CRISPR-CAS9 UTILITY IN GENOME ENGINEERING

Michelle McCullers

Faculty Advisor: Dr. Brian Chadwick

Department of Biological Sciences

Abstract:

The field of genomic engineering and manipulation has made great strides in recent years through developing genome-altering techniques that alleviate disease by flexing control on an epigenetic scale. Facioscapulohumeral muscular dystrophy (FSHD) poses a series of points within its pathophysiology that make it possible to examine the utility of these manipulation techniques. This paper specifically focuses on how three approaches can be applied to stop the expression of the full-length double homeobox 4 DUX4 gene transcript, which is thought to be responsible for the upper body muscular atrophy exhibited in most FSHD cases. With this information, we expect that epigenetics will drive more discoveries in molecular biology, including the purpose of repetitive DNA, the role of epigenetics in disease manifestation, and how to apply new genetic engineering techniques in creative ways.

I. Introduction

Adaptation of CRISPRS

The field of genome engineering has made great strides in recent years, especially due to the utility of the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) system modeled after adaptive bacterial immunity. Using the CRISPR system in modern genome engineering centers around the adaptive defense system of bacteria against invading viruses (Rath et al, 2015). Bacterial genomes contain tandem arrays of sequences that through the process of adaptation or spacer acquisition, are transcribed into short RNAs known as guide RNAs (gRNA) where new spacers are inserted. Unique spacers for each viral infection encountered from previous generations interrupt CRISPR sequences. These gRNAs have homology to viral genomes and if a viral genome enters a bacterium, the gRNA will hybridize to the complementary sequence and begin to express the cas genes mediated by CRISPR, CRISPR-associated protein 9 (Cas9) endonuclease then introduces double strand breaks into the viral genome, rendering it unable to cause damage to the bacterium.

Although CRISPR was initially discovered in the 1980s, work in exploring its possible uses for genome editing has become popular since

2012 when it was first shown to be effective in human cell lines (Ran et al., 2013). Researchers have found that co-expression of wild-type Cas9 and gRNA to a genomic region of interest in human cells can direct Cas9 to introduce a double strand break in the target DNA. Upon the formation of a double-stranded break, repair mechanisms like non-homologous end joining (NHEJ) are used to re-anneal the broken ends, resulting in errors in the now-disrupted genomic target. This is one way to use CRISPR to edit targeted genes.

To increase the versatility of Cas9, it is possible to introduce a catalytically inactive form of Cas9 that no longer possesses the ability to cut DNA by mutating key amino acid residues in the two endonuclease active sites of the enzyme. This does not impact the ability of Cas9 to be recruited specifically to the gRNA target. This "dead" version, dubbed dCas9, gives two additional modes to utilize CRISPR: activation or suppression. It is possible to fuse a functional part of a protein sequence (a protein domain), of interest (an activation or suppression domain) to the dCas9 and direct this to the target sequence. The addition of a protein domain to the CRISPR-dCas9 complex allows for further control of the target region without directly disrupting the sequence.

> " The term epigenetics means 'above the genome,' which refers to key level of control that is not pure sequence. "

The term epigenetics means "above the genome," which refers to key level of control that is not pure sequence. Rather, it is a biochemical alteration made to the sequence. The most common epigenetic markers include DNA methylation and histone modifications, which dramatically alter gene expression as well as overall regulation. The use of suitable gRNAs, along with dCas9 tagged with such effector domains, helps in epigenetic regulation of target regions.

Facioscapulohumeral muscular dystrophy

Facioscapulohumeral Muscular Dystrophy (FSHD), is a genetic disease characterized by upper body muscle atrophy in early adulthood (Tawil, 1998; Lemmers et al., 2010). Methodologies of eliminating the genetic causes of this disease are of primary importance, both in terms of alleviating suffering caused by the muscle atrophy, and for determining how we can optimize CRISPR at various points in a given disease pathway. There are two different forms of the disease (FSHD1 and FSHD2) to be discussed in the coming section.

The physiological cause for the weakening of the muscles in other

CRISPR-CAS9 UTILITY IN GENOME ENGINEERING

common forms of muscular dystrophy is abnormal or even absent dystrophin proteins (Emery, 2008). Dystrophin facilitates the maintenance of the fibrous membrane in muscle cells. Without a supportive membrane, muscle fibers are biochemically altered in a way that reduces their ATP output and causes them to eventually die and become scar tissue. Over time, the scar tissue provides further complications within the muscle's ability to move properly, and leads to severe cases in the most pronounced patients.

