Master's thesis

Activation of *KCNQ1* expression in HEK293 cells using inducible CRISPR-dCas9-VPR tripartite transcriptional activator domain

Ema Kuncheva

Master's Program in Genetics and Molecular Biosciences

Faculty of Biology and Environmental Sciences

University of Helsinki

April, 2022

Abbreviations

Cas	CRISPR associated protein
cDNA	complementary DNA
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRISPRa	CRISPR activation
CRISPRi	CRISPR interference
crRNA	CRISPR RNA
dCas9	Catalytically dead Cas9
Dox	Doxycycline
EDTA	Ethylenediaminetetraacetic acid
gDNA	genomic DNA
gRNA	guide RNA
HEK293	Human embryonic kidney cells 293
kbp	kilobase pairs
PAM	Protospacer adjacent motif
PB	PiggyBac
PCR	Polymerase Chain Reaction
RT-qPCR	Quantitative reverse transcription polymerase chain reaction
SAM	Synergistic Activation Mediator
sgRNA	single guide RNA
TAE	Tris-acetate-EDTA buffer
Tet	Tetracycline
Tm	Primer melting temperature
tracrRNA	Trans-activating CRISPR RNA
TRE	Tetracycline Response Element
VPR	VP64-p65-Rta activation system

Abstract

Tiedekunta – Fakultet – Faculty Biology and Environmental Sciences	Koulutusohjelma – Utbildningsprogram – Degree Programme Genetics and Molecular Biosciences	
Tekijä – Författare – Author Ema Kuncheva		
Työn nimi – Arbetets titel – Title Activation of *KCNQ1* expression in HEK293 cells using inducible CRISPR-dCas9-VPR tripartite transcriptional activator domain		
Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track Cell and developmental biology		
Työn laji – Arbetets art – Level Master's thesis	Aika – Datum – Month and year April, 2022	Sivumäärä – Sidoantal – Number of pages 42

Tiivistelmä – Referat – Abstract

Mutations in the *KCNQ1* gene have been implicated in the onset of hypopituitarism. Regulating *KCNQ1* expression would therefore enable future clinical research on the mechanism of the disease. CRISPR offers a flexible toolset for controlling genetic expression via knockout, knock-in, knockdown, and gene activation. Of these approaches, CRISPR activation (CRISPRa) is distinguished by its ability to induce gene overexpression in a cell's native context, making it a valuable tool in the interrogation of genetic disorder pathogenesis. This thesis therefore tested the efficacy of a CRISPRa subsystem in increasing *KCNQ1* expression.

The CRISPRa subsystem, VPR, was chosen because of its high activation efficiency and the ease of controlling the activation system of its doxycycline-inducible mode of action. The cell line used for the experiment, HEK293, was similarly chosen because of its ease of culture and transfection. To validate the proper functioning of the activation system, expression rates of the related genes *ASCL1* and *GHRH* were measured as positive controls. The activation system successfully upregulated the expression rates of the two genes. As the dCas9-VPR system is dependent on the Tet-ON operator for inducing activation in a controllable manner, a test for dCas9 leakage was conducted. RT-qPCR analysis showed the upregulation of *ASCL1* expression in the uninduced state of the system, confirming the presence of dCas9-VPR leakage. The dCas9-VPR system finally aimed to test the expression rate of *KCNQ1*. Although one novel guide RNA successfully upregulated *KCNQ1* expression, it did so inefficiently and its success was not shared by any of the other tested guide RNAs.

Altogether, the dCas9-VPR system was successfully established in HEK293 cells, and the leakage of the inducible system was confirmed, however, *KCNQ1* activation by CRISPRa requires further optimization.

Avainsanat – Nyckelord – Keywords KCNQ1, CRISPR/Cas, CRISPRa, dCas9-VPR, HEK293, Human embryonic kidney 293 cells
Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors Taneli Raivio and Yafei Wang
Säilytyspaikka – Förvaringsställe – Where deposited HELDA - Helsingin yliopiston digitaalinen arkisto / HELDA - Helsingfors universitets digitala publikationsarkiv /HELDA - Digital Repository of the University of Helsinki
Muita tietoja – Övriga uppgifter – Additional information

Introduction

KCNQ1 function and disorders

The Potassium Voltage-Gated Channel Subfamily Q Member 1 (*KCNQ1*) gene encodes the α-subunit of the Kv7.1 voltage channel, which is prominently expressed in both cardiac and epithelial tissues (Jespersen et al., 2005). Besides being responsible for the repolarization phase of the cardiac action potential (Barro-Soria et al., 2014; Wu and Larsson, 2020), KCNQ1 channels fulfill a myriad of purposes, including the regulation of water and salt transport in the epithelium (Jespersen et al., 2005) and the maintenance of the endocochlear potential (Nakajo and Kubo, 2011). Mutations in *KCNQ1* have been associated with the pathogenesis of cardiac arrythmias, deafness, and epilepsy (Barro-Soria et al.; 2014, Goldman et al., 2009). Beckwith-Wiedemann syndrome is also a disease related to a *KCNQ1* mutation, specifically the loss of its imprinting center (Valente et al., 2019). Additionally, Tommiska et al. discovered that *KCNQ1* mutations are implicated in maternally-inherited gingival fibromatosis and pituitary hormone deficiency (Tommiska et al., 2017). Despite the wide array of diseases in which *KCNQ1* is implicated, the underlying mechanisms of gene pathogenicity are still unknown. This study sought to test a tailored gene activation system that would ensure the future ability to reliably control *KCNQ1* expression while observing concomitant changes in the cellular phenotype.

CRISPR systems

Out of the multitude of gene manipulation strategies on a DNA, RNA, and protein level, the DNA-based technique Clustered Regularly Interspaced Short Palindromic Repeats activation (CRISPRa) was employed due to it being the most recent activation method, and, compared to Open Reading Frame-based methods, CRISPRa can upregulate transcription of endogenous genes. This allows researchers to study genes in their native context – potentially leading to more biologically relevant results. Therefore, its application will be the focus of this study.

DNA-recognition and effector components are necessary for gene manipulation at the DNA level, where the recognition component recognizes the specific gene, and the effector component either makes edits to the DNA sequence or alters gene expression in an epigenetic fashion (Gong et al.,

2021). In the case of CRISPR, guide RNAs (gRNAs) fulfill the role of recognition component, while the Cas (CRISPR associated protein) proteins serve as effectors (Gong et al., 2021).

The Cas proteins are nucleic acid endonucleases, whose purpose is to degrade foreign DNA or RNA (Liu et al., 2020). CRISPR-Cas systems can be categorized into two broad classes, where Class 1 systems rely on multiple Cas proteins working in tandem with each other, and Class 2 systems utilize a large Cas protein (Koonin et al., 2019). Class 2 systems are more widely utilized by genetic engineers because of their simplicity, and Cas9 is the most frequently used protein among them (Wang et al., 2020). The genome editing complex thereby is comprised of Cas9 (a DNA endonuclease), CRISPR RNA (crRNA), and trans-activating crRNA (tracrRNA) (Casas-Mollano et al., 2020; Hirakawa et al., 2020). The basic mode of action is that Cas9 binds crRNAs to tracrRNA scaffolds (also recognized by Cas9) in order to bind to the target DNA sequence and cleave it upon reaching its destination (Casas-Mollano et al., 2020; Hirakawa et al., 2020). The tracrRNA "escorts" the crRNA, allowing it to bind to the 20 nucleotide (nt) long target DNA (protospacer) region. This protospacer is located upstream of a protospacer adjacent motif (PAM) with a sequence specific to the Cas endonuclease being employed. SpCas9, the most commonly used such nuclease, has a PAM sequence of 5'-NGG-3', where "N" is any nitrogenous base, followed by two "G" guanine bases (Anders et al., 2014). The tracrRNA-crRNA complex may be combined into a single guide RNA (sgRNA), which is able to target nucleotide sequences of different DNA loci, allowing tremendous flexibility in genome editing (Casas-Mollano et al., 2020).

The DNA changes induced by CRISPR/Cas can be both random and specific, i.e., lead to non-homologous end-joining (NHEJ) and homology-directed repair (HDR), respectively (Xu and Qi, 2019). The formation of unpredictable insertions and deletions (InDels) with NHEJ can be problematic if the goal of an experiment relies on accuracy in genome editing (Xu and Qi, 2019).

Whereas basic CRISPR/Cas systems may be employed to achieve gene editing, there are also second generation CRISPR techniques that allow for gene knockdown and gene activation instead. These daughter applications are CRISPRi (CRISPR interference) for repression of gene expression and CRISPRa (CRISPR activation) for promotion of gene expression (Figure 1). Cas9 is both the most studied and the most widely utilized effector in Cas systems, taking advantage of its high specificity in editing in order to achieve reliable regulation of gene expression (Santos-Moreno et

al., 2020). dCas9 is a mutated, catalytically inactivated variant of Cas9, lacking an endonuclease component, but maintaining its RNA-guided DNA binding property (Jiang and Doudna, 2017). In the case of CRISPRi, dCas9 is targeted to a region near the transcription start site to inhibit gene expression via the physical obstruction of elongating RNA polymerase by the dCas9/sgRNA complex (Qi et al., 2013). While CRISPRi may see dCas9 employed by itself, it is common to bolster its action by fusing it with additional repressor domains such as KRAB (Kruppel-associated box) (MacLeod et al., 2019). CRISPRi employing dCas9 fused to KRAB domains typically results in significantly stronger inhibition due to KRAB interacting with endogenous chromatin remodeling factors to alter euchromatin into heterochromatin and therefore further reduce the interaction between transcription factors and binding domains (Ede et al., 2017; Du and Qi, 2016).

