Progress in Neurobiology* 130, 1–28 (2015).

The overall prevalence of considerable medical problems varies by age and ascertainment. In childhood, the triad of DiGeorge syndrome (although with highly variable degree of severity) is often ascertained: immunodeficiency (~75% of patients); congenital cardiac anomalies (~75% of patients); and hypocalcaemia due to hypoparathyroidism (~50% of patients). Other complications include palatal abnormalities (~75% of patients); manifest gastrointestinal, feeding and swallowing problems (~30% of patients); and genitourinary anomalies including renal agenesis (~30% of patients). Phenotypic expression is highly variable and ranges from severe life-threatening conditions to only a few less-severe associated features.[32] Additional complexities include considerable interfamilial and intrafamilial variability, even between identical twins.[33] Diagnosis on clinical grounds of a child with mild features requires familiarity with the condition. Dual diagnosis of other unrelated conditions (for example, ‘café-au-lait’ spots due to neurofibromatosis, skeletal disproportion due to achondroplasia or Marfan syndrome, or deep palmar and plantar creases due to trisomy 8 mosaicism) is possible, especially when features appear unusual, as 22q11.2DS is common.

Developmental and educational concerns are frequently reported in association with 22q11.2DS. Gross and fine motor difficulties, and expressive language delays and speech problems dominate in infants and toddlers.[34,35] Children with 22q11.2DS often demonstrate a significant delay in language onset; an early study indicated that ~70% of children did not speak or used only a few words or signs at 24 months of age or older.[35] Speech deficits should be discriminated from language disorders as the former often improves after velopharyngeal corrective surgery, whereas language disorders

may occur independently of palatal findings.[35] Intelligence in children and adolescents follows a normal distribution that is comparable to the general population.[34,36] However, mean IQ is only ~70, with about two-thirds of individuals falling in the IQ range of 55–85, compared with the reference IQ range of 85–115 (mean: 100) in the typically developing population. Thus, learning difficulties are very common in preschool and primary school, especially within the domains of mathematics and language comprehension.[37–39] More-severe levels of intellectual disability are uncommon.[40] However, pediatric patients with secondary insults (for example, following cardiac arrest, prolonged hypocalcemia or neonatal seizures) or primary brain malformations (for example, polymicrogyria) can have a poorer cognitive prognosis.[41] Several studies indicate that cognitive development varies with divergent trajectories and that the level of intellectual ability is not necessarily stable across the lifespan of the patient.[42] Although IQ is generally considered to be a more or less stable trait in the typically developing youth, an average decline of 7 full-scale IQ points is observed in individuals with 22q11.2DS between 8 years and 24 years of age.[43]

Individuals with 22q11.2DS are at an increased risk for developing several psychiatric disorders; the prevalence of anxiety, attention-deficit and autism spectrum disorders is increased in children with 22q11.2DS. Anxiety disorders are also profoundly increased in adults with 22q11.2DS. Conversely, bipolar disorder does not seem to be increased, and it remains unclear if major depressive disorder shows greater prevalence than in the general population. As for any phenotypic feature of 22q11.2DS, prevalence is likely to be influenced by ascertainment bias and the use of diverse assessment methods. Approximately 25% of individuals with 22q11.2DS are diagnosed

with schizophrenia, and, in turn, the 22q11.2 deletion can be found in ~1 per 100–200 individuals with schizophrenia, making 22q11.2 deletion the strongest known molecular genetic risk factor for schizophrenia.[44–46]

22q11.2 Deletion Genetics

The 22q11.2 region is one of the most structurally complex areas of the genome primarily due to several large blocks of LCRs or segmental duplications. These LCRs are >96% identical, thereby making the locus vulnerable to meiotic error. The two largest repeats, LCR22A and LCR22D, flank the typical 3-Mb 22q11.2 region that is hemizygous in ~85% of patients. The 22q11.2 deletion results from non-allelic homologous recombination between LCR22A and LCR22D. The same mechanism leads to the proximal (centromeric) nested 1.5-Mb (LCR22A–LCR22B) or 2-Mb (LCR22A–LCR22C) deletions. Patients with these nested deletions have major phenotypic features in common with patients with the typical LCR22A–LCR22D deletion. However, the frequency of the nested proximal deletions accounts for only 5–10% of all 22q11.2 deletions. Likewise, distal nested (LCR22B–LCR22D and LCR22C–LCR22D) deletions lead to overlapping phenotypic features, including cotruncal cardiac anomalies, palatal defects, and developmental differences, but the clinical features are less penetrant than those of the typical LCR22A–LCR22D deletion. Moreover, these deletions (LCR22B–LCR22D and LCR22C–LCR22D) are more frequently inherited.[1,17,47–49]

There are 90 known or predicted genes present in the typical 3-Mb 22q11.2 locus that are hemizygously deleted — including 46 protein-coding genes and seven

microRNAs (miRNAs), ten non-coding RNAs (including one read-through transcript) and 27 pseudogenes (per genome build GRCh37 — a human reference sequence produced by the Genome Reference Consortium).[50]

The most studied gene of interest in the 22q11.2 deletion region is *TBX1*, encoding a T-box transcription factor. *TBX1* was found to be a crucial gene in the LCR22A–LCR22B region using multiple mouse model approaches. Heterozygous loss-of-function mutations of *Tbx1* in the mouse result in partially penetrant cardiovascular, thymic and parathyroid defects that are reminiscent of congenital defects in 22q11.2DS.[50–52] *Tbx1*-null mice are embryonic lethal with a persistent truncus arteriosus, cleft palate and absence of the thymus and parathyroid glands. Conditional mutagenesis of *Tbx1* in the mesoderm, pharyngeal surface ectoderm or endoderm each leads to an overlapping subset of the abnormalities mentioned above, demonstrating the complexity of the tissue interactions that are required for morphogenesis of the pharyngeal derivatives.[53–55] *Tbx1* has been implicated in brain microvascular development and may play some part in cognitive and behavioural deficits.[56,57]

Another gene of interest is *DGCR8*, encoding the DGCR8 microprocessor complex subunit (also known as Pasha), a double-stranded RNA-binding protein that mediates the biogenesis of miRNAs. This observation implicates an miRNA-related mechanism in 22q11.2DS.[58,59] miRNAs are small non-coding RNAs that regulate the expression of target genes by binding to specific sites in mRNAs for translational repression or degradation. In mouse models, heterozygosity of *Dgcr8* results in neuronal deficits, whereas inactivation of both alleles in neural crest cells results in

heart defects.[60,61] Subtle alterations in miRNA expression levels can have profound effects on brain development and plasticity, especially involving synapses.[62,63] Recent studies propose that *DGCR8* may play a part in modifying the expression of genes outside of the 22q11.2 deletion region that contribute to the neuropsychiatric and other phenotypes associated with 22q11.2DS.[60,64–66]

In addition to *DGCR8*-related changes in miRNAs, the high density of miRNAs in the 22q11.2 deletion region and the accumulative insight into their function indicate that these functional non-coding RNAs may themselves have a role in the variable expression of 22q11.2DS.[66–68] These effects on expression are likely to involve not only the central nervous system (CNS) but also the cardiovascular system and other aspects of embryonic development.