In FSHD, however, the muscular atrophy is caused by the inappropriate expression of the double homeobox transcription factor DUX4. DUX4 expression interrupts muscle repair and replenishment by signaling a cascade of transcription factors to inappropriately activate endogenous retroviruses (ERV elements) (Geng et al., 2012). FSHD, an autosomal dominant disease, occurs in 1 out of 15,000 live births, making it the third most common form of muscular dystrophy (Flanigan et al., 2001). The onset of FSHD primarily occurs in early adulthood and increases in pathophysiology
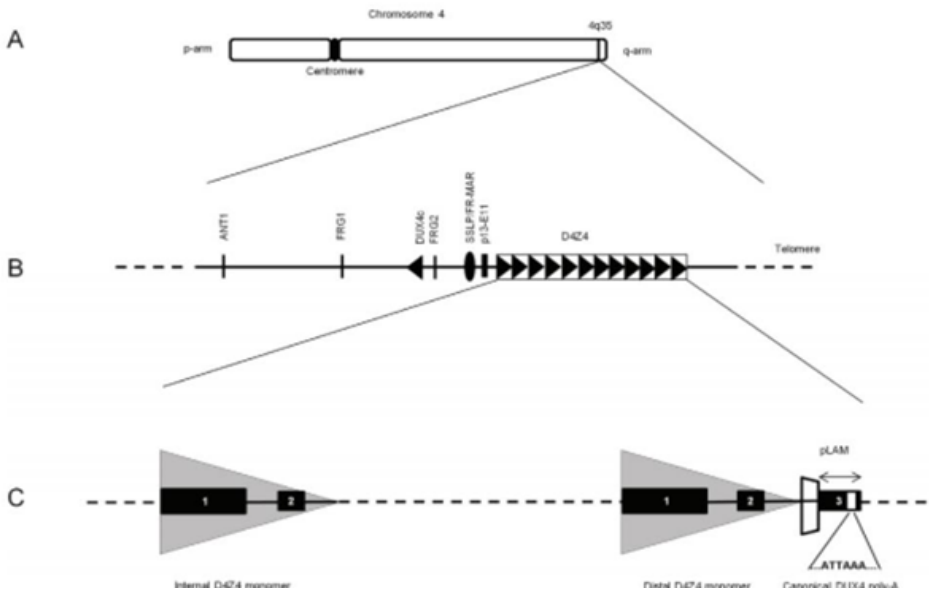

Figure 1 (Das, 2015): Chromosome 4q35 and its components, which houses macrosatellite D4Z4. (A) The entire chromosome 4 with the subtelomeric region separated next to D4Z4. (B) Regulatory components flanking either side of the D4Z4 macrosatellite. (C) The individual monomers of D4Z4, where unaffected people have 11-150 copies, and affected people have 1-10 monomer copies. The last D4Z4 repeat adjacent to the telomere is unique since it has a pLAM region next to the poly-A sequence that can be disrupted as a possible avenue of DUX4 repression.

MCCULLERS

as time goes on and normal muscle repair fails to occur.

There are two common forms of FSHD, both of which are associated with chromatin changes at the macrosatellite repeat D4Z4. D4Z4 is a tandem repeat composed of a variable number of 3.3 kb GC-rich repeat units located on chromosome 4q35 (Fig. 1A) along with a highly homologous, yet non-pathogenic pair of arrays on 10q26 (Chadwick, 2008). Notably, each repeat unit contains the full DUX4 open reading frame. FSHD1 is caused by contraction in the size of the D4Z4 tandem repeat specifically on chromosome 4 due to a reduction in the number of repeat units within the satellite array (Ottaviani, et al., 2009). Unaffected individuals have 11-150 copies of the repeat, while affected individuals only have 1-10 copies of the repeat. However, at least one copy of the repeat is required for disease manifestation (Tupler et al., 1996). This poses a unique disease model where genomic engineering techniques can be applied to gain the best understanding of how the disease works on an epigenetic scale. Severity of the disease is also influenced by the disease-associated contraction size, with disease severity inversely related to repeat copy number, meaning that FSHD1 is "dosage dependent", where patients with only one copy of the repeat are expected to present with more severe signs of atrophy than patients with ten copies of the repeat.

FSHD2 is caused by mutations in the protein encoding gene Structural Maintenance Of Chromosomes Flexible Hinge Domain Containing 1 (SMCHD1) located on chromosome 18, resulting in chromatin relaxation at D4Z4 (since a decrease in SMCHD leads to inappropriate increase in the formation of euchromatin, the open relaxed state (Lemmers et al., 2010). Both forms of the disease result in the activation of DUX4 that is normally only expressed in testes (Snider et al., 2010; Lemmers et al., 2010; Lemmers et al., 2012). DUX4 activates various downstream target genes that result in muscle atrophy (Geng et al., 2010, Zhao et al., 2014). Since the disease is caused by the inappropriate expression of DUX4, there are a number of methods to silence it.

In healthy people, SMCHD1 is maintained in a methylated or closed off state, repressing DUX4-fl. (B) In both types of FSHD 1 and 2 patients, SMCHD1 is altered to allow DUX4 expression. In FSHD1, there is contraction in one allele that results in demethylation of CpG islands that decrease SMCHD1 levels, and in FSHD2 there is a mutation in the SMCHD1 gene that leads to loss of methylation.

Targeting Options to Decrease DUX4 activity

Biological Sciences

CRISPR-CAS9 UTILITY IN GENOME ENGINEERING

With the end goal of suppressing the full-length DUX4 transcript, three approaches in genome engineering emerge: (1) targeting the poly-A tail, (2) increasing SMCHD1 expression, or (3) nucleating the heterochromatin marker Histone 3 Lysine 9 trimethylation (H3K9me3). The approaches come into play at various points in the pathophysiology of FSHD in terms of transcription, regulation, and access, respectively. Each approach is examined individually and their effectiveness is addressed later on in the paper.
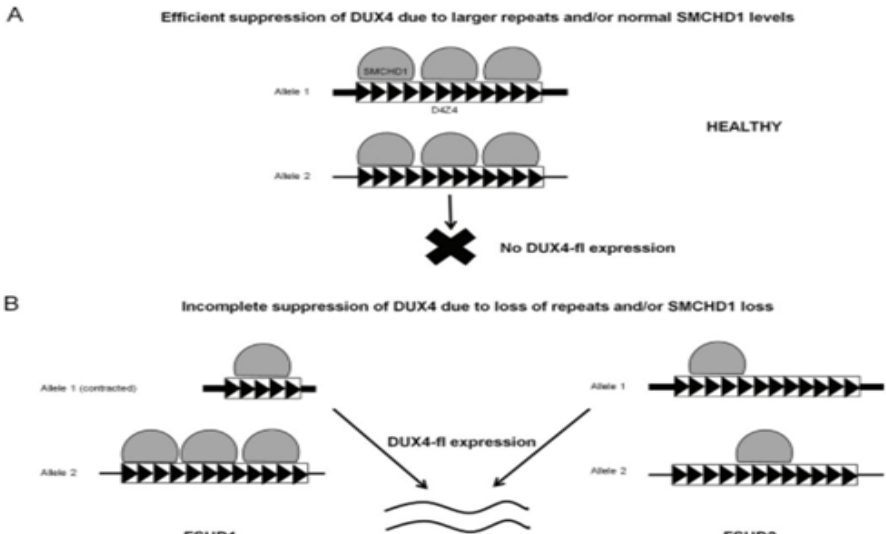