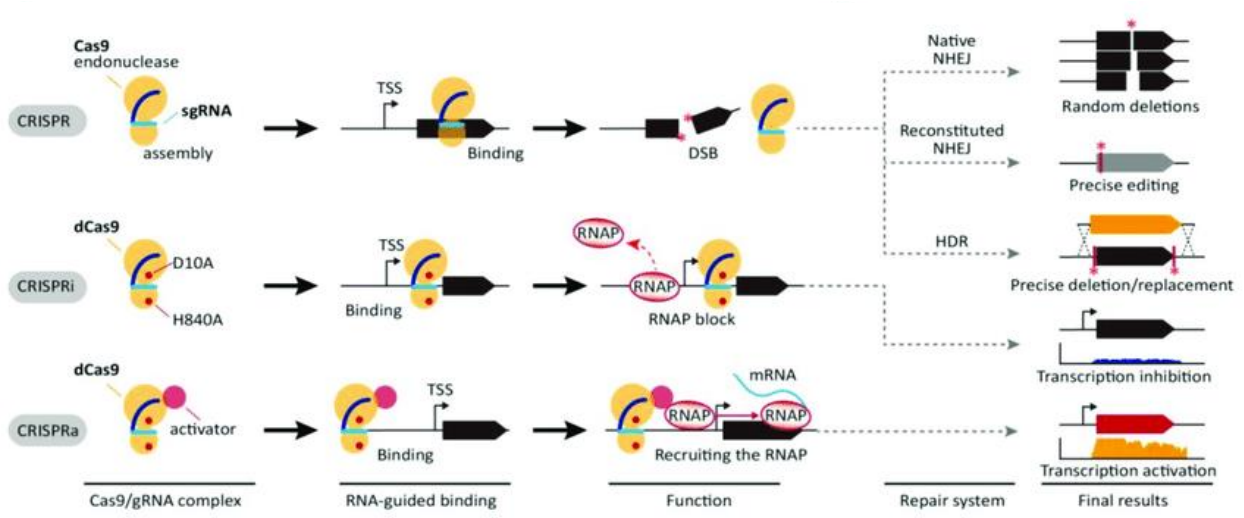

Figure 1. Schematic of CRISPR, CRISPRi, and CRISPRa systems. The CRISPR-Cas9 complex binds to a target site and forms a double-stranded break (DSB) upon reaching its destination. This can induce both the NHEJ and HDR repair mechanisms, which can lead to both random and specific genome editing. CRISPRi and CRISPRa circumvent the induction of a repair system altogether by employing dCas9. In the case of CRISPRi, dCas9 binds upstream of the gene of interest, but downstream of the transcription start site (TSS), blocking RNA polymerase (RNAP) activity and inhibiting the gene. In CRISPRa, dCas9 (alongside an activator) binds upstream of both the TSS and the gene of interest, promoting the activity of RNAP and activating the gene. Image taken from Cho et al. (2018) under the Creative Commons Attribution 4.0 License.

CRISPRa, on the other hand, induces gene expression via association with an effector (Figure 1). In the following section, the efficiency and individual advantages of prominent CRISPRa systems are briefly compared and this paper's choice of VP64-p65-Rta (VPR) is justified.

Design and effectiveness of CRISPRa systems as inducers of gene expression

The various CRISPRa systems rely on the dCas9 effector. The creation of dCas9 involves inactivating both endonuclease domains of Cas9 via D10A point mutations for the RuvC domain and H840A point mutations for the HNH domain, resulting in the removal of its DNA cutting capabilities without impacting is DNA binding ability (Brezgin et al., 2019). In order to facilitate its guiding to a target region, dCas9 is fused to many effector domains, particularly transcription activation domains (Brocken et al., 2018). These domains work in tandem with gRNAs to deliver the protein to the target site and bind to it, instead of cleaving it like in Cas9-based methods (Brocken et al., 2018).

Specifically, the CRISPRa targeting mechanism first sees gRNA bind to apo-dCas9 and induce conformational changes through RNA-protein interactions (Rahman and Tollefsbol, 2021). As a result of these interactions, the dCas9's PAM-interacting domain is repositioned for subsequent DNA interrogation, while the entire dCas9-gRNA complex is further stabilized (Rahman and Tollefsbol, 2021). The activated dCas9/gRNA complex then collides and binds with sequences that match its PAM, with the dCas9 proceeding to melt DNA at the adjacent nucleation site (Jiang and Doudna, 2017). Subsequently, the gRNA carries out strand invasion, wherein it invades the target DNA and base-pairs with the complementary strand, proceeding from the PAM-proximal toward the PAM-distal end and unwinding the target DNA in search of any mismatches with the gRNA. (Jiang and Doudna, 2017; Wilkinson et al., 2019). Upon successful binding, the protein becomes a CRISPR activator, subsequently gaining the function of inducing the expression of genes near the target site (Casas-Mollano et al., 2020).

A typical dCas9-based activator is comprised of the dCas9 protein and a transcriptional activator domain fused together (Tuttle and Chavez, 2016). Two generations of dCas9-activators have been utilized, with the second being an adaptation of the first. First generation activators mostly employed the VP16 domain in either its four tandem repeat iteration (VP64) or ten tandem repeat

iteration (VP160) (Casas-Mollano et al., 2020). The activation effect was not strong enough, however, when using either iteration, as only mild activation effects could be observed. Fusing different activation domains to dCas9, such as p65, yielded similarly modest results (Casas-Mollano et al., 2020).

Second generation dCas9-activators solve this single domain issue by instead recruiting multiple activation domains and fusing them to dCas9, combining several of them in a homogenous or heterogenous manner. One such dCas9-activator is dCas9-VPR (VP64-p65-Rta), developed after determining the most optimal combination of activation domains to achieve the highest efficiency in gene activation (Casas-Mollano et al., 2020). The effectiveness of this tripartite activator was demonstrated when dCas9-VPR was employed to induce the differentiation of human pluripotent stem cells into neurons by increasing the expression of responsible genes *neurogenin 2* and *neurogenic differentiation factor 1*, a feat that its predecessor systems could not achieve. (Casas-Mollano et al., 2020). The activator has also been observed to yield a nearly 4000-fold improvement in activation efficiency in comparison to VP64 (Shakirova et al., 2020).

The individual CRISPRa systems are reliant on different mechanisms of action – p300 maintains euchromatin arrangements by way of histone acetylation (Brezgin et al., 2019), whereas VP64 recruits p300 and other acetyltransferases such as the SAGA and NuA4 complexes while simultaneously facilitating the assembly of the pre-initiation complex (Hirai et al., 2010; Perez-Pinera et al., 2013). SunTag co-expresses epitope-tagged dCas9 with antibody-activated activator proteins. Here, dCas9 is fused to an array of GCN4 epitope motifs whose subsequent interactions with linked single-chain variable fragments serve to recruit effector molecules such as VP64, p300, and others thanks to the high affinity and specificity of antibody binding (Brezgin et al., 2019; Casas-Mollano et al., 2020). SAM appends modified, activator domain-tethering MS2 hairpin loops to the tetraloop and stem-loop 2 of the sgRNA using dimers of MS2 bacteriophage coat protein (MCP) (Casas-Mollano et al., 2020). Moreover, p65 and HSF1 activators are also fused to MCP while VP64 is fused to dCas9 to further increase transcriptional activation (Casas-Mollano et al., 2020). Finally, VPR builds on the individual strengths and mechanisms of VP64, p65, and Rta by sequentially fusing these activation domains into a tripartite activator (Casas-Mollano et al., 2020). There are many other approaches to CRISPRa beyond these, but a major advantage and

justification for the use of newer techniques such as VPR, SAM, and SunTag is their low rate of off-target effects (Du and Qi, 2016).

When VPR is compared with other previously-mentioned second generation approaches, such as SAM and SunTag, all three systems demonstrate activation efficiencies that are within an order of magnitude of each-other (Chavez et al., 2016). While SAM sometimes outperforms its competitors in certain contexts, Chavez et al. note that other practical concerns such as ease of delivery, availability of reagents, and familiarity may ultimately be the deciding factor in choosing between VPR, SAM, and SunTag (Figure 2) (Chavez et al., 2016). The choice of system is therefore highly contextual to both the experiment and the laboratory, with closer consideration revealing minute but important distinctions between approaches. For example, VPR is uniquely advantaged in that it combines the high activation efficiency of second-generation systems and the small transgene size of first-generation ones, which is beneficial in instances where the transgene vector is of limited capacity, such as adeno-associated viruses. (Shakirova et al., 2020).

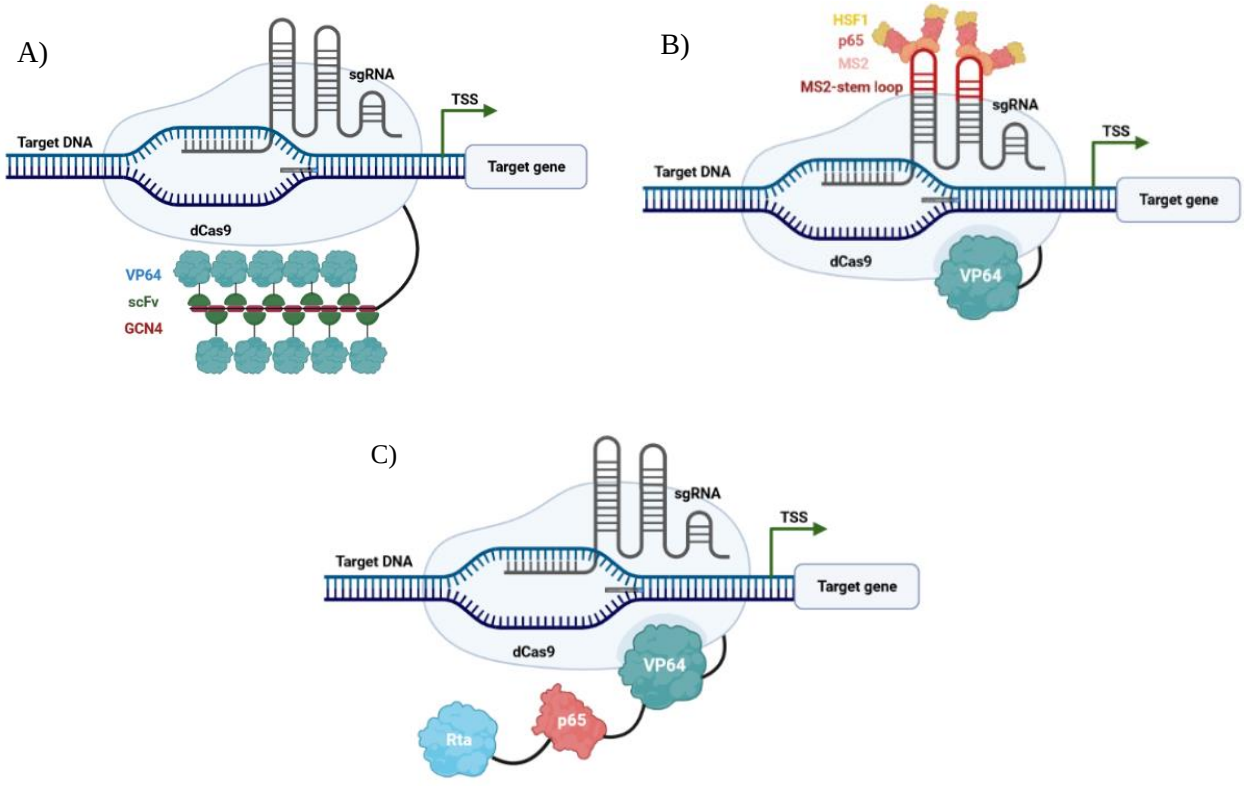

Figure 2. An outline of transcriptional activators that utilize dCas9. A) The SunTag system uses epitope-tagged (GCN4) dCas9 that recruits an antibody-tagged activator protein (VP64). B) The SAM system employs a fusion of dCas9 and VP64, as well as an sgRNA with MS2 hairpin loops tethered to the tetraloop by way of MCP dimers. The activators p65 and HSF1 are also recruited though the partner protein MCP. C) The dCas9-VPR system works by fusing the tripartite activation domain (VP64-p65-Rta) to dCas9. Images created with BioRender.com and adapted from Li et al. (2020).