Much evidence has been accumulated for a role of other individual 22q11.2 region protein-coding genes in major phenotypes of 22q11.2DS. These genes include v-crk avian sarcoma virus CT10 oncogene homologue-like (*CRKL*), encoding a cytoplasmic adaptor protein to growth factor signalling, which maps to the LCR22B–LCR22D region and acts in a dosage-sensitive manner.[69,70] Human and mouse model data indicate that haploinsufficiency of *CRKL* could be responsible for the aetiology of cardiac anomalies in individuals with nested distal deletions and seems to modulate natural killer cell function.[71] Another gene, synaptosomal-associated protein 29 kDa (*SNAP29*), encodes a soluble SNARE (soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor) protein that is predicted to mediate vesicle fusion at the endoplasmic reticulum or Golgi apparatus membranes, is highly expressed in

myelinating glial cells, is required for lamellar body formation in the skin, and is indirectly required for β1-integrin endocytosis and cell migration. Mutations in this gene have been associated with cerebral dysgenesis, neuropathy, ichthyosis and palmoplantar keratoderma (CEDNIK), Kousseff and Opitz G/BBB syndromes.[72] Scavenger receptor expressed by endothelial cells 2 protein (SREC2), encoded by *SCARF2*, contains putative epidermal growth factor-like domains in its extracellular domain, along with numerous positively charged residues in its intracellular domain, indicating that it may be involved in intracellular signalling. Homozygous or compound heterozygous mutations of *SCARF2* underlie Van den Ende–Gupta syndrome characterized by severe contractural arachnodactyly and distinctive facial dysmorphism, including triangular face, as well as skeletal anomalies.[73]

With respect to neuropsychiatric phenotypes, there are multiple gene candidates as the majority of genes in the 22q11.2 deletion region are expressed in the brain.[74] *COMT* encodes catechol-*O*-methyltransferase, one of several enzymes that degrade catecholamines, including dopamine. Its activity is of particular importance in brain regions with low expression of the presynaptic dopamine transporter, such as the prefrontal cortex. A polymorphism associated with a different level of enzymatic activity (that is, the *COMT* Val/Met functional polymorphism) has been explored in 22q11.2DS with respect to cognition and susceptibility to schizophrenia. However, multiple studies have found no association of the *COMT* functional Val/Met common allele with schizophrenia in adults with 22q11.2 DS.[75–77] There may be some effects of this common variant on frontal lobe functioning and anatomy in 22q11.2 DS, although results for overall intellect are mixed.[32] *PRODH*, encoding the enzyme proline

dehydrogenase, which breaks down proline, has also been studied, in part because pathogenetic mutations in *PRODH* are known to cause type I hyperprolinaemia, which in severe forms can cause seizures and intellectual disability. Approximately one-third of patients with 22q11.2 DS have increased levels of proline, and several studies, although not all, have shown significant associations between high proline levels and various brain outcome measures in 22q11.2 DS.[78–82] However, studies of common variants in *PRODH*, as for those in *COMT*, in 22q11.2DS show contradictory results with respect to risks for intellectual disability or schizophrenia. Zinc-finger DHHC-type-containing 8 (*ZDHHC8*), which encodes a palmitoyltransferase, has shown interesting results in studies of mutant mouse models, with effects on axonal growth and terminal arborization, and potential functional implications for synaptic connections and working memory. Another 22q11.2 region candidate is *RANBP1*, encoding a binding protein for the small GTPase Ran. As a regulator of the Ran complex, this protein has multiple functions — including cilia formation and modulation of mitosis — that may contribute to the CNS and other phenotypes of 22q11.2DS. Evidence for a role in neurogenesis places *RANBP1* as a candidate for the cortical circuits implicated in disorders associated with 22q11.2DS, such as attention-deficit disorders, autism, and schizophrenia.[80]

Due to the complexity of the genes affected by the haploinsufficiency, we focused on the genes associated with the neurological and behavioral phenotypes. **Figure 4.3** shows a list of all the genes associated with the 22q11.2 DS deletion, as well as an indication of which set of symptoms are associated with the disorder. Overall, we

identified a list of 16 genes based off this evidence as potential targets for rescue through CRISPR activation. A table of these genes are shown in **Figure 4.3**.

Figure 4.3- About 16 different genes have been shown from the literature to be associated with the neurological/cognitive phenotypes of the disorder.

	Interval					
DGCR6	A–B	–	–	++	–	–
PRODH	A–B	–	++	–	–	–
DGCR2	A–B	++	++	–	–	–
DGCR14	A–B	–	–	–	–	+
TSSK2	A–B	–	–	–	–	–
GSC2	A–B	–	–	–	–	–
SLC25A1	A–B	++	++	++	++	–
CLTCL1	A–B	++	++	–	–	–
HIRA	A–B	++	–	++	–	–
MRPL40	A–B	–	–	–	–	–
UFD1L	A–B	–	++	++	–	–
CDC45	A–B	++	–	–	++	–
CLDN5	A–B	–	–	++	–	–
SEPT5	A–B	–	++	–	–	–
GP1BB	A–B	–	–	–	–	–
SEPT5-GP1BB	A–B	++	++	–	–	–
TBX1	A–B	++	++	++	++	++
GNB1L	A–B	–	++	–	–	–
TXNRD2	A–B	–	–	++	–	–
COMT	A–B	–	++	–	–	–
ARVCF	A–B	–	++	–	+	–
TANGO2	A–B	++	++	++	–	–
DGCR8	A–B	++	++	++	–	–
TRMT2A	A–B	–	–	–	–	–
RANBP1	A–B	++	++	–	++	–
ZDHHC8	A–B	++	++	–	–	–
RTN4R	A–B	?	++	–	–	+
DGCR6L	A–B			++		
ZNF74	B–D	–	–	–	–	–
SCARF2	B–D	++	–	–	++	–
KHL22	B–D	–	–	–	–	–
MED15	B–D	–	–	–	–	–
PI4KA	B–D	++	++	+	–	–
SERPIND1	B–D	–	–	–	–	–
SNAP29	B–D	++	++	–	++	–
CRKL	B–D	++	–	++	++	++
LZTR1	B–D	–	–	++	++	–
THAP7	B–D	–	–	–	–	–
P2X6R	B–D	–	–	–	–	–
SLC7A4	B–D	–	–	–	–	–
HIC2	B–D	–	–	++	–	–

Motahari et al. Journal of Neurodevelopmental Disorders
https://doi.org/10.1186/s11689-019-9267-z

Cas12a CRISPR activation

Cas12a (also known as Cpf1) is a type of CRISPR-associated endonuclease enzyme that can be used for genome editing, just like Cas9. However, Cas12a differs from Cas9 in several ways.

PAM sequence: Cas9 requires a short PAM (Protospacer Adjacent Motif) sequence, typically NGG, immediately adjacent to the target DNA sequence. In contrast, Cas12a recognizes a different PAM sequence, typically TTTV (where V can be A, C, or G), which can be located further away from the target sequence.

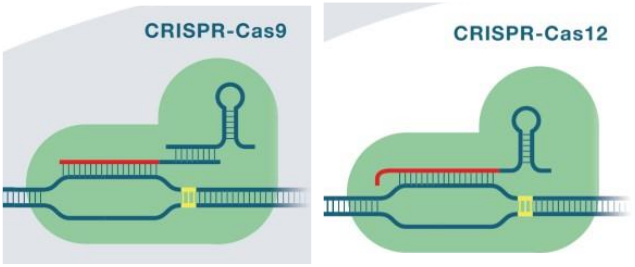
Cleavage position: Cas9 cleaves both DNA strands at the same position, generating blunt ends or overhangs depending on the target site. In contrast, Cas12a cleaves the DNA strand further away from the PAM sequence, generating staggered ends with a 5' overhang on the target site.

sgRNA structure: Cas9 requires two components, a single guide RNA (sgRNA) and a tracrRNA, to form a complex that recognizes the target DNA. In contrast, Cas12a requires only a single, shorter sgRNA to form a complex that recognizes the target DNA.

Target specificity: Cas12a is known to have a higher target specificity compared to Cas9, which means that it is less likely to generate off-target effects. This is because Cas12a recognizes longer and more complex target sequences compared to Cas9.[8]

These differences make Cas12a a valuable tool for genome editing, especially in situations where high target specificity is important. Additionally, Cas12a can also be used for CRISPR activation, a technique that involves using a deactivated Cas12a enzyme fused to a transcriptional activator domain to upregulate gene expression at specific target sites.[8]

Figure 4.4 - Cas12a has several differences from the more widely studied Cas9 that make it more beneficial to use for multiplexed editing for 22q11.2 DS.