Figure 2 (Das, 2015): The interplay between SMCHD1 and DUX4 full length expression. (A) In healthy people, SMCHD1 is maintained in a methylated or closed off state, repressing DUX4-fl. (B) In both types of FSHD 1 and 2 patients, SMCHD1 is altered to allow DUX4 expression. In FSHD1, there is contraction in one allele that results in demethylation of CpG islands that decrease SMCHD1 levels, and in FSHD2 there is a mutation in the SMCHD1 gene that leads to loss of methylation.

Targeting the polyadenylation signal

At the end of every transcribed mRNA, there is an added polyadenylation (poly-A) signal that stabilizes the transcript for proper translation into a final protein product. By disrupting the polyadenylation signal in genomic DNA, a poly-A tail fails to be added to the end of the DUX4 transcript, dramatically reducing the stability of the RNA such that it is degraded before it can be translated. This is the most direct way to interrupt the disease pathway, since it involves targeting the transcript itself. In FSHD, the DUX4-fl poly-A signal is associated with the pLAM region of the gene locus, which is located immediately distal to the D4Z4 repeats (Richards,

MCCULLERS

et. al, 2012). Graduate student Sunny Das has spearheaded this approach through using Cas9 to disrupt the poly-A sequence using gRNAs that flank the poly-A signal and by simultaneous cutting, can excise the intervening DNA, including the poly -A signal.

Upregulating SMCHD1

Another key epigenetic player is SMCHD1. SMCHD1 is a gene product that is responsible for chromosomal structural maintenance by in part maintaining DNA hypermethylation, a function that is compromised when protein levels are reduced due to mutation of one allele. Therefore, it is possible that increasing SMCHD1 levels may compensate for the reduced protein levels in FSDH2. I aim to increase SMCHD1 levels by targeting the transcriptional activator domain VP64 to the SMCHD1 promoter using gRNAs designed to the promoter in combination with a dCas9-VP64 fusion (Fig. 3).
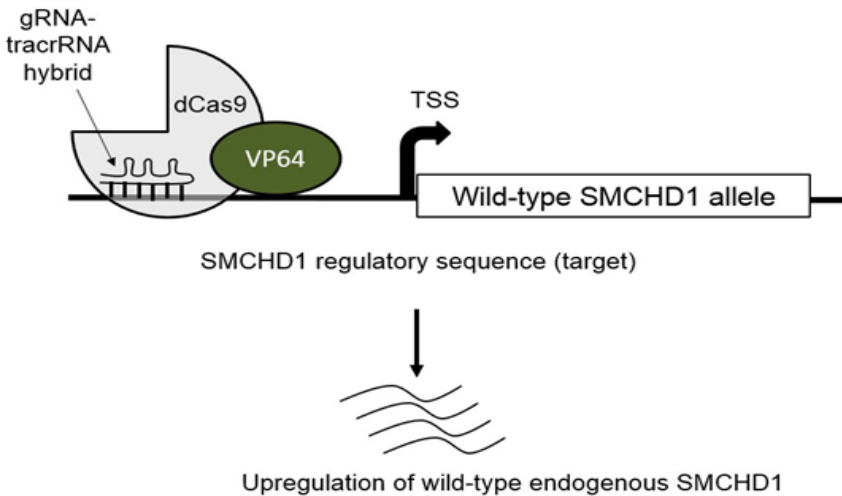

Figure 3 (Das, 2015): Schematic of how binding a transcriptional activator such as VP64 can achieve increased SMCHD1 levels.

VP64 is a protein that attaches to the dCas9 targeted for a gRNA and serves as a transcriptional activator. It was developed by fusing to-gether four copies of the transcriptional activation domain (VP16) from the herpes virus, and attaching it to CRISPR-dCas9 for targeting purposes (Tanenbaum, et al., 2014). In molecular biology, VP64 is commonly used to

68

CRISPR-CAS9 UTILITY IN GENOME ENGINEERING

attach to a DNA binding domain to upregulate transcription of a targeted gene, in this case, SMCHD1.