In designing gRNAs for CRISPRa systems, there is evidence that even though an individual gRNA is effective in inducing gene expression, activation efficiency increases when gRNAs are combined (Shakirova et al., 2020). Additionally, some genes, like *KCNQ1*, possess multiple transcription start sites (TSS), making gRNA targeting more difficult, as different tissues could be affected differently. This is in contrast with the other examined genes in this study, *ASCL1* and *GHRH*, which both have one transcription start site each. It had been previously suggested that if the transcription of a gene is initiated from several different TSSs, the function of the gene would be different each time (Xu et al., 2019). Nonetheless, it is instead proposed that this is due to molecular errors, rather than adaptations, and that the higher a gene's expression is, the less "secondary" TSSs are used (Xu et al., 2019). For this reason, when studying *KCNQ1*, more focus was put into its "primary" TSS.

Aims

The overall aim of this study was to set up the CRISPR-dCas9-VPR tripartite transcriptional activator domain gene activation system in easily cultured and transfectable HEK293 cells and induce the expression of *KCNQ1* as a target gene.

The following goals were pursued over the course of this project:

I. Validation of the HEK TRE-VPR cell line
II. Activation efficiency testing of *ASCL1* and *GHRH*
III. Tet-ON system leakage testing
IV. Activation efficiency testing of *KCNQ1*
V. Construction and testing of new *KCNQ1* gRNAs

Methods

Cell culture

HEK293 cells (Product Code: CRL-1573, Manufacturer: ATCC) were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Sigma-Aldrich, D5796-1L), supplemented with 10% fetal bovine serum (FBS) (Sigma-Aldrich, F7524) and 1% penicillin-streptomycin (Pen-Strep) (Sigma-Aldrich, P0781) at 37°C in a humidified 5% CO_2 incubator. The cells were maintained in a T-75 cell culture flask (ThermoFisher Scientific, 156499), set horizontally in the incubator.

Cells were split twice a week, once 80-90% confluency was reached.

1xPBS (ThermoFisher Scientific, 10010023), Trypsin-EDTA (Sigma-Aldrich, T4174), and DMEM, supplemented with 10% FBS and 1% Pen-Strep, were warmed in a 37°C water bath 30 minutes before each passaging of cells. Old media was aspirated without disturbing the attached cells and they were, subsequently, carefully washed with 1xPBS. 1xPBS was aspirated and 2ml Trypsin-EDTA was added to cover the surface of the flask. The flask was then incubated at 37°C for 2 minutes, so that the cells could detach from the surface. After incubation, new supplemented DMEM was added, and the cells were suspended in the DMEM-Trypsin-EDTA mixture. Supplemented DMEM was also added to a new T-75 flask, and cell suspension from the old one was added to the new one in a 1:6 ratio.

Transfection and drug selection

On Day-1, the HEK293 cells were seeded from a T-75 flask to a 24-well plate (Greiner Bio-One, 662160) in a standardized 1:5 ratio, i.e., 2ml of cell suspension dissolved into 10ml supplemented DMEM. Each well received 500µl of this mixture and was left to incubate for 24h.

On Day 0, Lipofectamine™ 3000 (ThermoFisher Scientific, L3000015) was used for transfection following the instruction. On Day 1, cell media was changed with 15µg/ml Doxycyline-supplemented media. On Day 3, the cells were collected for analysis.

Plasmid preparation

PB-TRE-dCas9-VPR was a gift from George Church (Addgene plasmid #63800; http://n2t.net/addgene:63800; RRID: Addgene_63800). The plasmid was expanded in LB medium, supplemented with 1µM amplimycin, on a shaker at 37°C overnight. A plasmid mini-prep system (Promega PureYield™ Plasmid Miniprep System, A1223) was utilized to isolate and purify the plasmid. DNA concentration was measured using the NanoDrop™ 2000c Spectrophotometer (ThermoFisher Scientific, ND-2000C). The sample values for concentration, the ratio of the absorbance at 260 and 280 nm ($A_{260/280}$), and the ratio of the absorbance at 260 and 230 nm ($A_{260/230}$) were recorded.

Agarose gel electrophoresis and imaging

DNA fragments were separated by their length in an agarose matrix via electrophoresis. The products were run in gels of a 1% agarose (Bio-Nordika, BN-50004) in 1xTAE buffer (ThermoFisher Scientific, 24710030) mixture. This mixture was heated up in a microwave and subsequently stirred until clear; afterwards, Ethidium Bromide (ThermoFisher Scientific, 11846714) was homogenized into the agarose liquid mixture (1 Pasteur pipette droplet/50ml mixture). The liquid was then poured into a suitable gel template and comb, and it was let to rest for 30 minutes. 1kb DNA ladder (ThermoFisher Scientific, SM0314) and each sample, supplemented with 6xDNA loading dye (Product Code: B7024S, Manufacturer: New England BioLabs), were interred into the now-solid gel, and the electrophoresis machine (Bio-Rad, 043BR02625) was run at 100V/30 minutes. DNA band detection was done using a ChemiDoc imaging system (ChemiDoc™ MP Imaging System, 12003154), with the Ethidium Bromide detection setting. Both original and pictures in a different contrast were taken.

Genomic DNA (gDNA) isolation

The NucleoSpin gDNA Clean-up (Macherey-Nagel, 740230.250) protocol was followed for the gDNA isolation process. As per the protocol, Binding Buffer was added to the sample, DNA was bound to a silica membrane, and the membrane was washed and subsequently dried. For elution

of gDNA, 60µl of pre-warmed (at 70ºC as per the protocol) BE buffer was passed through the column twice. Afterwards, gDNA concentration was measured via Nanodrop.

Polymerase Chain Reaction (PCR)

PCR was conducted in order to test the dCas9 expression properties of the isolated gDNA. A master mix (Table 1) was made and used in PCR reaction programs, outlined in Table 2.

Table 1. Reaction volumes in each mixture for the PCR reaction to test the presence of dCas9 in HEK TRE-VPR cell gDNA.

Reaction system	Volume
10X PCR buffer I	2.5 µl
10mM dNTPs	0.5 µl
10uM Forward Primer	0.5 µl
10uM Reverse Primer	0.5 µl
AmpliTaq Gold DNA polymerase	0.125 µl
Template (previously isolated HEK TRE-VPR gDNA)	250 ng
H_2O	to 25 µl

Table 2. PCR reaction program for the established optimal temperature to produce a dCas9 band in HEK TRE-VPR cell gDNA.

Reaction program	Time	Cycles
95°C	10 min	1
95°C	15 s	35
62°C	30 s	
72°C	2.5 min	
72°C	5 min	1
4°C	hold	∞

gRNA Design

Lab-provided gRNAs (Table 3), targeting *ASCL1*, *GHRH*, and *KCNQ1*, were tested for the experiment prior to further design of *KCNQ1* gRNAs.

Table 3. Lab-provided gRNAs, targeting *ASCL1*, *GHRH*, and *KCNQ1*.

Name	Target	crRNA sequence
ASCL1-gRNA1	*ASCL1*	AGCCGCTCGCTGCAGCAGCG
ASCL1-gRNA2	*ASCL1*	CTCCCCGCTGCTGCAGCGAG
ASCL1-gRNA3	*ASCL1*	GGAGGGGGAGTTTAGGGAGT
GHRH-gRNA1	*GHRH*	ACAAACCCTGTTGGTGAAAG
GHRH-gRNA2	*GHRH*	ACAGTCCTTTGGTTGACTTG
GHRH-gRNA3	*GHRH*	ATTACCCACAAGTCAACCAA
KCNQ1-gRNA1	*KCNQ1*	CAGGGACGCCCTGTGCGCGC
KCNQ1-gRNA2	*KCNQ1*	CGAGGGCCGGTCCCATGGGA
KCNQ1-gRNA3	*KCNQ1*	CGTCCCATGGGACCGGCCCT

The online tool Benchling (https://www.benchling.com/crispr/) was used for *KCNQ1* gRNA design, targeting four transcription start sites (TSS) of the gene. These sites were determined using the FANTOM genome browser (FANTOM, Genome Exploration Research Group, Genomic Sciences Center, Yokohama Institute, RIKEN, Yokohama, Kanagawa, Japan; https://fantom.gsc.riken.jp/zenbu) and all of them form the putative TSS region for the gene. The target for gRNAs in all four cases was the sequence around 100nt upstream of the respective TSS (Table 4).

The off-target score measures specificity, i.e., how precisely dCas9 binds to the target sequence and avoids binding to unrelated sequences in the genome (Pellegrini, 2016). The on-target score is also called the efficiency score, and it determines how well a gRNA conducts its cleavage activity (Pellegrini, 2016). gRNAs possessing higher scores, up to a maximum of 100 (and a minimum of 0), are considered more effective (Pellegrini, 2016). Typically, off-target scores above

50 are mandatory for a functional gRNA, while a mid-to-high on-target score is preferred, but not required (Kahn and East, 2018). Considering on-target (efficiency), and off-target (specificity) scores, six gRNAs were chosen for building.

Table 4. Designed gRNAs, targeting four TSS of *KCNQ1*.

Name	Sequence	Target	Strand	Position	On-Target Score	Off-Target Score
gRNA1	GGCACGCCGTCCCATGGGAC	TSS1	1	2466151	40.51	68.40
gRNA2	CCCTGTGCGCGCCGGACCGG	TSS1	1	2466103	40.37	89.37
gRNA3	GGCGGGGCACGCCGTCCCAT	TSS2	1	2466146	45.84	90.84
gRNA4	TGGCCGAGGGCCGGTCCCAT	TSS2	-1	2466150	40.93	84.70
gRNA5	CGGGAAGGGAACCTTGAGTG	TSS3	1	2482611	62.14	67.15
gRNA6	CGGGGCTGTGAGTCATTCCG	TSS4	1	2798750	69.63	78.79

Assembly of gRNA transcriptional cassettes by PCR

gRNA-PCR products were generated using the designed 59bp oligos by PCR (Balboa Alonso, 2018). The reaction mix is outlined in Table 5, while the PCR program is shown in Table 6.

Table 5. Generation of gRNA-PCR products from 59bp oligos.

Reagent	Volume [µl]
5xHF buffer	20
dNTP (2.5mM)	2
gRNA (1µM)	2
1aggc Fw (10µM)	5
1aggc Rv (10µM)	5
Phusion (Thermo)	1
U6 promoter tailed (2ng/µl)	2.5
Term tailed (2ng/µl)	2.5
PCR-grade H_2O	60

Table 6. Assembly PCR reaction program.