	Cas9	Cas12a
Size	1600 aa	1100 aa
Guide Length	18-24 nt	18-25 nt
Total Guide Length	100 nt	42-44 nt
PAM	NGG- 3'	5'- TTTV
Cut	Blunt-ended dsDNA break	5 nt 5' overhang dsDNA break
Guide RNA Processing	Requires RNAse III + tracrRNA	Self-process guide RNAs

Multiplex CRISPR activation with Cas12a (also known as CRISPRa) is a technique used to upregulate the expression of multiple genes simultaneously. It involves the use of a deactivated Cas12a enzyme fused to a transcriptional activator domain, such as VP64 or p65, which is guided to specific target sites in the genome using a single guide RNA (sgRNA). For this study, we used the VPR activator, consisting of VP65, p65, and Rta. More information regarding VPR can be found in

Chapter 3 of the dissertation, where it was employed to rescue the Syngap1 haploinsufficiency.[8]

The steps involved in multiplex CRISPRa with Cas12a are as follows:

1. Design sgRNAs: Design sgRNAs that target the promoter regions of the genes of interest. The sgRNA sequence should be specific to the target gene, and it should be followed by the Cas12a recognition site (TTTV) to allow for Cas12a binding and activation of gene expression.
2. Expression of Cas12a-activator fusion protein: Express the deactivated Cas12a-activator fusion protein and the sgRNAs in the cells of interest using plasmid transfection or viral transduction.
3. Cas12a binding and activation: The Cas12a-activator fusion protein binds to the sgRNA, and the complex is guided to the target gene promoter region. This leads to the recruitment of transcriptional activators, such as VP64, p65, and Rta, to the promoter region, thereby activating gene expression.
4. Multiplex activation: The same approach can be used to activate multiple genes simultaneously by expressing sgRNAs that target different gene promoter regions along with the Cas12a-activator fusion protein.
5. Analysis: The expression levels of the target genes can be measured using techniques such as qPCR.

Multiplex CRISPRa with Cas12a offers a powerful tool for studying gene function and regulatory networks in a high-throughput manner. By activating multiple genes simultaneously, researchers can gain insights into complex cellular processes and

identify potential drug targets. This approach is particularly useful for 22q DS because in order to rescue this haploinsufficiency, we hypothesize that 16 genes will need to be upregulated.

MATERIALS AND METHODS

CRISPR/Cas12a Guide Design

CRISPR guide RNAs were designed for a Cas12a system. The bioinformatics tool ChopChop was utilized to identify all the potential guides 400bp upstream of the promoter for each gene of interest.[83] **Table 4.1** shows the specific sequence for each guide for al 16 genes. Due to the stringent nature of the TTTV PAM, nearly all guides identified for each promoter were screened. In total, about 2-6 guides were chosen for each gene. Although the ChopChop algorithm ranks guides based off potential off-target affects and predicted activity levels, guides were manually chosen to include guides in different regions of the promoter to account for any variation in activity due to steric blocking by other transcription factors.

Following identification of guide sequences *in silico*, each guide was cloned into a vector expressing the sgRNA, dCas12a, and VPR (dCas12a-VPR-sgRNA). This vector was obtained from Campa *et al*, and cloned using their preferred protocol.[8] Oligonucleotides were ordered from IDT containing the 23bp spacer designed from ChopChop, with overhangs enabling sticky-end ligation into the dCas12a-VPR-sgRNA digested with *SpaI*.[84] The dCas12a-VPR-sgRNA vector backbone was digested with *SpaI*, gel purified using QIAquick Gel Extraction Kit. Oligonucleotides were

Table4.1 - Cas12a Guide Sequences for all 16 genes implicated in the neurological and/or cognitive symptoms of 22q11.2 deletion syndrome. All guides were designed to target the mouse promoter of their respective gene, and were preceded by a TTTV PAM, characteristic of guide sequences for the Cas12a system.

Tbx1_004	Tbx1	Mouse	TTTAGGGCCTGTAAAATGGTCTAGTCA
Tbx1_005	Tbx1	Mouse	TTTGCGCACAGGGGTCCTGAGCTTACA
DGCR2_001	DGCR2	Mouse	TTTCCTGCTAGCCCCGCCCTTGGCCCC
DGCR2_002	DGCR3	Mouse	TTTGGTACCTCGCTGCCAAGGCTGCGT
SLC25A1_001	SLC25A1	Mouse	TTTAGAGACACGTGCGTCTGGAGCGCG
SLC25A1_002	SLC25A1	Mouse	TTTCCCGGCTGGCACGGGGCGTGGCGT
PRODH_001	PRODH	Mouse	TTTCCTGTCCCCAAGCTGACCAGGTGG
PRODH_002	PRODH	Mouse	TTTGGGCTGGGGAGGCGGGGTCAGCAA
PRODH_003	PRODH	Mouse	TTTGGCGTACTTCCCGTTGCTGACCCC
PRODH_004	PRODH	Mouse	TTTGTCCAGCACAGAAAGTAGCTCCAT
PRODH_005	PRODH	Mouse	TTTAAGAAGCTTTGTCCAGCACAGAAA
PRODH_006	PRODH	Mouse	TTTAGCTCCGGGATTTTCCTGTCCCCA
COMT_001	COMT	Mouse	TTTCCCCAGGTGTCTAAAACGTTGGCG
COMT_002	COMT	Mouse	TTTGTCCGTAATCCCACCCCCATGTAG
COMT_003	COMT	Mouse	TTTGTTTTTGTCCGTAATCCCACCCCC
COMT_004	COMT	Mouse	TTTAGACACCTGGGGAAAAAGCAGCCC
COMT_005	COMT	Mouse	TTTGTGTCCCCAATATTAATCGTAGCC
SEPT5_001	SEPT5	Mouse	TTTATAGAATGGCCACTCGATGGAAGG
SEPT5_002	SEPT5	Mouse	TTTGGTTCTCCCGGCCGCGCGAGTGCA
DGCR8_001	DGCR8	Mouse	TTTGCAAGGGCCTTCGCGGGTTGTGGC
DGCR8_002	DGCR8	Mouse	TTTGAGTTTGCGAATGACAATGGACCT
HIRA_001	HIRA	Mouse	TTTGCGGGCGTCCCGCCTTCTGTGGAT
HIRA_002	HIRA	Mouse	TTTCACGCCACCCGCTGGCGGGCAGCT
UFD1_001	UFD1	Mouse	TTTCCAATAAGTTTCCAATACACGGTT
UFD1_002	UFD1	Mouse	TTTAGATAGTACTCTCTTCCAAACAAC
UFD1_003	UFD1	Mouse	TTTGGAAGAGAGTACTATCTAAACATT
UFD1_004	UFD1	Mouse	TTTCCAATACACGGTTCGTCGAATTAG
UFD1_005	UFD1	Mouse	TTTCTGTTTTTCCGCCGCTGACTTCAG
UFD1_006	UFD1	Mouse	TTTCCGCCGCTGACTTCAGGCCTCCGT
CDC45_001	CDC45	Mouse	TTTCCAATAAGTTTCCAATACACGGTT
CDC45_002	CDC45	Mouse	TTTATAATGTTTAGATAGTACTCTCTT
CDC45_003	CDC45	Mouse	TTTGGAAGAGAGTACTATCTAAACATT
CDC45_004	CDC45	Mouse	TTTCCGGCGTTGATCCCGGCCATCATG
CDC45_005	CDC45	Mouse	TTTCCAATACACGGTTCGTCGAATTAG
CDC45_006	CDC45	Mouse	TTTGAATGTGTTTCCAATAAGTTTCCA
GNB1_001	GNB1	Mouse	TTTCCGCCCCGCCCCTCCTTCCCTCTG
GNB1_002	GNB1	Mouse	TTTCGGCCCTCCCGGGCTGGTCATTCA
GNB1_003	GNB1	Mouse	TTTAAGCCTCTCGCGCACGCGCAGAGG
ARVCF_001	ARVCF	Mouse	TTTCCCATTGGTAGAGAGACGCGCCCC
ARVCF_002	ARVCF	Mouse	TTTGTATTCCGGGAGGCTGAACTGGGG
ARVCF_003	ARVCF	Mouse	TTTGAGGTGCGCCCCCCTGCTGTAGTC
ARVCF_004	ARVCF	Mouse	TTTCCCACCTCATCTCCTTAGCTTTGT
ARVCF_005	ARVCF	Mouse	TTTGCTCTAAGACATAGTTCTGGTATG