VP64 is versatile because it can also be adapted for the relatively new Sun Tagging system. SunTag, named for its great amplification capabilities and dubbed "SUperNova Tag," is an epitope tagging system of up to 24 peptide copies for which a desired tag, like green florescent protein (GFP) or transcriptional activator like VP64, can be attached (Fig. 4). In a recent study focusing on FSHD utilizing SunTag to target dCas9-KRAB to exon 1 of DUX4, the expression of full length DUX4 was reduced (Himeda, 2015). By recruiting 24 copies of VP64 to the SunTag epitope anchored to dCas9 at the SMCHD1 promoter, a large increase in expression (and therefore DUX4 repression) can be expected.
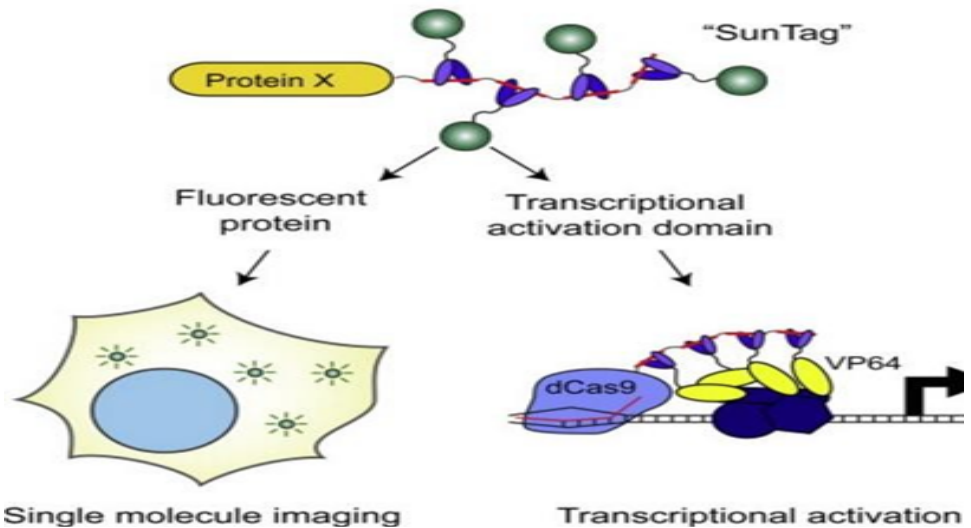

Figure 4 (Perkel, 2014): The SunTag epitope can serve as a way to amplify GFP for fluorescent visualization or can add many transcription factors like VP64 to upregulate a chosen gene.

Nucleating heterochromatin formation

Lastly, focusing on the formation of heterochromatin may be a good strategy for a number of reasons. Heterochromatin and euchromatin are ways in which the DNA can be packaged in a cell's nucleus. In somatic cells, euchromatin allows enzymes like helicase and RNA polymerase to access the sequence for transcription and replication, while heterochromatin is pack-

aged in such a way to make it less accessible transcription factors and other DNA binding factors. Therefore, recruiting proteins that can promote heterochromatin formation at D4Z4 should be able to block or reduce DUX4 expression (Volpe, et al., 2002; Zeng, et al., 2009). Directing dCas9-KRAB to the DUX4 gene should therefore re-establish gene silencing and reverse the negative impact of DUX4 expression. Directive guide RNAs (gRNAs) would recruit a "dead" Cas9 (dCas9) protein that binds to a regulatory region (eg. promoter) in D4Z4, where the Krüppel associated box (KRAB) domain fuses to dCas9, which will recruit the protein Tripartite motif-containing 28 (TRIM28), which in turn will recruit the Histone-lysine N-methyltransferase (SETDB1) that will ultimately lay down the heterochromatin mark H3K9me3 at and around D4Z4. Modifications to the DUX4 target sites may result in sustained heterochromatin that is maintained throughout multiple somatic cell divisions.
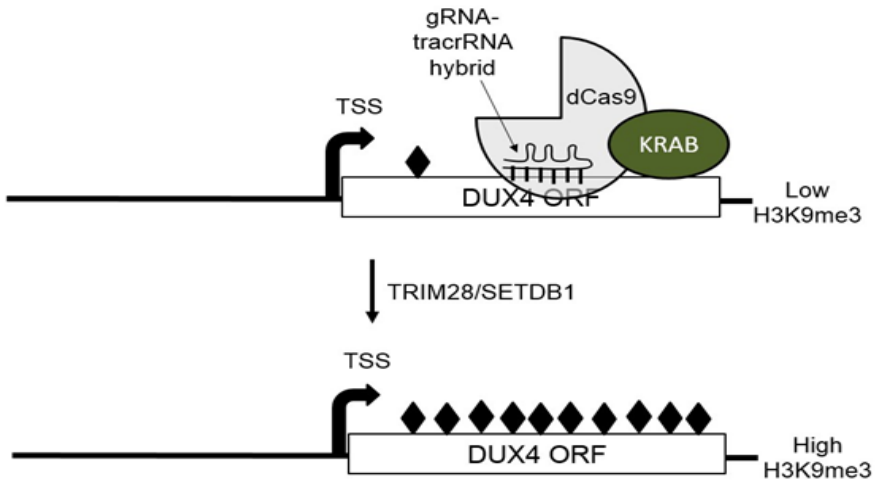

Figure 5 (Das, 2015): Schematic of using KRAB to recruit factors that bring about heterochromatin formation around D4Z4.

II. Purpose

The goal of this thesis is to demonstrate the utility of the CRISPR system in genome engineering and manipulation by examining its applications to FSHD. There are a number of ways to use CRISPR, involving either Cas9 for editing purposes, or a dCas9 to alter regulation to either activate or suppress a target gene (Ran et al., 2013). In the context of FSHD where the goal is to silence the DUX4 transcript, the three parallel routes of

disrupting the poly-A tail, boosting SMCHD1 transcription, or nucleating H3K9me3 are each considered. The primary focus of this project is to use a KRAB domain to assist in the formation of heterochromatin to silence DUX4. The chromatin structure will be evaluated using a D4Z4-specific Chromatin Immunoprecipitation (ChIP) analysis for H3K9 methylation (heterochromatin marker) and for H3K4me2 (euchromatin marker). The decreased expression of DUX4 can be confirmed by conducting qRTPCR. This approach is expected to promote heterochromatin nucleation at D4Z4. Successful demonstration that these approaches can manipulate DUX4 expression with FSHD will open up the possibility of applying these methods as therapeutic approaches in patients. At the cellular basis of future genetic therapeutics, it is crucial to be able to target disease-contributing regions of the genome. The three parallel methods described above provide a starting point for comparing approaches individually.