Reaction program	Time	Cycles
98°C	3 min	1
98°C	10 s	35
52°C	30 s	
72°C	20 s	
72°C	8 min	1
4°C	hold	∞

Agarose gel electrophoresis was done as in the "Agarose gel electrophoresis" section in order to test and confirm the gRNA size. Both a 1kb plus ladder (Product Code: SM1331, Manufacturer: ThermoFisher Scientific) and a 100bp ladder (Product Code: SM0243, Manufacturer: ThermoFisher Scientific) on each end of the comb were applied for measuring the size in this experiment.

The resulting gRNA-PCR products were purified using the NucleoSpin® Gel and PCR Clean-up kit (Takara Bio, 740609.250). The protocol 5.1 "PCR Clean-up" was followed to produce the final, transfectable gRNA products.

RNA extraction, reverse transcription and real-time quantitative PCR (RT-qPCR)

The NucleoSpin RNA plus (Macherey-Nagel, 740984.250) protocol was followed for the RNA extraction process. Cells had media aspirated, and 350µl Lysis Buffer (reagent in NucleoSpin RNA Plus kit (Macherey-Nagel, 740984.250)) was added in each well. The cells were then disrupted and suspended in the buffer, and were placed on ice. For elution of RNA, 30µl RNAase-free H_2O was added to the RNA-containing column, and the flow through RNAase-free H_2O was pipetted back in the column after centrifugation. RNA concentration was measured with Nanodrop.

Reverse transcription was performed with the iScript™ cDNA Synthesis Kit (Bio-Rad, 1708891), using the following volume measurements for a total complementary DNA (cDNA) sample volume of 20µl (Table 7). The PCR reaction protocol used for the completion of reverse transcription is outlined in Table 8.

Table 7. Reverse transcriptase reaction mix for a 20 µl reaction. Adapted from "iScript™ cDNA Synthesis Kit" document supplement at the Bio-Rad website.

Component	Volume per Reaction [µl]
5x iScript Reaction Mix	4
iScript Reverse Transcriptase	1
RNA template	$\frac{1\mu g\ RNA\ [\mu g]}{concentration\ of\ sample\ [\mu g/\mu l]}$
Nuclease-free H_2O	Total volume – volume of other three components
Total volume	20

Table 8. Reverse transcription PCR reaction protocol. Adapted from "iScript™ cDNA Synthesis Kit" document supplement from the Bio-Rad website.

Step	Procedure
Priming	5 min at 25ºC
Reverse Transcription	30 min at 46ºC
RT inactivation	1 min at 95ºC
Upon completion	Hold at 4ºC

RT-qPCR was done by combining 5x HOT FIREPol EvaGreen qPCR Mix Plus (Solis Biodyne, 08-25-00001), cDNA, a 2µM primer pair mix (Table 9), and PCR grade H_2O to a total volume of 20µl, as outlined in Table 10. Pipetting was done using a multichannel pipette, and samples were pipetted in duplicate or triplicate. The cycle program used is shown in Table 11.

Table 9. Primer sequences used for conducting PCR (dCas9 (1) and dCas9 (2)) and RT-qPCR (GAPDH, ASCL1, GHRH, KCNQ1-O, and KCNQ1-N).

Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')	Notes
dCas9 (1)	CGGTTGATCTATCTCGCGCT	ACATCGCTGTTGTCTGGGTT	PCR product size: 89bp.
dCas9 (2)	GGACAGTCTTCACGAGCACA	GGTGGCCTTGCCTATTTCCT	PCR product size: 949bp.
GAPDH	GCTTGGTCGTCAGGTAGACAG	GAGTTCCTGCTGTGCTTCTGC	Housekeeping gene.
ASCL1	CGCTCGGCGGTCGAGTA	AGATGCAGGTTGTGCGATCA	
GHRH	GCTTGGTCGTCAGGTAGACAG	GAGTTCCTGCTGTGCTTCTGC	
KCNQ1-O	CATTGGGAAGCCCTCACTGT	CTTCTACTCGGTTCAGGCGG	
KCNQ1-N	TGTCCACCATCGAGCAGTATG	CCGTCCCGAAGAACACCAC	Adapted from Fan et al., 2018.

Table 10. RT-qPCR reaction Mix for a 20µl reaction.

Reagent	Volume [µl]
5x HOT FIREPol EvaGreen qPCR Mix Plus	4
Primer Pair (Forward+Reverse) [2µM]	0.4
cDNA	0.4
PCR grade H_2O	15.2
Total volume	20

Table 11. RT-qPCR cycle program used. Adapted from "HOT FIREPol® EvaGreen® qPCR Supermix, 5x" Solis Biodyne datasheet.

Cycle step	Temperature	Time	Cycles
Initial activation	95ºC	12 min	1
Denaturation	95ºC	15 s	40
Annealing	60ºC	25 s	
Extension	72ºC	25 s	

Statistical analysis

Statistical analysis was performed using the analysis and graphing software GraphPad Prism version 8.0.0 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com. Transfections were done in biological and technical duplicates. Graphs, analyzing qPCR data, used the log10 of $2^{(-\Delta\Delta CT)}$ and calculated the mean with 95% CI on the resulting graph. The $2^{(-\Delta\Delta CT)}$ method was employed due to it being a practical method for measuring the changes in relative gene expression (Livak and Schmittgen, 2001). Finally, a one-way ANOVA test and a Dunnett multiple comparison test was performed to determine the *p*-values and statistical significance of each sample, compared to the control group (Dunnett, 1955; Ross and Wilson, 2017). For comparing between leakage conditions, a Tukey multiple comparisons test was used to determine *p*-values, with comparisons being made between both the control and test conditions, as well as between individual test conditions (Tukey, 1949).

Results

PB-TRE-dCas9-VPR plasmid validation

Before transfection, the PB-TRE-dCas9-VPR plasmid was verified using a diagnostic restriction enzyme digest. The plasmid was linearized with NheI enzyme (red box indicated) to be suitable for checking the exact plasmid size (Figure 3). The linearized plasmid was expected to produce a single agarose gel band with a size equal to 13808bp.

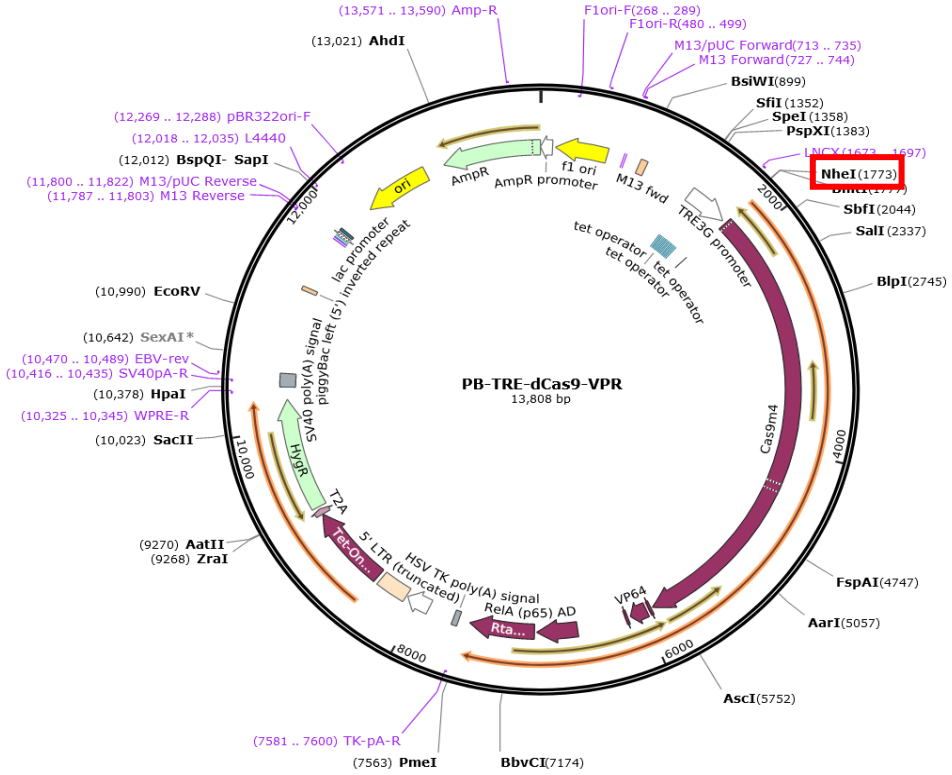

Figure 3. Schematic of the PB-TRE-dCas9-VPR plasmid, as illustrated by Addgene and reproduced for noncommercial use according to the Addgene Terms of Use.

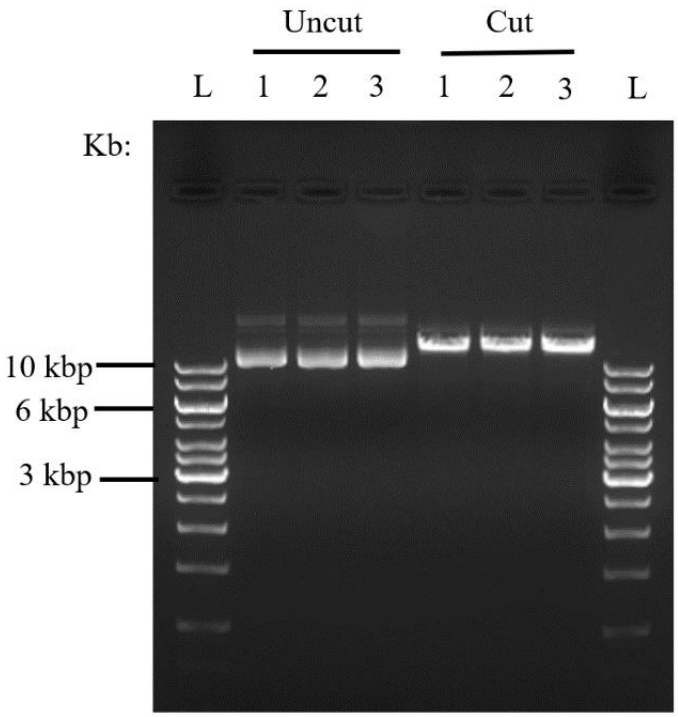

Figure 4. TAE 1% agarose gel of NheI enzyme cut. L = 1 kb ladder. Samples "Uncut" 1-3 contain an uncut (control) plasmid, while the plasmid cut with NheI is in samples "Cut" 1-3.

The results reflected what was expected (Figure 4). In the control wells, containing the uncut plasmid (Uncut 1-3), there were two observed bands. This meant that the bacterial DNA was in two conformations: nicked above 10kbp (top band) and supercoiled at ~10kbp (bottom band), with the majority being in the supercoiled conformation based on band intensity.