annealed, and phosphorylated oligonucleotides were then diluted at 1:200 before being ligated into the *SpaI* digested dCas12a-VPR-sgRNA vector using T4 DNA Ligase (NEB). Successful insertion of the oligonucleotides containing the sgRNA sequence was confirmed via Sanger sequencing.

Cell Transfection

In order to screen the guides for the strongest potential activator for each gene, Neuro2A cells were transfected using Lipofectamine 3000, with subsequent RT-qPCR for relative gene mRNA levels. Neuro2A cells were maintained in DMEM media supplemented with 10% FBS and Penicillin/Streptomycin at 37C under 5% CO2. A total of 1 million cells at 60-70% confluency were transfected using Lipofectamine 3000 following the manufacturer's instructions. Transfections were performed in 12-well plates using 500ng of the dCas12a-VPR-sgRNA expressing plasmid. Cells were harvested for RNA extraction at day 3 post lipofection. RNA was harvested using Qiagen's RNeasy kit.

Quantitative PCR (qPCR)

RNA was reverse transcribed using the AB high-capacity cDNA synthesis kit (ThermoFisher) according to the manufacturer's instructions. Primers for quantitative PCR (qPCR) were designed using Primer3. qPCR was performed in triplicate using PowerUp SYBR Green Master Mix (ThermoFisher) with the CFX384 Real-Time System C1000 Touch system (Bio-Rad). Gene expression analysis was performed with gene specific primers using three biological replicates. Relative target gene expression was calculated as the difference between the target gene and the GAPDH reference gene

(dCq = Cq[target] − Cq[GAPDH]). Gene expression results are indicated as fold change to a reference sample (dCas12a combination without targeting gRNA), using the ddCq method. Results from the guide screen are shown in **Figure 4.5**.

DISCUSSION

This chapter of the thesis accomplished the first step in developing an epigenetic therapy towards treating the cognitive/behavioral deficits in 22q11.2 DS: screening guide RNAs for their ability to upregulate 16 genes implicated in the disorder. From the results in **Figure 4.5**, the strongest activator for each gene can be utilized for a multiplex approach: Tbx1_003, DCGR2_002, SLC25A1_002, PRODH_003, COMT_004, SEPT5_002, DGCR8_002, HIRA_002, UFD1_006,CDC45_001, GNB1_001, ARVCF_004, TANGO_005, RANBP1_001, ZDHHC8_002, RTN4R_001. These 16 guides can then be cloned into the vector shown in **Figure 4.6**. Each of these guides were the strongest activators for their respective genes. Although the therapeutic activation target is still an open question, starting the with the strongest activator makes sense as studies have shown that *in vitro* activation often overestimates *in vivo* activation.[6] This single vector will express dCas12a, VPR, and each of these 16 guide RNAs.

Once multiplex activation is achieved *in vitro*, the next step towards proof-of-concept of a viable therapy is to assess multiplex gene activation and phenotypic correction in a 22q DS murine model *in vivo.* The dCas12a-VPR and validated multiplexed crRNA array can be packaged in AAV-PHP.eB, allowing widespread brain distribution upon systemic injection into WT and Df1/+ mice, a mouse model of the

disease.[85] Biodistribution and gene activation can be assessed by RNA-seq, Western blot, and IHC. Df1/+ mice display many physical and behavioral phenotypes that are analogous to human symptoms. One robust and potentially rescuable phenotype is the prepulse inhibition (PPI) of the acoustic startle response.[74] PPI is the phenomenon in which a weak prestimulus or “prepulse” suppresses the response to a subsequent startle stimulus. Reduced PPI of the startle response is indictive of impaired sensorimotor gating. Deficits in PPI have been observed in subjects with schizophrenia and other neuropsychiatric diseases.[57]

One problem that may arise is that the 16 genes chosen are not sufficient for behavioral rescue the haploinsufficiency. A potential fix for this issue is to increase the number of guides in the multiplex array to target additional genes in the deleted region. If need be, multiple AAV vectors can be developed to potentially deliver guides for every deleted gene in the Df1 deletion region. Additionally, the therapeutic window chosen correlating to the 3–5-year age range for human development may not be early enough to see reversal of PPI. Further experimentation with dosing at different ages would need to be conducted to determine the ideal therapeutic window.

The development of a multiplex vector that activates multiple genes and demonstration of phenotypic correction in a mouse model for 22qDS would be the first steps towards developing a viable therapy for the neurological and cognitive phenotypes of 22q DS. Additionally, the framework established by these aims would enable a better understanding of the genetics involved in 22qDS, helping unearth the gene-to-phenotype relationship for the disorder. Future steps should include refining the

therapeutic window for the CRISPR-based therapeutic. Better models of 22qDS are needed that allow assessment of the anxiety, autism spectrum disorders, ADHD and other prevalent symptoms involved in 22qDS.

Figure 4.5- Panel of 16 graphs showing the results of the guide screen for each 22q11.2 DS gene conducted in Neuro2A cells. The strongest activator for each gene can be utilized for future multiplex applications. Relative gene expression was determined by comparing each guide to a no guide control consisting of just dCas12a-VPR. All guides are compared to this control, whose expression was normalized to 1.