III. Materials and Methods

Cell culture

HCT116, a near diploid male colon carcinoma cell line, was obtained from the American Type Culture Collection (ATCC; No. CCL-247) Type Culture Collection (ATCC; No. CCL-247) (www.atcc.org), as was human embryonic kidney 293 cells (CRL-1573). DKO was obtained from Dr. Bert Vogelstein's laboratory at Johns HopkinsUniversity School of Medicine.

Table 1: Primers

Category	Primer Name	Primer Sequence (5' to 3')	Product Size
DUX4-fl qRT-PCR Exons 2-3	DUX4-cDNA-F10	ACCGCGGAGAACTGC-CATTC	159 bp
	DUX4-cDNA-R4	GACATTCAGC-CAGAATTTCACG	
SMCHD1 qRT-PCR	qSMCHD1-F1	TCCGGATATGAG-GAAGAAAAAG	435 bp
	qSMCHD1-R1	TGTCGTCT-CAACCTTTGGTG	
GAPDH Control qRT-PCR	GAPDH-q-Fwd	CCCAATACGAC-CAAATCCGT	119 bp
	GAPDH-q-Rev	TCTCTGCTCCTCCT-GTTCGA	

MCCULLERS

Oligonucleotides

    All oligonucleotides were synthesized using the service of Eurofins Genomics. Oligonucleotide sequences and their applications are listed above in in Table 1. Within the exons of DUX4, publically available programs were used to identify those targets with the greatest potential for success. After having found the targets, gRNA oligos were designed using web tools available through Addgene and these gRNAs will be cloned into the plasmids optimized for experiments in mammalian systems.

Designing gRNAs

    First, we identified targets for designing a number of unique gRNAs to our regions of interest. We used publically available genome sequence data from the UCSC Genome browser; and the computer sequence alignment software Sequencher (Gene Codes Corporation) to design the gRNAs. For suppression of DUX4 by dCas9-KRAB, we designed 3 gRNAs to Exon 3

Table 2: gRNA oligos and vectors, with the SunTag vector from AddGene being "dCas9-24xGCN4" (originally "pHRdSV40-NLS-dCas9-24xGCN4 v4-NLS-P2A-BFP-dWPRE") and the antibody-GCN vector being "Anti-GCN4-VP64" (originally "pHRdSV40-scFv-GCN4-sfGFP-VP64-GB1-NLS").

Category	gRNA Name	gRNA Sequence (5' to 3')	Vector used for cloning gRNA	Vector used for cloning dCas9-effector
dCas9/KRA B mediated DUX4 suppression	KRAB- CR2A-T	CACCGCACCCCGGCT-GACGTGCA	pLenti-gRNA	pHR- SFFV- dCas9- BFP-KRAB
	KRAB-CR2A-B	AAACTGCACGTCAGC-CGGGGTGC		
	KRAB-pA- CR5A-T	CACCGTTCTTCCGT-GAAATTC	pLenti-gRNA	
	KRAB-pA- CR5A-B	AAACGAATTTCACG-GAAGAAC		
	KRAB-pA- CR11-T	CACCGTAGACT-GAACCTAGAGAA	pLenti-gRNA	
	KRAB- pA-CR11-B	AAACTTCTCTAG-GTTCAGTCTAC		

CRISPR-CAS9 UTILITY IN GENOME ENGINEERING

dCas9/VP6 4 mediated SMCHD1 activation	SMCHD 1-CR-3-T	CACCGCTCAGTC-GGGATCCTGGA G	pLenti-gRNA	pMLM370 5 (for regular VP64)
	SMCHD 1-CR-3- B	AAACCTCCAG-GATCCCGACTGAG C		
	SMCHD 1-CR-4-T	CACCGGGATCCTG-GAGAGGCGG GA	pLenti-gRNA	
	SMCHD 1-CR-4- B	AAACTCCCGCCTCTC-CAGGATCC C		
	SMCHD 1-CR-5-T	CACCGGAGAGGC-GGGAAGGCAG CG	pLenti-gRNA	
	SMCHD 1-CR-5- B	AAACCGCT-GCCTTCCCGCCTCTC C		
	SMCHD 1-CR-6-T	CACCGGCCTCTGAG-GACTACCCG C	pLenti-gRNA	
	SMCHD 1-CR-6- B	AAACGCGGG-TAGTCCTCAGAGGC C		
	SMCHD 1-CR-7-T	CACCGCCTCTGAG-GACTACCCGC A	pLenti-gRNA	
	SMCHD 1-CR-7- B	AAACTGCGGG-TAGTCCTCAGAGG C		
	SMCHD 1-CR-9-T	CACCGCCTCCTTC-CAACTTCGCG A	pLenti-gRNA	
	SMCHD 1-CR-9- B	AAACTCGCGAAGTTG-GAAGGAGG C		
	SMCHD 1-CR-15- T	CACCGGGCAGGAGC-GCGTTTGAA T	pLenti-gRNA	
	SMCHD 1-CR-15- B	AAACATTCAAACGC-GCTCCTGCC C		