In the enzyme-cut wells, containing the cut plasmid (Cut 1-3), there was a single band at about the expected size at above 10kbp. This meant that the restriction digest worked correctly, and the plasmid size is as expected.

Detecting dCas9 presence in HEK TRE-VPR

The HEK TRE-VPR cell line was established previously in Raivio lab by transfecting HEKs with the plasmids (PB-TRE-dCas9-VPR and PBase). To prove the cell line has been incorporated with the activation system, the presence of dCas9 in the HEK TRE-VPR cell line was detected.

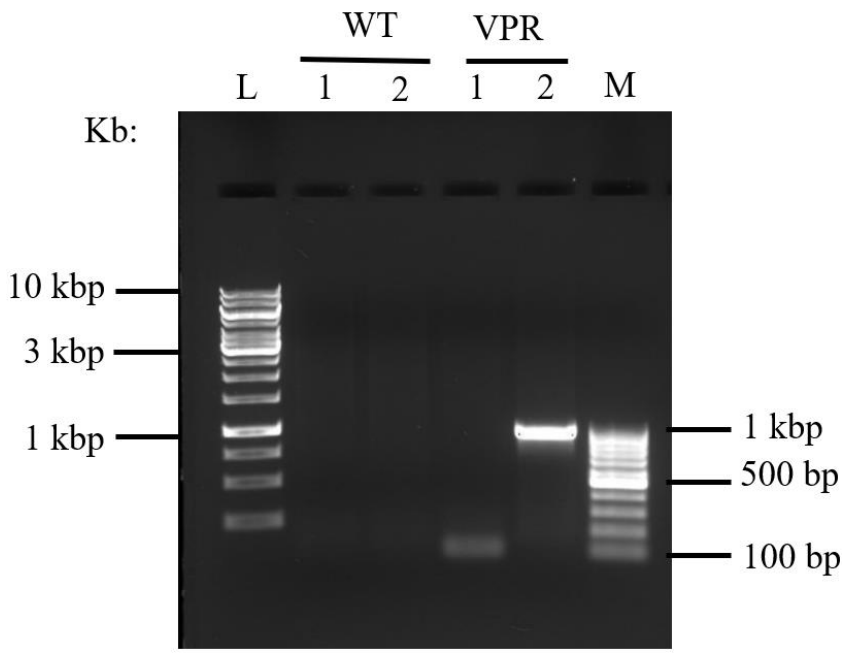

Figure 5. TAE 1% agarose gel validating the presence of dCas9 in HEK TRE-VPR gDNA. L = 1 kb ladder. M = 100 bp ladder. Sample WT1 contains HEK293 wild type cells, tested with primer dCas9 (1), and sample WT2 contains HEK293 wild type cells, tested with primer dCas9 (2) (Table 9). Sample VPR1 contains the HEK293 TRE-VPR cell line, tested with primer dCas9 (1), and sample VPR2 contains the HEK293 TRE-VPR cell line, tested with primer dCas9 (2). Tm in all wells = 62ºC.

Two cell lines, HEK wild type (WT) and HEK TRE-VPR were tested for the presence of dCas9 in this follow-up, with a constant Tm = 62ºC being used for all samples. Two new dCas9 primers were tested, as a previous dCas9 detection experiment in the HEK TRE-VPR cell line yielded no bands on the agarose gel. No bands were expected in the HEK WT sample wells (WT1 and WT2), as it does not express dCas9.

There was a noticeable band observed in each primer well at ~100bp (VPR1) and ~1000bp (VPR2), respectively (Figure 5). A band of 89bp was expected in the VPR1 sample, and a band of 949bp was expected in the VPR2 sample. This meant that both primers work correctly and the HEK TRE-VPR gDNA was confirmed to express dCas9.

HEK TRE-VPR cell line activation test of *ASCL1* and *GHRH*

After the VPR system was validated, the genes *GHRH* and *ASCL1* were tested in order to determine the activation efficiency of the system. Targeting *GHRH* led to the successful activation of the gene (Figure 6). The neg1 blank gRNA was used as a control group, by which the other groups would be measured, and, thus, it assumed a relative *GHRH* expression of 0. The transfection of the neg2 blank gRNA aimed to serve as a contrasting group to the gene-targeting gRNAs – a control of the control. It fulfilled its purpose, as its expression was x-0.19. Overall, the *GHRH*-targeting gRNAs showed a net increase in the gene's expression, proving the efficacy of the VPR system and the potency of the gRNAs. The GHRH-1 group demonstrated the lowest expression of the three gene-targeting groups, with an expression of x1.93. GHRH-2 had the highest expression of x2.77, while GHRH-3 had an expression of x2.12. A Dunnett multiple comparisons test showed all three GHRH conditions as statistically significant, with adjusted p = 0.0087, p = 0.0018, and p = 0.0058 for GHRH-1, GHRH-2, and GHRH-3, respectively. The neg2 condition was not statistically significant, with p = 0.9491.

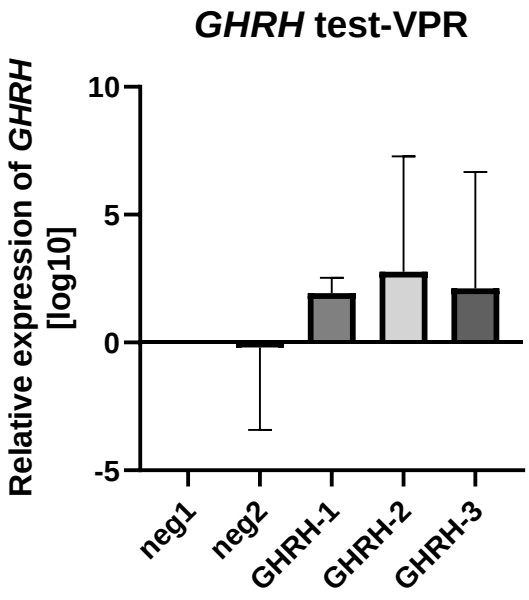

Figure 6. RT-qPCR test of GHRH expression levels in the established HEK TRE-VPR system. Single gRNAs neg1 and neg2 are blank guides, while gRNAs GHRH-1-3 target *GHRH*. Measure of increase of

GHRH expression: neg1 = 0; neg2 = x-0.19; GHRH-1 = x1.93; GHRH-2 = x2.77; GHRH-3 = x2.12. Error bars: 95% CI.

The *ASCL1*-targeting gRNA transfection data, conveys similar results as with *GHRH* (Figure 7). The neg1 blank gRNA was the control, and neg2 served to contrast the other positive datasets, as in the previous experiment, and it had an expression of x-1.39. ASCL1-1 and ASCL1-2 had a similar expression level of x2.29 and x2.40, respectively. ASCL1-3 had the highest expression level of x2.71. A Dunnett multiple comparisons test unexpectedly demonstrated all conditions as statistically significant, with adjusted $p < 0.0001$ in all four cases (neg2, ASCL1-1, ASCL1-2, and ASCL1-3). Overall, it could be concluded that *ASCL1* was able to be activated by certain gRNAs, with differing efficacy, and that the constructed VPR system functions correctly.

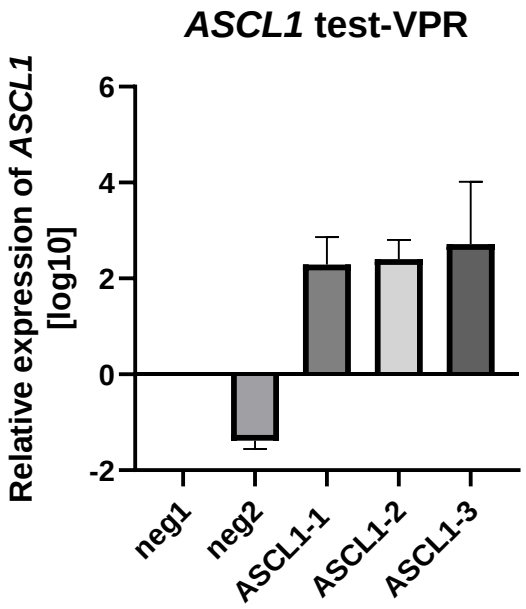

Figure 7. RT-qPCR test of *ASCL1* expression levels in the established HEK TRE-VPR system. Single gRNAs neg1 and neg2 are blank guides, while gRNAs ASCL1-1-3 target *ASCL1*. Measure of increase of *ASCL1* expression: neg1 = 0; neg2 = x-1.39; ASCL1-1 = x2.29; ASCL1-2 = x2.4; ASCL1-3 = x2.72. Error bars: 95% CI.

Testing dCas9 leakage

A leaky inducible system overexpresses a gene even when the inducer (here: doxycycline) is not present. Since leakage may decrease the reliability and accuracy of experimental data from a given system, it is important to account for its presence. The activation system in this study is under a Tet-ON switch, with which there is an associated leakage problem (Pham et al., 2008). Therefore, the next step for quality control of the VPR system was testing for the severity of the leakage problem in this cell line.

The best-performing *ASCL1* gRNA from before, ASCL1-3 (Figure 7), was chosen for the experiment. Several conditions were tested: ctrl (the control group with no transfected gRNAs), neg1-Dox (a blank gRNA with no addition of doxycycline), neg1-Dox (the same blank gRNA with the addition of doxycycline), ASCL1-Dox (an *ASCL1*-targeting gRNA with no addition of doxycycline), and ASCL1+Dox (the same *ASCL1*-targeting gRNA with the addition of doxycycline). A graph of the RT-qPCR data analysis showed the expression rates of each condition (Figure 8).

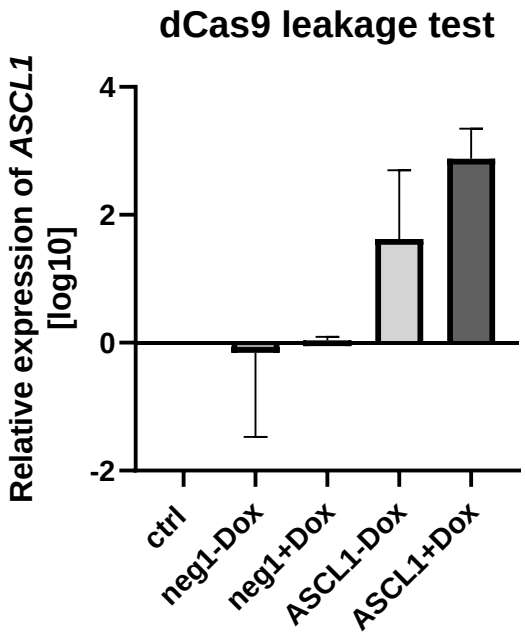

Figure 8. RT-qPCR test of dCas9 expression levels in the established HEK TRE-VPR system, examining whether there is leakage present. Ctrl were untransfected cells, single gRNA neg1 was a blank guide, with or without the supplementation with doxycycline (neg1-Dox/+Dox), and gRNA ASCL1 targets *ASCL1*, with or without the supplementation with doxycycline (ASCL1-Dox/+Dox). Measure of increase of *ASCL1*

expression: ctrl = 0; neg1-Dox = x-0.16; neg1+Dox = x0.04; ASCL1-Dox = x1.62; ASCL1+Dox = x2.88. Error bars: 95% CI.