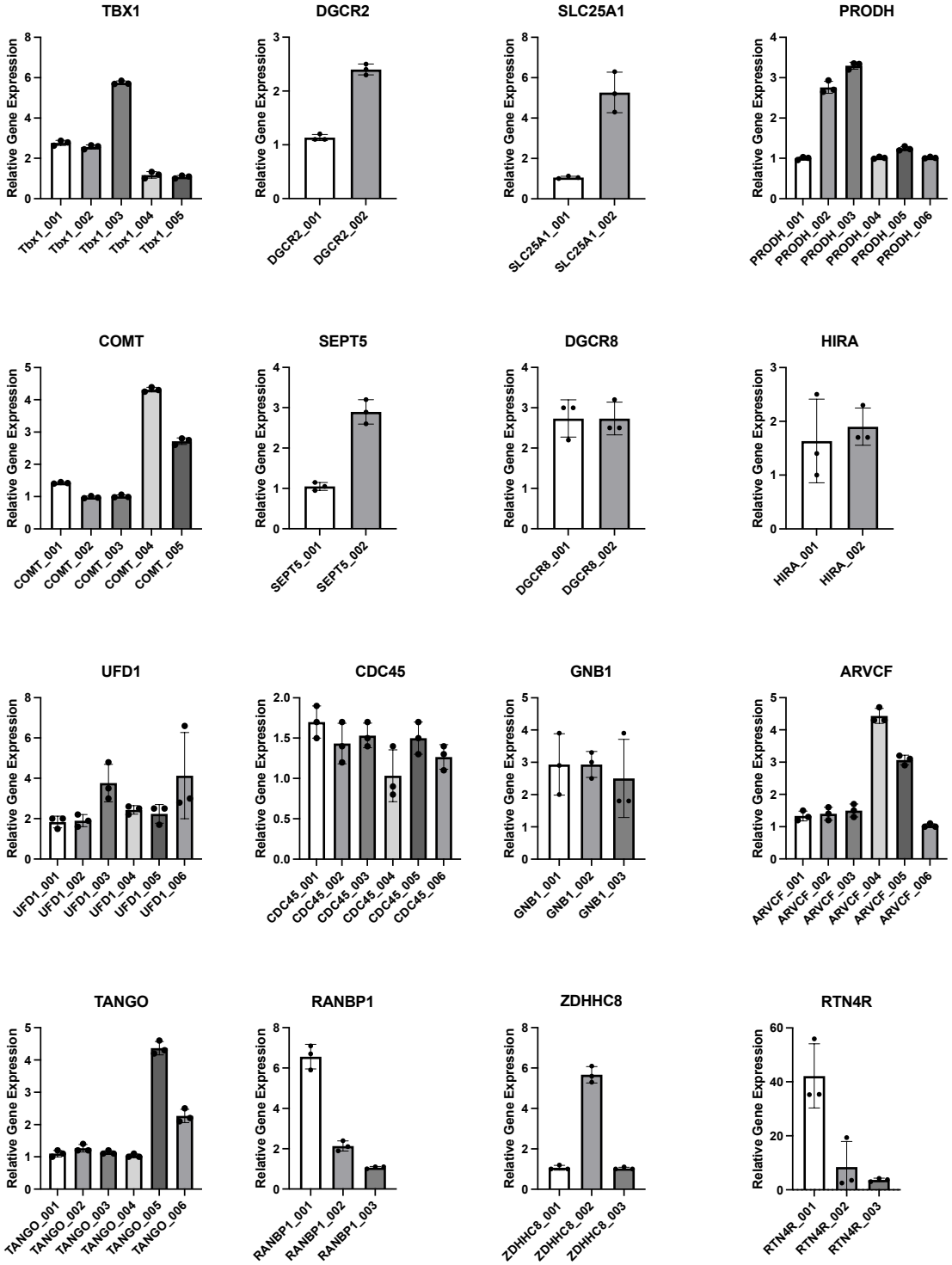
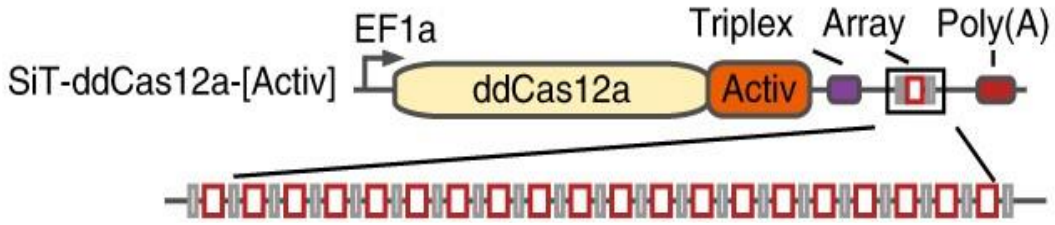


Figure4.6 - Multiple genes can be upregulated with a Cas12a Array



Campa, C.C., Weisbach, N.R., Santinha, A.J. *et al.* Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nat Methods* **16,** 887–893 (2019) doi:10.1038/s41592 -019-0508-6

REFERENCES

1. Morrow, B. E., McDonald-McGinn, D. M., Emanuel, B. S., Vermeesch, J. R. & Scambler, P. J. Molecular genetics of 22q11.2 deletion syndrome. *Am. J. Med. Genet. A.* **176**, 2070–2081 (2018).
2. Prasad, S. E., Howley, S. & Murphy, K. C. Candidate genes and the behavioral phenotype in 22q11.2 deletion syndrome. *Dev. Disabil. Res. Rev.* **14**, 26–34 (2008).
3. Motahari, Z., Moody, S. A., Maynard, T. M. & LaMantia, A.-S. In the line-up: deleted genes associated with DiGeorge/22q11.2 deletion syndrome: are they all suspects? *J. Neurodev. Disord.* **11**, 7 (2019).
4. Pickar-Oliver, A. & Gersbach, C. A. The next generation of CRISPR–Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.* **20**, 490–507 (2019).
5. Tak, Y. E. *et al.* Inducible and multiplex gene regulation using CRISPR–Cpf1-based transcription factors. *Nat. Methods* **14**, 1163–1166 (2017).
6. Matharu, N. *et al.* CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency. *Science* **363**, eaau0629 (2019).
7. Zetsche, B. *et al.* Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat. Biotechnol.* **35**, 31–34 (2017).
8. Campa, C. C., Weisbach, N. R., Santinha, A. J., Incarnato, D. & Platt, R. J. Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nat. Methods* 1–7 (2019) doi:10.1038/s41592-019-0508-6.
9. Botto, L. D. A population-based study of the 22q11.2 deletion: phenotype, incidence, and contribution to major birth defects in the population. *Pediatrics* **112**, (2003).
10. Devriendt, K., Fryns, J. P., Mortier, G., van Thienen, M. N. & Keymolen, K. The annual incidence of DiGeorge/velocardiofacial syndrome. *J Med Genet* **35**, (1998).
11. Goodship, J., Cross, I., LiLing, J. & Wren, C. A population study of chromosome 22q11 deletions in infancy. *Arch Child* **79**, (1998).
12. Oskarsdottir, S., Vujic, M. & Fasth, A. Incidence and prevalence of the 22q11 deletion syndrome: a population-based study in Western Sweden. *Arch Child* **89**, (2004).
13. zenas Du Montcel, S., Mendizabai, H., Ayme, S., Levy, A. & Philip, N. Prevalence of 22q11 microdeletion. *J Med Genet* **33**, (1996).
14. McDonald-McGinn, D. M. Phenotype of the 22q11.2 deletion in individuals identified through an affected relative: cast a wide FISHing net! *Genet Med* **3**, (2001).
15. Costain, G., Chow, E. W., Silversides, C. K. & Bassett, A. S. Sex differences in reproductive fitness contribute to preferential maternal transmission of 22q11.2 deletions. *J Med Genet* **48**, (2011).
16. Repetto, G. M. Case fatality rate and associated factors in patients with 22q11 microdeletion syndrome: a retrospective cohort study. *BMJ Open* **4**, (2014).
17. McDonald-McGinn, D. M. et al. The perplexing prevalence of familial nested 22q11.2 deletions. ASGH[online], (2014).
18. Goldmuntz, E. Microdeletions of chromosomal region 22q11 in patients with congenital conotruncal cardiac defects. *J Med Genet* **30**, (1993).
19. Peyvandi, S. 22q11.2 deletions in patients with conotruncal defects: data from 1,610 consecutive cases. *Pediatr Cardiol* **34**, (2013).