SMCHD 1-CR-16- B	AAACGGGAACC-GATTCAAACGCG C		
SMCHD 1-CR-19- T	CACCGCCCC-GGGAGCTGGAGCT GA	pLenti-gRNA	
SMCHD 1-CR-19- B	AAACTCAGCTC-CAGCTCCCGGGG C		
SMCHD 1-CR-20- T	CACCGCGCACCT-CAGCCCTGAGC C	pLenti-gRNA	
SMCHD 1-CR-20- B	AAACGGCTCAGGGCT-GAGGTGCG C		
SMCHD 1-CR-22- T	CACCGCGGGCCCC-GGGCGCGCG CG	pLenti-gRNA	
SMCHD 1-CR-22- B	AAACCGCGCGCGCCC-GGGGCCC GC		

of DUX4, surrounding the poly-A site. For SMCHD1 upregulation, we designed 11 gRNAs in a 500 bp region of the SMCHD1 promoter. These gRNAs were designed using already established criteria to maximize targeting success (Ran et al., 2013) and synthesized by Eurofins Genomics (Table 2).

Cloning and expression of gRNAs

Top and bottom oligonucleotides constituting gRNAs were cloned into suitable plasmids available through Addgene (Table 2) that are optimized for expression in mammalian systems. Post-cloning sequence analysis (Eurofins Genomics) verified the presence and orientation of the correct gRNAs. The gRNA oligo plasmids, along with dCas9-effector plasmids, were finally transfected into established model cell lines. The cell lines were DKO, which is a DNA methyltransferase double knockout of HCT116 and is able to express DUX4 transcripts constitutively, and 293T cells. Transfection was carried out using the Lipofectamine 3000 reagent (Invitrogen), using standard protocols. For all experiments, 72 hours post-transfection (except for first attempt with SMCHD1 gRNAs where cells were harvested at 48 hours, 72 hours and 96 hours), cells were pelleted and assessed for DUX4 or SMCHD1 expression using qRT-PCR.

Isolation of RNA and preparation of cDNA

Biological Sciences

CRISPR-CAS9 UTILITY IN GENOME ENGINEERING

Total RNA was isolated from cells using the NucleoSpin RNA II kit (Machery-Nagel, Bethlehem, PA). First-strand cDNA was prepared from equal amounts of starting RNA (2ug total RNA) with random hexamers with and without M-MuLV reverse transcriptase (RT) according to the manufacturer's instructions (NEB). cDNAs prepared with and without RT were used as templates for both qualitative and quantitative PCR. cDNA was diluted to half its original concentration prior to qualitative RT-PCR and one-sixth of its original concentration prior to qRT- PCR. Both RT-PCR and qRT-PCR reactions were carried out with equal starting amounts of cDNA for each sample.

qRT-PCR for DUX4 expression analysis

qRT-PCR for DUX4 transcripts was carried out using two independent sets of primers. While a combination of DUX4-UTR-Fwd and DUX4-A-Rev amplified any transcript containing the 3' end of Exon 1 of DUX4, the combination of DUX4-cDNA-F10 and DUX4-cDNA-R4 amplified transcripts containing Exons 2 and 3. qGAPDH-Fwd and qGAPDH-Rev were used as a control for normalization during analysis of results. qRT-PCR was performed with the same reagents and conditions as for qChIP. All oligonucleotides were synthesized using the services of Eurofins MWG Operon. DUX4 expression for each sample was normalized to GAPDH expression and displayed using the ΔΔCt method. Triplicates were used for each sample.

IV. Results

Targeting the polyadenylation signal

In the approach to disrupt the poly-A signal, using the wild type Cas9 vector pX459 with gRNAs CR-2A and CR-11 introduced into HCT116 cells resulted in simultaneous cutting, deleting a 148 bp sequence encompassing the poly-A site as determined by qRT-PCR. For post-verification of the efficiency of targeting, single-cell clones of successfully targeted HCT116 cells were isolated and are currently being assayed for repression of DUX4-fl expression.compassing the poly-A site as determined by qRT-PCR. For post-verification of the efficiency of targeting, single-cell clones of successfully targeted HCT116 cells were isolated and are currently being assayed for repression of DUX4-fl expression.

MCCULLERS

Upregulating SMCHD1

     In the upregulating SMCHD1 approach with VP64, 12 different gRNAs designed to the SMCHD1 promoter were pooled and tested together, and the level of transcription of SMCHD1 was assessed at varying times: 48 hours post-transfection, 72 hours post-transfection, and 96 hours post-transfection. Figure 7 depicts the transcription levels of SMCHD1 against time.
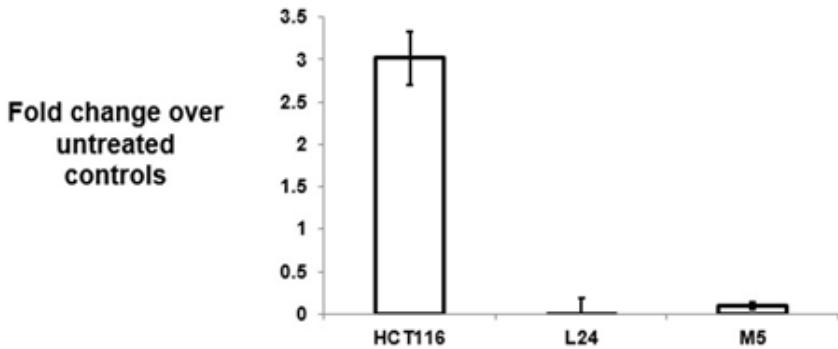