The relative expression of ctrl was 0, as the control group. The neg1-Dox and neg1+Dox groups had a low gene expression of x-0.16 and x0.04, as was expected. The difference between neg1 conditions with and without doxycycline was negligible, and this was also reflected in the Tukey multiple comparisons test for the sample, which affirmed them as not statistically significant with $p = 0.4610$ for the neg1-Dox condition and $p = 0.9911$ for the neg1+Dox condition. The ASCL1-Dox group had a relative expression of x1.62, while ASCL1+Dox showed a 1.78 times increased expression of x2.88. The Tukey multiple comparisons test also demonstrated the ASCL1-Dox and ASCL1+Dox conditions to both be statistically significant, with both having $p < 0.0001$. The ASCL1-Dox and ASCL1+Dox conditions were further compared between each other, with the results from this comparison revealing a statistically significant $p = 0.0002$. The results showed a definite increase of *ASCL1* expression while exposed to doxycycline, but a substantial activation of the gene could also be observed in the non-doxycycline treated group (ASCL1-Dox).

HEK TRE-VPR cell line activation test of *KCNQ1*

After validating the correct functioning of the HEK TRE-VPR activation system, the next aim was to use this system to activate *KCNQ1* expression. In contrast to the previous two tests, the *KCNQ1* data analysis showed an overall negligible overexpression of the gene (Figure 9). The neg1 blank gRNA was the control, and neg2 contrasted the other negative controls, as in the previous experiment, with an increase in gene expression of x0.55. KCNQ1-1 and KCNQ1-3 had a similarly low expression level of x-1.83 and x-1.27, respectively. KCNQ1-2 had the highest relative expression level of x1.12. A Dunnett multiple comparisons test revealed that there are no statistically significant results, with the $p > 0.7000$ in all cases ($p = 0.9922$, $p = 0.7125$, $p = 0.9158$, and $p = 0.8804$ for neg2, KCNQ1-1, KCNQ1-2, and KCNQ1-3, respectively). These results demonstrate that the three pre-designed gRNAs are incapable of achieving sufficient *KCNQ1* upregulation, necessitating the design of new guides.

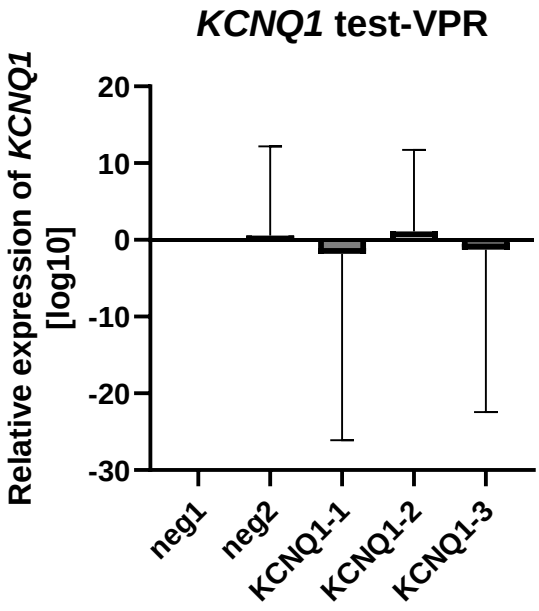

Figure 9. RT-qPCR test of *KCNQ1* expression levels in the established HEK TRE-VPR system, using the KCNQ1-O primer. Single gRNAs neg1 and neg2 are blank guides, while gRNAs KCNQ1-1-3 target *KCNQ1*. Measure of increase of *KCNQ1* expression: neg1 = 0; neg2 = x0.55; KCNQ1-1 = x-1.83; KCNQ1-2 = x1.12; KCNQ1-3 = x-1.27. Error bars: 95% CI.

Construction of the gRNA-PCR products

In order to investigate whether the gRNAs themselves were not efficient enough in activating *KCNQ1* (Figure 9), six new gRNAs were designed and tested. After the construction of the gRNA-PCR products (Table 4), verifying the correct guide size was necessary to determine whether the design and assembly was successful. The resulting band size in the gRNA1-6 wells is reflected in an agarose gel image (Figure 10). The band size in all gRNA1-6 wells was marginally less than 500 bp, which was consistent with the expected band size of 450 bp. This indicated that the construction of the gRNA-PCR products was successful and that the gRNAs were suitable for transfection.

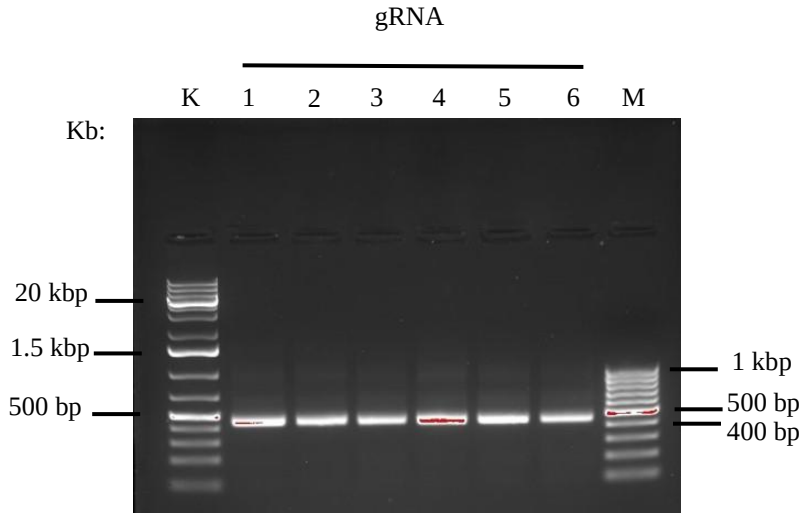

Figure 10. TAE 1% agarose gel validating the correct product band size (450bp) in the amplified gRNA-PCR products. K = 1 kb plus ladder. M = 100 bp ladder. Samples gRNA1-6 denote the six constructed gRNA-PCR products.

Testing of newly designed guides in HEK TRE-VPR cell line

The first test of the newly designed *KCNQ1* gRNAs aimed to compare how well they can activate the gene in the VPR system. The gene expression levels for these gRNAs from both KCNQ1-O and KCNQ1-N primers were compared.

The results of the test analysis were shown in a bar graph (Figure 11). The ctrl group had an expression of 0, as in previous experiments. The overall *KCNQ1* expression in the gRNA transfected cells was lower than expected, but gRNA2 and gRNA5 stood out as the best out of the group, with an expression of x0.30 and x0.26, respectively. gRNAs 1, 3, and 6 had a middling expression of x0.14, x0.24, and x0.13 respectively, while gRNA4 had the lowest expression of x0.07. A Dunnett multiple comparisons test confirmed the low expression rates, and no conditions were found to be statistically significant, with the most significant $p = 0.1523$ belonging to gRNA2. The other gRNAs had $p = 0.8061$, $p = 0.3123$, $p = 0.9843$, $p = 0.2479$, $p = 0.8216$ (gRNAs 1, 3, 4, 5, and 6, respectively).

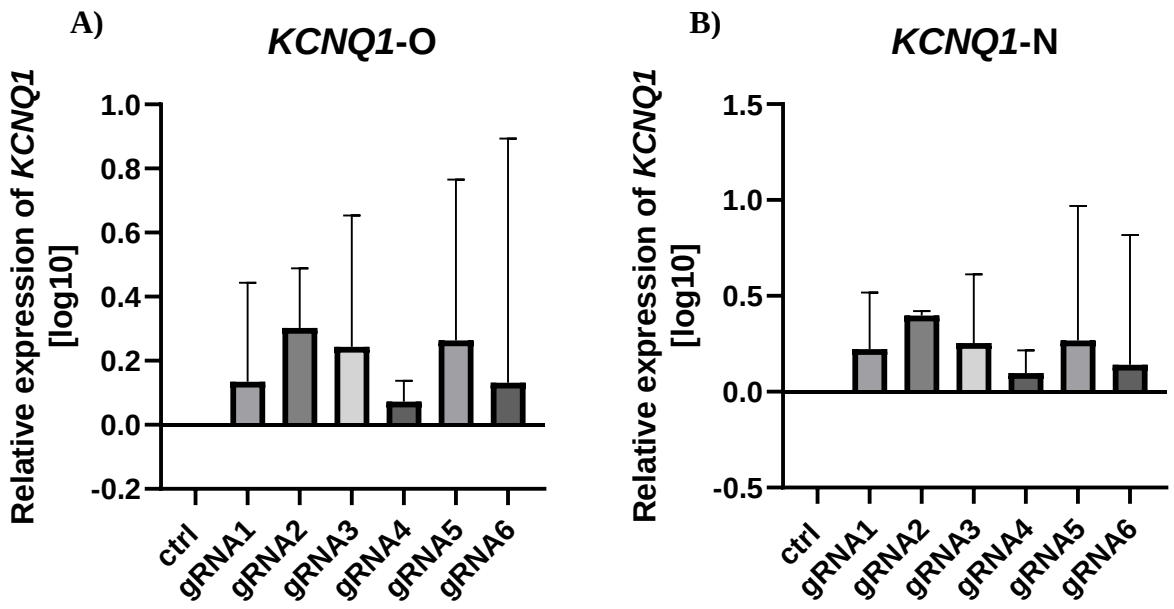

Figure 11. RT-qPCR tests of *KCNQ1* expression levels in the established HEK TRE-VPR system. A) The primer used is "KCNQ1-O". Control were untransfected cells, and single gRNAs gRNA1-6 target different TSSs of *KCNQ1*. Measure of increase of *KCNQ1* expression: ctrl = 0; gRNA1 = x0.14; gRNA2 = x0.30; gRNA3 = x0.24; gRNA4 = x0.07; gRNA5 = x0.26; gRNA6 = x0.13. B) The primer used is "KCNQ1-N". Control were untransfected cells, and single gRNAs gRNA1-6 target different TSSs of *KCNQ1*. Measure of increase of *KCNQ1* expression: ctrl = 0; gRNA1 = x0.22; gRNA2 = x0.40; gRNA3 = x0.25; gRNA4 = x0.10; gRNA5 = x0.27; gRNA6 = x0.14. Error bars: 95% CI.