20. Zori, R. T. Prevalence of 22q11 region deletions in patients with velopharyngeal insufficiency. *Am J Med Genet* **77**, (1998).
21. Boorman, J. G., Varma, S. & Mackie Ogilvie, C. Velopharyngeal incompetence and chromosome 22q11 deletion. *Lancet* **357**, (2001).
22. Ruiter, E. M., Bongers, E. M., Smeets, D., Kuijpers-Jagtman, A. M. & Hamel, B. C. No justification of routine screening for 22q11 deletions in patients with overt cleft palate. *Clin Genet* **64**, (2003).
23. Rauch, A. Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. *Am J Med Genet* **140**, (2006).
24. Bassett, A. S. Clinically detectable copy number variations in a Canadian catchment population of schizophrenia. *J Psychiatr Res* **44**, (2010).
25. Horowitz, A., Shifman, S., Rivlin, N., Pisante, A. & Darvasi, A. A survey of the 22q11 microdeletion in a large cohort of schizophrenia patients. *Schizophr Res* **73**, (2005).
26. McDonald-McGinn, D. M. The 22q11.2 deletion in African-American patients: an underdiagnosed population? *Am J Med Genet A* **134**, (2005).
27. Bassett, A. S. Premature death in adults with 22q11.2 deletion syndrome. *J Med Genet* **46**, (2009).
28. Bassett, A. S. Practical guidelines for managing patients with 22q11.2 deletion syndrome. *J Pediatr* **159**, (2011).
29. McDonald-McGinn, D. M. The Philadelphia story: the 22q11.2 deletion: report on 250 patients. *Genet Couns* **10**, (1999).
30. Swillen, A. & McDonald-McGinn, D. Developmental trajectories in 22q11.2 deletion syndrome. *Am J Med Genet C Semin Med Genet* **169**, (2015).
31. McDonald-McGinn, D. M. The 22q11.2 deletion: screening, diagnostic workup, and outcome of results; report on 181 patients. *Genet Test* **1**, (1997).
32. Philip, N. & Bassett, A. S. Cognitive, behavioural and psychiatric phenotype in 22q11.2 deletion syndrome. *Behav Genet* **41**, (2011).
33. Vergaelen, E. 3 generation pedigree with paternal transmission of the 22q11.2 deletion syndrome: intrafamilial phenotypic variability. *Eur J Med Genet* **58**, (2015).
34. Swillen, A. Early motor development in young children with 22q.11 deletion syndrome and a conotruncal heart defect. *Dev Med Child Neurol* **47**, (2005).
35. Solot, C. B. Communication issues in 22q11.2 deletion syndrome: children at risk. *Genet Med* **3**, (2001).
36. De Smedt, B. Intellectual abilities in a large sample of children with velo-cardio-facial syndrome: an update. *J Intellect Disabil Res* **51**, (2007).
37. De Smedt, B., Swillen, A., Verschaffel, L. & Ghesquiere, P. Mathematical learning disabilities in children with 22q11.2 deletion syndrome: a review. *Dev Disabil Res Rev* **15**, (2009).
38. Wang, P. P., Woodin, M. F., Kreps-Falk, R. & Moss, E. M. Research on behavioral phenotypes: velocardiofacial syndrome (deletion 22q11.2). *Dev Med Child Neurol* **42**, (2000).
39. Glaser, B. Language skills in children with velocardiofacial syndrome (deletion 22q11.2). *J Pediatr* **140**, (2002).
40. Evers, L. J. Psychopathology in adults with 22q11 deletion syndrome and moderate and severe intellectual disability. *J Intellect Disabil Res* **58**, (2014).

41. Cheung, E. N. Neonatal hypocalcemia, neonatal seizures, and intellectual disability in 22q11.2 deletion syndrome. *Genet Med* **16**, (2014).
42. Duijff, S. N. Cognitive development in children with 22q11.2 deletion syndrome. *Br J Psychiatry* **200**, (2012).
43. Vorstman, J. A. Cognitive decline preceding the onset of psychosis in patients with 22q11.2 deletion syndrome. *JAMA Psychiatry* **72**, (2015).
44. Schneider, M. Psychiatric disorders from childhood to adulthood in 22q11.2 deletion syndrome: results from the International Consortium on Brain and Behavior in 22q11.2 deletion syndrome. *Am J Psychiatry* **171**, (2014).
45. Fung, W. L. Elevated prevalence of generalized anxiety disorder in adults with 22q11.2 deletion syndrome. *Am J Psychiatry* **167**, (2010).
46. Bassett, A. S. & Chow, E. W. C. 22q11 deletion syndrome: A genetic subtype of schizophrenia. *Biological Psychiatry* vol. 46 882–891 (1999).
47. Edelmann, L., Pandita, R. K. & Morrow, B. E. Low-copy repeats mediate the common 3-Mb deletion in patients with velo-cardio-facial syndrome. *Am J Hum Genet* **64**, (1999).
48. Shaikh, T. H. Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Hum Mol Genet* **9**, (2000).
49. Saitta, S. C. Aberrant interchromosomal exchanges are the predominant cause of the 22q11.2 deletion. *Hum Mol Genet* **13**, (2004).
50. Jerome, L. A. & Papaioannou, V. E. DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. *Nat Genet* **27**, (2001).
51. Merscher, S. TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. *Cell* **104**, (2001).
52. Lindsay, E. A. Tbx1 haploinsufficieny in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature* **410**, (2001).
53. Zhang, Z. Tbx1 expression in pharyngeal epithelia is necessary for pharyngeal arch artery development. *Development* **132**, (2005).
54. Zhang, Z., Huynh, T. & Baldini, A. Mesodermal expression of Tbx1 is necessary and sufficient for pharyngeal arch and cardiac outflow tract development. *Development* **133**, (2006).
55. Papangeli, I. & Scambler, P. The 22q11 deletion: DiGeorge and velocardiofacial syndromes and the role of TBX1. *Wiley Interdiscip Rev Dev Biol* **2**, (2013).
56. Cioffi, S. Tbx1 regulates brain vascularization. *Hum Mol Genet* **23**, (2014).
57. Paylor, R. Tbx1 haploinsufficiency is linked to behavioral disorders in mice and humans: implications for 22q11 deletion syndrome. *Proc Natl Acad Sci USA* **103**, (2006).
58. Xu, B., Hsu, P. K., Stark, K. L., Karayiorgou, M. & Gogos, J. A. Derepression of a neuronal inhibitor due to miRNA dysregulation in a schizophrenia-related microdeletion. *Cell* **152**, (2013).
59. Xu, B., Karayiorgou, M. & Gogos, J. A. MicroRNAs in psychiatric and neurodevelopmental disorders. *Brain Res* **1338**, (2010).
60. Stark, K. L. Altered brain microRNA biogenesis contributes to phenotypic deficits in a 22q11-deletion mouse model. *Nat Genet* **40**, (2008).

61. Chapnik, E., Sasson, V., Blelloch, R. & Hornstein, E. Dgcr8 controls neural crest cells survival in cardiovascular development. *Dev Biol* **362**, (2012).
62. Petri, R., Malmevik, J., Fasching, L., Akerblom, M. & Jakobsson, J. miRNAs in brain development. *Exp Cell Res* **321**, (2014).
63. Beveridge, N. J., Gardiner, E., Carroll, A. P., Tooney, P. A. & Cairns, M. J. Schizophrenia is associated with an increase in cortical microRNA biogenesis. *Mol Psychiatry* **15**, (2010).
64. Earls, L. R. & Zakharenko, S. S. A synaptic function approach to investigating complex psychiatric diseases. *Neuroscientist* **20**, (2013).
65. Brzustowicz, L. M. & Bassett, A. S. miRNA-mediated risk for schizophrenia in 22q11.2 deletion syndrome. *Front Genet* **3**, (2012).
66. Merico, D. MicroRNA dysregulation, gene networks, and risk for schizophrenia in 22q11.2 deletion syndrome. *Front Neurol* **5**, (2014).
67. Zhao, D. MicroRNA profiling of neurons generated using induced pluripotent stem cells derived from patients with schizophrenia and schizoaffective disorder, and 22q11.2 del. *PLoS ONE* **10**, (2015).
68. Sellier, C. Decreased DGCR8 expression and miRNA dysregulation in individuals with 22q11.2 deletion syndrome. *PLoS* **9**, (2014).
69. Guris, D. L., Fantes, J., Tara, D., Druker, B. J. & Imamoto, A. Mice lacking the homologue of the human 22q11.2 gene CRKL phenocopy neurocristopathies of DiGeorge syndrome. *Nat Genet* **27**, (2001).
70. Racedo, S. E. Mouse and human CRKL is dosage sensitive for cardiac outflow tract formation. *Am J Hum Genet* **96**, (2015).
71. Zheng, P. Molecular mechanisms of functional natural killer deficiency in patients with partial DiGeorge syndrome. *J Allergy Clin Immunol* **135**, (2015).
72. McDonald-McGinn, D. M. Hemizygous mutations in SNAP29 unmask autosomal recessive conditions and contribute to atypical findings in patients with 22q11.2DS. *J Med Genet* **50**, (2013).
73. Bedeschi, M. F. Unmasking of a recessive SCARF2 mutation by a 22q11.12 de novo deletion in a patient with Van den Ende–Gupta syndrome. *Mol Syndr.* **1**, (2010).
74. Guna, A., Butcher, N. J. & Bassett, A. S. Comparative mapping of the 22q11.2 deletion region and the potential of simple model organisms. *J Neurodev Disord* **7**, (2015).
75. Bassett, A. S., Caluseriu, O., Weksberg, R., Young, D. A. & Chow, E. W. Catechol-O-methyl transferase and expression of schizophrenia in 73 adults with 22q11 deletion syndrome. *Biol Psychiatry* **61**, (2007).
76. Murphy, K. C., Jones, L. A. & Owen, M. J. High rates of schizophrenia in adults with velo-cardio-facial syndrome. *Arch Gen Psychiatry* **56**, (1999).
77. Gothelf, D. Risk factors and the evolution of psychosis in 22q11.2 deletion syndrome: a longitudinal 2-site study. *J Am Acad Child Adolesc Psychiatry* **52**, (2013).
78. Goodman, B. K., Rutberg, J., Lin, W. W., Pulver, A. E. & Thomas, G. H. Hyperprolinaemia in patients with deletion (22)(q11.2) syndrome. *J Inherit Metab Dis* **23**, (2000).