Figure 6 (Das, 2015). Sunny's preliminary results indicate success in decreasing DUX4-fl by cutting out with CR-2A and CR-11. Assay for DUX4-fl expression following transfection with pX459 vector with gRNAs CR-2A and CR-11.
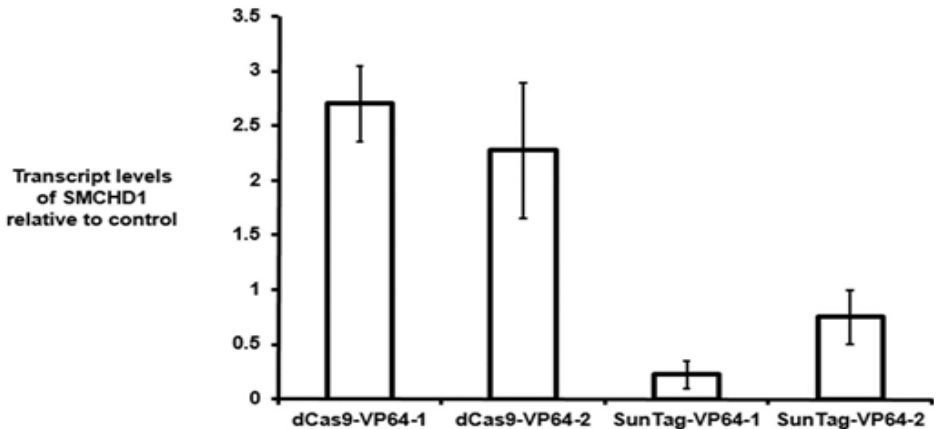
Figure 7 (Das, 2015). All 11 different gRNAs designed to the SMCHD1 promoter were pooled and tested together, cells were harvested and SMCHD1 transcript levels were assessed. Compared to levels of SMCHD1 in Mock (no gRNA) 293T cells, there was an appreciable difference in the levels of SMCHD1

Nucleating Heterochromatin Formation

     In the nucleation of heterochromatin approach, initial testing of three different gRNAs to Exon 3 (CR-2A, 5A and 11) revealed that CR2A

CRISPR-CAS9 UTILITY IN GENOME ENGINEERING

and 5A but not CR-11 alone, resulted in appreciable (>50%) DUX4-fl repression (Fig. 8A). Repeating this experiment without CR-11 yielded similar results, showing that CR-5A alone can repress transcript levels by up to 85% (Fig. 8B). Initial qRT-PCR results show that the use of dCas9-KRAB in conjunction with both gRNAs CR-2A and CR-5A alone resulted in a decrease in transcript levels of DUX4-fl compared to baseline. Conversely, CR-11 alone resulted in normal to slightly increased expression of DUX4-fl, suggesting that it is non-viable as a gRNA in this case. The decrease in transcription is not as pronounced when all three gRNAs are combined, which may be due to the presence of CR-11. Transcript levels are normalized to DUX4-fl transcription in DKO-Mock (no gRNA) whose expression value is 1.
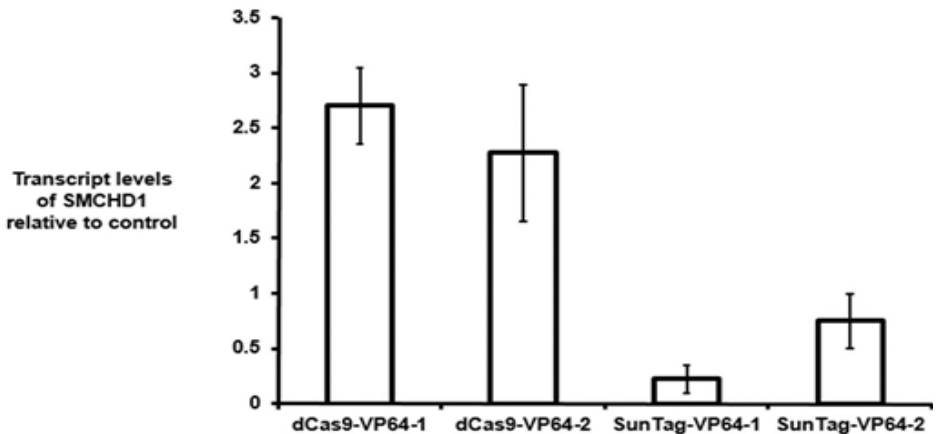

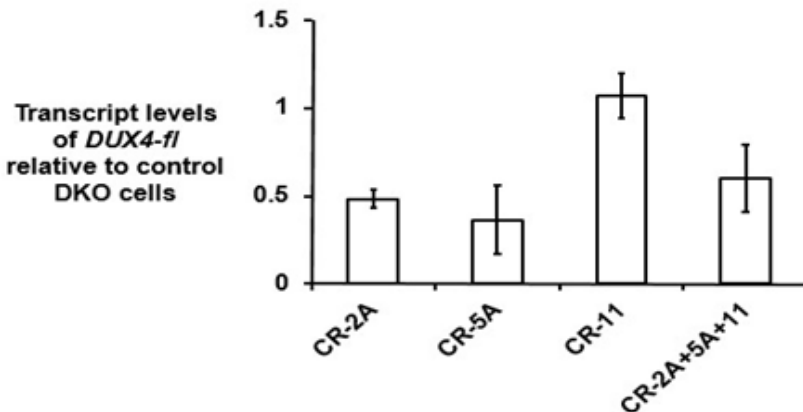*Figure 8 A (Das 2015): qRT-PCR results for the same transfection repeated only with CR-2A and CR-5A.*

Figure 8 B (Das, 2015): Expression trends are similar to that in Figure 8 A.