Similar conclusions could be drawn from the data of the KCNQ1-N primer to that of the KCNQ1-O primer (Figure 11). The ctrl group had an expression of 0. gRNA2 was still the best-performing guide, with an expression of x0.40. gRNA5 fell into the category of middling expression at x0.27, alongside gRNAs 1 and 3, with expressions of x0.22 and x0.25, respectively. gRNAs 4 and 6 had low expression of x0.10 and x0.14, respectively, with gRNA4 being once again the worst performing one. Dunnett multiple comparisons analysis confirmed similar results to the KCNQ1-O data, as all gRNA conditions, bar that of gRNA2, were statistically insignificant. gRNA2 once again had the lowest $p = 0.0471$, while the other gRNAs had $p = 0.4129$, $p = 0.2935$, $p = 0.9438$, $p = 0.2527$, $p = 0.7946$ (gRNAs 1, 3, 4, 5, and 6, respectively).

The results from both tests showed that *KCNQ1* was difficult to activate when using this activation method and system, but gRNA2 seemed to be the most promising gRNA, due to it achieving upregulation with a low *p*-value.

Testing of gRNA1 and gRNA2

The best-performing gRNA, gRNA2, was chosen for further testing alongside gRNA1, which shares the same TSS of *KCNQ1*. The results of this test (Figure 12) drew similar conclusions to those in the previous experiment, testing all six gRNAs (Figure 11). The KCNQ1-O primer was tested first. The ctrl group had a gene expression of 0, as in previous experiments. Both gRNA1 and 2 had a low *KCNQ1* activation level, at x0.24 and x0.31, respectively. A Dunnett multiple comparisons test showed neither gRNA to yield statistically significant results, with gRNA1 having $p = 0.3201$ and gRNA2 having $p = 0.2112$.

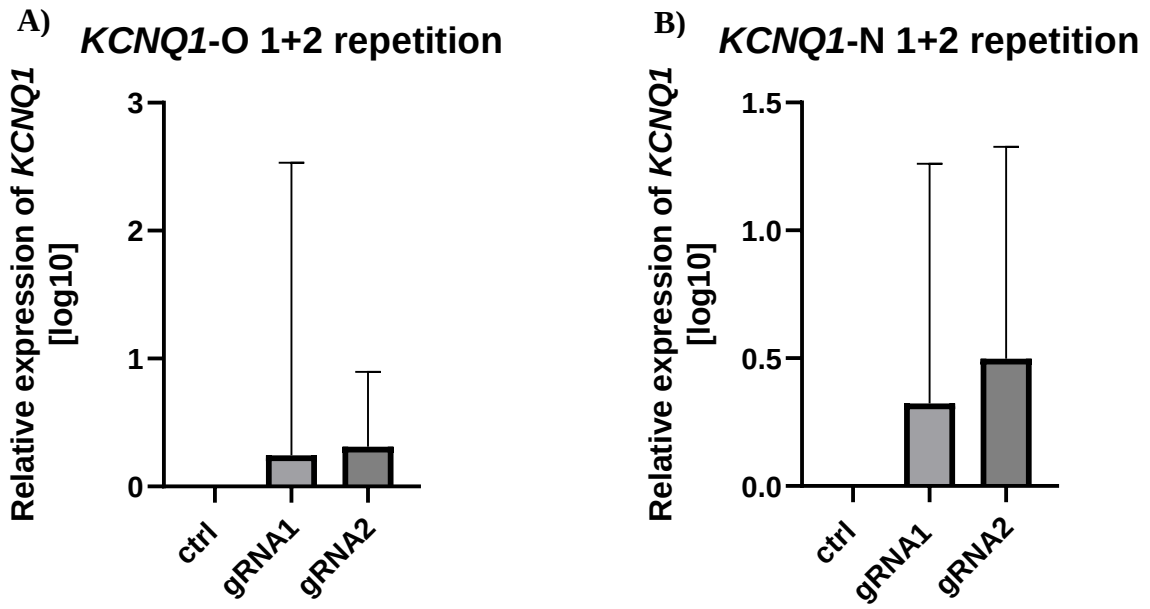

Figure 12. RT-qPCR tests of *KCNQ1* expression levels in HEK TRE-VPR cell line with gRNA target *KCNQ1*. A) The primer used is "KCNQ1-O". Control were untransfected cells, and single gRNAs gRNA1 and gRNA2 target the first TSS of *KCNQ1*. Measure of increase of *KCNQ1* expression: ctrl = 0; gRNA1 = x0.24; gRNA2 = x0.31. B) The primer used is "KCNQ1-N". Control were untransfected cells, and single

gRNAs gRNA1 and gRNA2 target the first TSS of *KCNQ1*. Measure of increase of *KCNQ1* expression: ctrl = 0; gRNA1 = x0.32; gRNA2 = x0.5. Error bars: 95% CI.

A similar expression rate of *KCNQ1* could be observed with the KCNQ1-N primer to the KCNQ1-O primer (Figure 12). The ctrl group had a gene expression of 0, as in previous experiments. There was a small amount of positive gene expression in both gRNA1 and 2 at x0.32 and x0.5, respectively, compared to the ctrl condition, but it was not enough to consider it proof of the gene's activation. A Dunnett multiple comparisons test showed both gRNA1 and 2 datasets as statistically significant, in direct contrast to the KCNQ1-O data. The *p*-values for gRNAs 1 and 2 were $p = 0.0451$ and $p = 0.0140$, respectively.

The discrepancy between the results yielded by the two primers made it difficult to draw a reliable conclusion regarding the efficacy with which *KCNQ1* is activated. Considered together, the results from both experiments showed that *KCNQ1* is difficult to activate and that the assembled gRNAs and the VPR system are not able to induce significant activation.

Discussion

The CRISPRa VPR system was tested to gauge its capacity to induce an increase in *KCNQ1* expression in modified HEK293 (HEK TRE-VPR) cells, in order for it to potentially be used as a tool to treat *KCNQ1*-related diseases, such as pituitary hormone deficiency. To this end, relative genes *ASCL1* and *GHRH* were tested alongside *KCNQ1* in order to compare their expression rates, resulting in the conclusion that *KCNQ1* has a substantially smaller increase in activation than its counterpart genes. The doxycycline-inducible VPR system was then tested for dCas9 leakage and a leaky expression was confirmed, indicating that the system may need further improvement to be able to induce a higher gene expression. Nonetheless, new *KCNQ1* gRNAs were designed and tested, with the resulting *KCNQ1* expression increase being similarly low. Finally, two of the well-performing gRNAs, targeting the first TSS of the gene, were tested once again, with similarly low expression being observed.

Overall, the results of this experiment show that *KCNQ1* has a marginal increase in activation when activated in HEK TRE-VPR cells, but not enough to make a statistically significant difference, compared to the controls. Improving future iterations of this experiment would entail refining the technical aspect, as well as determining the reason behind the low activation efficiency of the gene.

Leakiness of the TET-inducible system

'Leaky' gene expression occurs in an inducible system when gene activity is upregulated without the addition of an inducer, such as the commonly used tetracycline or, its homolog, doxycycline (Sar and Dalai, 2021). In prior experiments, leakage was dealt with using vectors, which reduce the gene copy number in various amounts. However, this approach negatively affects gene expression, which did not achieve sufficiently high levels after the gene copy number reduction (Pham et al., 2008).

In this paper, the Tet (tetracycline)-ON/OFF inducible system was employed for its gene activation/deactivation effect, and no leakage prevention method was employed, as *KCNQ1* expression was already decidedly low (Pham et al., 2008). Nonetheless, leakage of dCas9 was observed to occur when the expression of *ASCL1* became upregulated, even without the addition

of doxycycline (Figure 8). The presence of leakage is a common issue in inducible systems, subtracting from the reliability of the results they yield. Some methods to reduce leakage while maintaining a high gene expression have been developed, including single amino-acid substitutions in the doxycycline-inducible transactivator, which help to reduce the transactivator's rigidity and increase its DNA binding affinity, and the induction of adenylate/uridylate-rich mRNA destabilizing elements into the 3' untranslated regions of tetracycline-inducible elements (Pham et al., 2008; Roney et al., 2016). The successful incorporation of any such methods in a repetition of this study could give a more fruitful and reliable *KCNQ1* expression rate.

Factors that influence CRISPRa efficiency

When testing the assembled HEK TRE-VPR system for the first time, the goal was to monitor the activation efficiencies of several genes and choose one that might be feasible to try to activate. The x2.71 and x2.77-fold increase in expression of *ASCL1* and *GHRH*, respectively, proved the correct functioning of the system, while *KCNQ1* with its minimal x1.12-fold increase required improvements in the activation system's gRNAs. After follow-up experiments showed that newly designed, different gRNAs targeting several TSSs are also ineffective in activating the gene, the question of why *KCNQ1* was so difficult to activate was posed.

As previously stated, six gRNAs, targeting four different TSSs were constructed for this experiment. Most gRNAs failed to produce the desired *KCNQ1* activation effect, with only gRNA2 eliciting statistically significant upregulation (Figure 11). As it is unclear whether the gene itself is preventing activation from occurring, some considerations need to be had regarding the effectiveness of the constructed gRNAs. Some papers suggest that the most favorable range in a gene's nucleotide sequence which CRISPRa gRNAs can target is the sequence 100nt upstream of the TSS (Fontana et al., 2020). The constructed gRNAs all target that region, but other factors influence the efficacy with which they conduct activation.

The difficulty in activating *KCNQ1* could also be related to the shortcomings of CRISPRa and reflect the advantages and disadvantages of different gene activation methods. This distinction is reflected by the varying degree of gene activation in the experiments, described in the Results section. Both *GHRH* and *ASCL1* are activated by a factor of around 2-3, in comparison to the

controls, while *KCNQ1* has an irregular and undetermined expression (Figure 6, Figure 7, Figure 9).

Additionally, before undertaking any further experiments, an understanding of chromatin and its effect on gene expression would be paramount. Chromatin remodeling can determine whether a DNA target can be accessed by the CRISPR-Cas9 complex (Jiang and Doudna, 2017). Since gene expression may thus depend on the "openness" of chromatin, it possible that this factor comes into play in the case of *KCNQ1*, where a "closed" chromatin state may block interactions with the CRISPR-dCas9 complex. Indeed, studies have implicated heterochromatic factors in the regulation of the *KCNQ1* domain, with, e.g., non-coding RNA-mediated heterochromatinization establishing repressive chromatin structures that silence *KCNQ1* (Kanduri, 2011). If these higher structures have hindered the CRISPRa system's action in the present experiment, future iterations may benefit from epigenetic editing to improve *KCNQ1*'s chromatin accessibility for the CRISPR-dCas9 complex (Kang et al., 2019).