79. Magnée, M. J., Lamme, V. A., de Sain-van der Velden, M. G., Vorstman, J. A. & Kemner, C. Proline and COMT status affect visual connectivity in children with 22q11.2 deletion syndrome. *PLoS* **6**, (2011).
80. Paronett, E. M., Meechan, D. W., Karpinski, B. A., LaMantia, A.-. S. & Maynard, T. M. Ranbp1, deleted in DiGeorge/22q11.2 deletion syndrome, is a microcephaly gene that selectively disrupts layer 2/3 cortical projection neuron generation. *Cereb Cortex* **25**, (2014).
81. Raux, G. Involvement of hyperprolinemia in cognitive and psychiatric features of the 22q11 deletion syndrome. *Hum Mol Genet* **16**, (2007).
82. Vorstman, J. A. S. Proline affects brain function in 22q11DS children with the low activity COMT158 allele. *Neuropsychopharmacology* **34**, (2008).
83. Labun, K. *et al.* CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* **47**, W171–W174 (2019).
84. Stringer, B. W. *et al.* A reference collection of patient-derived cell line and xenograft models of proneural, classical and mesenchymal glioblastoma. *Sci. Rep.* **9**, 4902 (2019).
85. Sivagnanasundaram, S. Differential gene expression in the hippocampus of the Df1/+ mice: a model for 22q11.2 deletion syndrome and schizophrenia. *Brain Res* **1139**, (2007).

Chapter 5- Conclusion

CONCLUSION

This thesis began by introducing the concepts of CRISPR/Cas9 genome engineering, it's promise in treating genetic disease, and how it can be modified for epigenetic editing to expand the list of potentially treatable diseases. This dissertation provided preliminary results in the development of epigenetic therapies for three genetic disorders (Neurofibromatosis type 1, SYNGAP1-related intellectual disability, and 22q11.2 deletion syndrome). NF1 utilized a CRISPR repression system, by designing guide RNAs for *Map2k1* and *Map2k2*. SYNGAP1-related intellectual disability was addressed by utilizing a CRISPR activation system, for the *Syngap1* gene itself. Finally, a multiplex CRISPR activation approach was explored for 22q11.2 deletion syndrome, targeting 16 different genes.

Every chapter followed a similar order of operations in developing epigenetic therapies. First, an analysis of the genetics of the disease of interest is essential. Do genes need to be upregulated or downregulated to rescue whatever genetic deficiency plagues patients? Next, guide RNAs are designed *in silico*. The exact sequence that is targeted is primarily dictated by a PAM sequence specific for the various Cas protein intended to be utilized. In this thesis, we explored epigenetic editing with multiple Cas variants (dCas9, dCas12f, dCas12a). Each required a unique PAM sequence, which limited or expanded the list of potential guides that could be developed. Once guides are designed *in silico*, they were empirically screened for either the strongest repressor or activator, depending on the disease investigated. Although *in silico* analysis identifies ideal guides, often the guide that empirically proved to be the most ideal wasn't the one the algorithm identified. Therefore, this thesis shows that guides for downstream

applications must be tested empirically. I believe the ease in which the CRISPR/Cas9 system can be adopted for different diseases is perfectly encapsulated in this thesis. Simply by finding the right guide RNA for the right loci, it's possible to fine tune the expression of genes through activation or repression.

Overall, gene therapy and epigenetic therapies have the potential to be a promising medical treatment. While gene/epigenome therapy has shown some success in treating certain genetic disorders, there are still several challenges in implementing it to patients.

Some of these challenges include delivery. One of the main challenges in gene therapy is delivering the therapeutic genes to the target cells in the body. The delivery method must be safe and effective, as well as capable of getting the therapeutic genes into the cells where they can work. Several delivery methods are currently being researched, including both viral and non-viral vectors. Modified viral vectors are used to delivery genes of interest to the brain, with vectors derived from adeno-associated viruses (AAVs) having been used almost exclusively in clinical trials for CNS disorders.[1–5]

AAV gene therapy is a promising approach to treat genetic diseases, but it also faces several challenges, particularly related to its delivery. AAV vectors have a limited capacity to carry genetic material, which can be a significant challenge when trying to deliver larger genes or multiple genes.[1,3,4] In chapter 3, we explored the potential of using a split-Cas9 system to bypass the packaging limitation of AAV vectors. Another challenge is that the immune system can recognize AAV vectors as foreign and mount an immune response, which can reduce the efficacy of the therapy.[6] This can also limit

the ability to re-dose patients with AAV vectors in a clinical setting.[7] Some individuals may have pre-existing immunity to AAV vectors, which can limit their efficacy in these patients.[8] AAV vectors have limited tissue specificity, which can result in off-target effects or limited transduction in target tissues.[9] The route of administration can also impact the efficacy of AAV gene therapy. For example, intravenous administration may result in liver transduction, while intramuscular administration may result in muscle transduction.[9] Additionally, AAV vectors can be challenging to manufacture at scale, which can limit their availability for clinical use. Despite these challenges, researchers are actively working to overcome these challenges through the development of new AAV vectors, modifications to existing vectors, and improved manufacturing processes.[10] Despite whatever advancements are made molecularly in terms of developing an epigenetic therapy, delivery remains a major bottleneck in taking these tools to the clinic. Non-viral delivery vehicles, such as the polymer nanoparticles discussed in Chapter 2 may provide an alternative to viral delivery of gene/epigenetic therapies.[10]

It is my hope that the findings from this thesis ultimately make its way into pre-clinical studies by testing the guides identified either in disease relevant cell lines, or disease relevant mouse models. Additionally, off-target effects of all the CRISPR systems (either activation or repression) need to be investigated. Preliminary mouse studies were conducted for SYNGAP1, but the effects in NF1 and 22q11.2 DS mouse models are still unexplored. Luckily each respective community has developed robust mouse models that mimic the haploinsufficiencies, so the pipeline for these epigenetic therapies seem to have a clear pre-clinical destination.