MCCULLERS

Biological Sciences

VI. Discussion

The normal VP64 worked as anticipated, but the SunTag results did not reflect the large improvement that was expected. In the SunTagging system, two vectors were used instead of a single vector like a normal activator system, so it is possible that the transfection of the second vector was unsuccessful, accounting for the unexpected decrease in transcript levels of SMCHD1 for SunTag. Another possibility is the sheer bulk of the protein (i.e. 24 copies of VP64) may have inhibited proper access to the SMCHD1 promoter.

Given the fact the DUX4-fl transcript is produced by Exons 1 and 2 splicing into Exon 3, it is possible that recruitment of KRAB by dCas9 at Exon 3 results in increased H3K9me3 at the Exon 3 splice site, thereby inhibiting this splicing event. Previous research has suggested that altered H3K9me3 levels might be crucial in governing splice site choice in DUX4-fl expression (Snider at al., 2010). Nevertheless, seeing substantial DUX4-fl repression in pooled cells was encouraging. The next step would be to repeat the same experiment, possibly with just CR-5A to see if repression can be achieved in a clinically relevant cell type such as DUX4-fl expressing FSHD myoblast.

In the progress made thus far, each approach (i.e. disrupting the polyadenylation signal, upregulating SMCHD1, and nucleating heterochromatin formation) demonstrates various levels of efficiency in accomplishing the goal of repressing the DUX4-fl transcript. The primary limitation is whether the therapies would stick throughout the many series of nuclear recplcaton to come. Especially in vivo where muscle cells are multinucleated, one nucleus could be affected and may slowly spread to the others through the inappropriate expression of ERV elements. The next step towards forming a model to combat FSHD would be to apply one of these approaches to human myoblasts to ensure that the changes are lasting and effective. Once this is demonstrated to work in human myoblasts, in vivo gene therapy approaches may ultimately be applied in patients.

Developing such a model for FSHD is important as a guide for future medical applications of epigenetic techniques, both those that have emerged in the past five years and those that are still coming at more progressive

CRISPR-CAS9 UTILITY IN GENOME ENGINEERING

rates. This project demonstrates that there is much flexibility when working with genetic diseases, since there are many steps in the disease pathway that pose possibilities as targets for gene therapy. Each of the three approaches discussed here takes an in-depth look into a given step, so the different approaches can be compared and the right one can be chosen based on the patient's genome. The unique aspect of DUX4 being a retrogene can allow scientists to apply reverse engineering techniques to shed light on how to tackle other genetic diseases, and how therapies can be tailored within the disease pathway to be most effective for each induvial patient.. With developments in gene therapy and an expanded list of possible ways to manipulate disease, the future of molecular biology remains promising.

VII. Conclusion

FSHD is an interesting disease to study because of the nature of this myopathy. This work and other research focusing on D4Z4's role throughout the genome suggest that such repetitive DNA is anything but "junk," as it is often portrayed in the current literature. The nature of this myopathy allows for a variety of strategies for treatment on a molecular level, such as the role of D4Z4, and each strategy is a glance at current achievements in the development of genetic engineering techniques. Clearly more research needs to be done in order to confirm which of the approaches will prove best for repressing the DUX4-fl expression implicated in FSHD pathology.

VIII. Acknowledgements

MCCULLERS

1. Rath, D., Amlinger, L., Rath, A., & Lundgren, M. (2015). The CRISPR-Cas immune system: Biology, mechanisms and applications. Biochimie, 117, 119-128.

2. Ran, F., et al. (2013). Double Nicking by RNA-Guided CRISPR Cas9 for EnhancedGenome Editing Specificity. Cell, 154, 1380-1389.

3. Emery, A. (2008). Muscular Dystrophy: The Facts. New York: Oxford University Press, Inc.

4. Chadwick, B.P. (2008). DXZ4 chromatin adopts an opposing conformation to that of the surrounding chromosome and acquires a novel inactive X-specific role involving CTCF and antisense transcripts. Genome Research, 18, 1259-1269.

5. Richards, M., Coppee, F., Thomas, N., Belayew, A., & Upadhyaya, M. (2012).Facioscapulohumeral muscular dystrophy (FSHD): an enigma unraveled? Human Genetics, 131, 325-340.

6. Ottaviani, Alexandre et al. (2009). The D4Z4 Macrosatellite Repeat Acts as a CTCF and A-Type Lamins-Dependent Insulator in Facio-Scapulo-Humeral Dystrophy. PLOS Genetics 5.2, 1-9.

7. Das, S., and Chadwick B.P. (2015). Epigenetics: Current Research and Emerging Trends.United Kingdom: Caister Academic Press, 217-254.

8. Tanenbaum, M., Gillbert, L., Qi, L., Weissman, J., & Vale, R. (2014). A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. Cell, 159,635-646.

9. Perkel, J. (2014). Here Comes the SunTag. Retrieved from http://www.biotechniques.com/news/Here-Comes-the-SunTag/biotechniques-355351.html#.VkNvZ1QrKUk

10. Himeda, C., Jones T., & Jones, P. (2015). CRISPR/dCas9-mediated transcriptional inhibition ameliorates the epigenetic dysregulation at D4Z4 and represses DUX4-fl in FSH muscular dystrophy. Molecular Therapy accepted article preview online 03 November 2015; doi:10.1038/mt.2015.200

11. Volpe, Thomas et al. (2002). Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi. Science, 297.13, 1833-1837.

12. Zeng, Weihua et al. "Specific Loss of Histone H3 Lysine 9 Trimethylation and HP1γ/Cohesin Binding at D4Z4 Repeats Is Associated with Facioscapulohumeral Dystrophy (FSHD)." PLOS Genetics 5.7 (2009): 1-14. Online.