Improving the assay

Although the goal of testing the CRISPRa HEK TRE-VPR system was ultimately achieved, some of its technical aspects could be improved upon further repetition. Having a consistent control in all experiments would provide more accuracy to the data and allow for its better readability. In this case, controls with no transfected gRNAs, as well as negative controls with a blank transfected gRNA, were used as controls. Additionally, it would be useful to compare old and newly constructed gRNAs in the same experiment. Therefore, optimally, a no template control (no transfection), a negative control (transfection with neg blank gRNA), and positive control (transfection with old gRNA) would lead to more uniform data.

Future perspectives

The expression of *KCNQ1* has been slightly upregulated, but validation via repetition of the experiment is required before any strong conclusions may be drawn. A possible further experiment, as mentioned previously, would be to examine pituitary hormone deficiency

progression in control cells and *KCNQ1* overexpressed cells, and determine whether a correlation can be made between gene overexpression and disease progression.

Bibliography

Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569–573 (2014).

Antoniou, P., Miccio, A. & Brusson, M. Base and Prime Editing Technologies for Blood Disorders. *Front. Genome Ed.* **3**, 618406 (2021).

Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).

Balboa Alonso, D. Human Pluripotent Stem Cells and CRISPR-Cas9 Genome Editing to Model Diabetes. Ph. D. University of Helsinki (2018).

Barro-Soria, R. *et al.* KCNE1 divides the voltage sensor movement in KCNQ1/KCNE1 channels into two steps. *Nat Commun* **5**, 3750 (2014).

Brezgin, S., Kostyusheva, A., Kostyushev, D. & Chulanov, V. Dead Cas Systems: Types, Principles, and Applications. *IJMS* **20**, 6041 (2019).

Brocken, D. J. W., Tark-Dame, M. & Dame, R. T. dCas9: A Versatile Tool for Epigenome Editing. *Current Issues in Molecular Biology* 15–32 (2018).

Cámara, E., Lenitz, I. & Nygård, Y. A CRISPR activation and interference toolkit for industrial Saccharomyces cerevisiae strain KE6-12. *Sci Rep* **10**, 14605 (2020).

Casas-Mollano, J. A., Zinselmeier, M. H., Erickson, S. E. & Smanski, M. J. CRISPR-Cas Activators for Engineering Gene Expression in Higher Eukaryotes. *The CRISPR Journal* **3**, 350–364 (2020).

Chavez, A. et al. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* **12**, 326–328 (2015).

Chavez, A. *et al.* Comparison of Cas9 activators in multiple species. *Nat Methods* **13**, 563–567 (2016).

Cho, S., Shin, J. & Cho, B.-K. Applications of CRISPR/Cas System to Bacterial Metabolic Engineering. *IJMS* **19**, 1089 (2018).

Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* **34**, 184–191 (2016).

Du, D. & Qi, L. S. An Introduction to CRISPR Technology for Genome Activation and Repression in Mammalian Cells. *Cold Spring Harb Protoc* **2016**, pdb.top086835 (2016).

Dunnett, C. W. A Multiple Comparison Procedure for Comparing Several Treatments with a Control. *Journal of the American Statistical Association* **50**, 1096–1121 (1955).

Ede, D. R., Farhang, N., Stover, J. D. & Bowles, R. D. 4.32 Gene Editing Tools. in *Comprehensive Biomaterials II* 589–599 (2017).

Fan, H., Zhang, M. & Liu, W. Hypermethylated KCNQ1 acts as a tumor suppressor in hepatocellular carcinoma. *Biochemical and Biophysical Research Communications* **503**, 3100–3107 (2018).

Fontana, J. *et al.* Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nat Commun* **11**, 1618 (2020).

Gil, N. & Ulitsky, I. Regulation of gene expression by cis-acting long non-coding RNAs. *Nat Rev Genet* **21**, 102–117 (2020).

Goldman, A. M. *et al.* Arrhythmia in Heart and Brain: KCNQ1 Mutations Link Epilepsy and Sudden Unexplained Death. *Sci. Transl. Med.* **1**, (2009).

Gong, L., Liu, X., Wu, J. & He, M. Emerging strategies for the genetic dissection of gene functions, cell types, and neural circuits in the mammalian brain. *Mol Psychiatry* (2021).

Hirai, H., Tani, T. & Kikyo, N. Structure and functions of powerful transactivators: VP16, MyoD and FoxA. *Int. J. Dev. Biol.* **54**, 1589–1596 (2010).

Hirakawa, M. P., Krishnakumar, R., Timlin, J. A., Carney, J. P. & Butler, K. S. Gene editing and CRISPR in the clinic: current and future perspectives. *Bioscience Reports* **40**, BSR20200127 (2020).

Jespersen, T., Grunnet, M. & Olesen, S.-P. The KCNQ1 Potassium Channel: From Gene to Physiological Function. *Physiology* **20**, 408–416 (2005).

Jiang, F. & Doudna, J. A. CRISPR–Cas9 Structures and Mechanisms. *Annu. Rev. Biophys.* **46**, 505–529 (2017).

Jo, Y.-H. & Chua, S. Transcription factors in the development of medial hypothalamic structures. *American Journal of Physiology-Endocrinology and Metabolism* **297**, E563–E567 (2009).

Kahn, R. A. & East, M. P. Generating plasmids for CRISPR/Cas9 insertion of frameshifting mutations. *ARF Consortium Blog* https://scholarblogs.emory.edu/arf-consortium/files/2018/09/CRISPR-KO-protocol-v2.pdf (2018).

Kanduri, C. Kcnq1ot1: A chromatin regulatory RNA. *Seminars in Cell & Developmental Biology* **22**, 343–350 (2011).

Kang, J. G., Park, J. S., Ko, J.-H. & Kim, Y.-S. Regulation of gene expression by altered promoter methylation using a CRISPR/Cas9-mediated epigenetic editing system. *Sci Rep* **9**, 11960 (2019).

Koonin, E. V. & Makarova, K. S. Origins and evolution of CRISPR-Cas systems. *Phil. Trans. R. Soc. B* **374**, 20180087 (2019).

Li, Z., Xiong, X. & Li, J.-F. The working dead: repurposing inactive CRISPR-associated nucleases as programmable transcriptional regulators in plants. *aBIOTECH* **1**, 32–40 (2020).

Liu, Z., Dong, H., Cui, Y., Cong, L. & Zhang, D. Application of different types of CRISPR/Cas-based systems in bacteria. *Microb Cell Fact* **19**, 172 (2020).

Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* **25**, 402–408 (2001).

Matsoukas, I. G. Prime Editing: Genome Editing for Rare Genetic Diseases Without Double-Strand Breaks or Donor DNA. *Front. Genet.* **11**, 528 (2020).

Nakajo, K. & Kubo, Y. Nano-environmental changes by KCNE proteins modify KCNQ channel function. *Channels* **5**, 397–401 (2011).

Pellegrini, R. How to Design gRNAs to Target Your Favorite Gene. *Benchtalk* https://www.benchling.com/2016/02/01/how-to-design-grnas-to-target-your-favorite-gene/ (2016).

Perez-Pinera, P. *et al.* Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nat Methods* **10**, 239–242 (2013).

Pham, D. H., Moretti, P. A. B., Goodall, G. J. & Pitson, S. M. Attenuation of leakiness in doxycycline-inducible expression via incorporation of 3′ AU-rich mRNA destabilizing elements. *BioTechniques* **45**, 155–162 (2008).

Porras, P. *et al.* Towards a unified open access dataset of molecular interactions. *Nat Commun* **11**, 6144 (2020).

Przybyla, L. & Gilbert, L. A. A new era in functional genomics screens. *Nat Rev Genet* (2021).

Qi, L. S. *et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152**, 1173–1183 (2013).

Rahman, M. M. & Tollefsbol, T. O. Targeting cancer epigenetics with CRISPR-dCAS9: Principles and prospects. *Methods* **187**, 77–91 (2021).

Roney, I. J., Rudner, A. D., Couture, J.-F. & Kærn, M. Improvement of the reverse tetracycline transactivator by single amino acid substitutions that reduce leaky target gene expression to undetectable levels. *Sci Rep* **6**, 27697 (2016).

Ross, A. & Willson, V. L. One-Way Anova. in *Basic and Advanced Statistical Tests* 21–24 (2017).

Santos-Moreno, J. & Schaerli, Y. CRISPR-based gene expression control for synthetic gene circuits. *Biochemical Society Transactions* **48**, 1979–1993 (2020).

Sar, P. & Dalai, S. CRISPR/Cas9 in epigenetics studies of health and disease. in *Progress in Molecular Biology and Translational Science* **181** 309–343 (2021).

Shakirova, K. M., Ovchinnikova, V. Y. & Dashinimaev, E. B. Cell Reprogramming With CRISPR/Cas9 Based Transcriptional Regulation Systems. *Front. Bioeng. Biotechnol.* **8**, 882 (2020).

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection. *Nature*, **409**, 685–690 (2001).

Tommiska, J. *et al.* Two missense mutations in KCNQ1 cause pituitary hormone deficiency and maternally inherited gingival fibromatosis. *Nat Commun* **8**, 1289 (2017).

Tukey, J. W. Comparing Individual Means in the Analysis of Variance. *Biometrics* **5**, 99 (1949).

Tuttle, M. & Chavez, A. CRISPR Activation: A Practical Guide. *AddGene Blog* https://blog.addgene.org/cas9-activators-a-practical-guide (2016).

Valente, F. M. *et al.* Transcription alterations of KCNQ1 associated with imprinted methylation defects in the Beckwith–Wiedemann locus. *Genetics in Medicine* **21**, 1808–1820 (2019).

Wang, J., Zhang, C. & Feng, B. The rapidly advancing Class 2 CRISPR-Cas technologies: A customizable toolbox for molecular manipulations. *J Cell Mol Med* **24**, 3256–3270 (2020).

Wilkinson, R. A., Martin, C., Nemudryi, A. A. & Wiedenheft, B. CRISPR RNA-guided autonomous delivery of Cas9. *Nat Struct Mol Biol* **26**, 14–24 (2019).

Wu, X. & Larsson, H. P. Insights into Cardiac IKs (KCNQ1/KCNE1) Channels Regulation. *IJMS* **21**, 9440 (2020).

Xu, C., Park, J.-K. & Zhang, J. Evidence that alternative transcriptional initiation is largely nonadaptive. *PLoS Biol* **17**, e3000197 (2019).

Xu, X. & Qi, L. S. A CRISPR–dCas Toolbox for Genetic Engineering and Synthetic Biology. *Journal of Molecular Biology* **431**, 34–47 (2019).

Zrimec, J., Buric, F., Kokina, M., Garcia, V. & Zelezniak, A. Learning the Regulatory Code of Gene Expression. *Front. Mol. Biosci.* **8**, 673363 (2021).