FUTURE STEPS

Neurofibromatosis Type 1

In Chapter 2, we showed preliminary evidence of the promise of using a CRISPR repression system to downregulate *Map2k1* and *Map2k2* in Neur2A cells. The immediate next steps are to complete similar work in Schwann cells, a more clinically relevant cell line for NF1. Efficient transfection protocols need to be developed using either lentiviral systems or polymer nanoparticles (PNPs) in order to deliver the CRISPR repression platform to Schwann cells, as lipofection has proven to be very inefficient for this cell type.[11]

In cell lines, Phospho-ERK (Thr202/Tyr204) fluorescence assay is a phenotypic assay that could be done in conjunction with molecular assays that measure *Map2k1/Map2k2* expression.[12] The Phospho-ERK (Thr202/Tyr204) assay measures ERK when phosphorylated at Thr202/Tyr204. The Phospho-ERK (Thr202/Tyr204) assay uses 2 labeled antibodies: one with a donor fluorophore, the other one with an acceptor. The first antibody is selected for its specific binding to the phosphorylated motif on the protein, the second for its ability to recognize the protein independent of its phosphorylation state. Protein phosphorylation enables an immune-complex formation involving both labeled antibodies and which brings the donor fluorophore into close proximity to the acceptor, thereby generating a FRET signal. Its intensity is directly proportional to the concentration of phosphorylated protein present in the sample, and provides a means of assessing the protein phosphorylation state.[12] If our epigenetic therapy is successful in repressing *Map2k1/Map2k2,* we would expect the Phospho-ERK fluorescence assay to reduce ERK phosphorylation.

Next, evaluating the efficacy of such treatment in a clinically relevant animal model would be insightful. Mice models developed by the community have with the *Nf1*+/- genetic background develop malignant tumors and high grade malignant peripheral nerve sheath tumors that are histologically similar to their human counterparts.[13,14] Successfully delivery of a Schwann cell specific *Map2k1/Map2k2* epigenetic therapy to this *Nf1*+/- may attenuate the severity of this tumor phenotype.

SYNGAP1-related Intellectual Disability

The results outlined in Chapter 3 is preliminary work in a large collaboration between the Fink, Silverman, and Segal labs. In the efforts towards creating an epigenetic therapy for the *Syngap1* haploinsufficiency, this project was successful in identifying a promising dCas9 guide RNA that has shown to increase the expression of *Syngap1 in vitro* in Neuro2A cells. In wild-type mice, there is a slight trend, although not-statistically significant, towards activation. Future work in the pipeline includes the testing this therapy in a *Syngap1* deficient mouse, which will be done in collaboration with the Silverman lab.

The experimental design for these future experiment include injecting 5-week-old *Syngap1* deficient mice with the mouse-targeted AAV-B10/CRISPRa. 21 days post-injection, rescue of motor skills and motor coordination will be assessed by rotarod for balance coordination, and the Digigait for temporal and spatial parameters of gait. To examine rescue of cognitive impairments relevant to SYNGAP1, treated mice will undergo a cognitive battery including a Y-maze assay, novel object recognition and the touchscreen assay. To examine rescue of behavioral seizures and EEG spiking using

continuous non-tethered recordings, treated mice will be implanted with a wireless device designed to collect EEG data in free-moving animals in their home cage. To examine sleep rescue, quality by sleep stages and quantity by time spent will be recorded using continuous non-tethered recording. EEG acquired from treated mice will also be analyzed for sleep including frequency and duration of all sleep stages and sleep spindle production. At the conclusion of this functional assessment, brain tissue will be isolated for either molecular or histological assessment to determine molecular mechanism of action for functional reactivation of SYNGAP1.

22q11.2 Deletion Syndrome

Chapter 4 of the dissertation identified 16 strong dCas12a guides for CRISPR activation of 16 genes thought to be associated with the neurological/cognitive phenotype of 22q DS. Immediate next steps for this project involve packaging these guides into a single Cas12a array for multiplexing. The development of a multiplex vector that activates multiple genes and demonstration of phenotypic correction in a mouse model for 22qDS would be the first steps towards developing a viable therapy for the neurological and cognitive phenotypes of 22q DS.

Once multiplex activation is achieved *in vitro*, the next step towards proof-of-concept of a viable therapy is to assess multiplex gene activation and phenotypic correction in a 22q DS murine model *in vivo.* The dCas12a-VPR and validated multiplexed crRNA array can be packaged in AAV-PHP.eB, allowing widespread brain distribution upon systemic injection into wild-type and *Df1/+* mice, a mouse model of the disease.[15] Biodistribution and gene activation can be assessed by RNA-seq, Western

blot, and immunohistochemistry. *Df1/+* mice display many physical and behavioral phenotypes that are analogous to human symptoms.[15] One robust and potentially rescuable phenotype is the prepulse inhibition (PPI) of the acoustic startle response.[15] That being said, PPI is a non-specific phenotype, so better models are needed that allow assessment of the anxiety, autism spectrum disorders, ADHD and other prevalent symptoms involved in 22qDS.

REFERENCES

1. Chan, K. Y. *et al.* Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. *Nature Neuroscience* **20**, 1172–1179 (2017).
2. Fischell, J. M. & Fishman, P. S. A Multifaceted Approach to Optimizing AAV Delivery to the Brain for the Treatment of Neurodegenerative Diseases. *Front Neurosci* **15**, 747726 (2021).
3. Goertsen, D. *et al.* AAV capsid variants with brain-wide transgene expression and decreased liver targeting after intravenous delivery in mouse and marmoset. *Nat Neurosci* **25**, 106–115 (2022).
4. Goertsen, D. *et al.* AAV capsid variants with brain-wide transgene expression and decreased liver targeting after intravenous delivery in mouse and marmoset. *Nat Neurosci* **25**, 106–115 (2022).
5. Hanlon, K. S. *et al.* Selection of an Efficient AAV Vector for Robust CNS Transgene Expression. *Molecular Therapy - Methods and Clinical Development* **15**, 320–332 (2019).
6. Martino, A. T. & Markusic, D. M. Immune Response Mechanisms against AAV Vectors in Animal Models. *Mol Ther Methods Clin Dev* **17**, 198–208 (2020).
7. Rabinowitz, J., Chan, Y. K. & Samulski, R. J. Adeno-associated Virus (AAV) versus Immune Response. *Viruses* **11**, 102 (2019).
8. Frontiers | Human Immune Responses to Adeno-Associated Virus (AAV) Vectors. https://www.frontiersin.org/articles/10.3389/fimmu.2020.00670/full.
9. AAV serotypes and AAV Tissue-specific Tropism | GeneMedi. https://www.genemedi.net/i/aav-serotypes-tissue-specific-tropism.
10. Colella, P., Ronzitti, G. & Mingozzi, F. Emerging Issues in AAV-Mediated In Vivo Gene Therapy. *Mol Ther Methods Clin Dev* **8**, 87–104 (2017).
11. Mukherjee, A. *et al.* Lipid-polymer hybrid nanoparticles as a next-generation drug delivery platform: state of the art, emerging technologies, and perspectives. *Int J Nanomedicine* **14**, 1937–1952 (2019).
12. Haastert, K., Mauritz, C., Chaturvedi, S. & Grothe, C. Human and rat adult Schwann cell cultures: fast and efficient enrichment and highly effective non-viral transfection protocol. *Nat Protoc* **2**, 99–104 (2007).
13. Phospho-ERK (Thr202/Tyr204) cellular HTRF kit | Cisbio. https://www.cisbio.com/phospho-erk-thr202-tyr204-cellular-kit-40565.
14. Cichowski, K. Mouse models of tumor development in neurofibromatosis type 1. *Science* **286**, (1999).
15. Vogel, K. S. *et al.* Mouse tumor model for neurofibromatosis type 1. *Science* **286**, 2176–2179 (1999).
16. Sivagnanasundaram, S. Differential gene expression in the hippocampus of the Df1/+ mice: a model for 22q11.2 deletion syndrome and schizophrenia. *Brain Res.* **1139**, (2007).