

Gene circuit integration in induced pluripotent stem cells for the study of neurodegeneration



Federico Caso
Wolfson College
University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2024

Gene circuit integration in induced pluripotent stem cells for the study of neurodegeneration

Federico Caso, Wolfson College, DPhil in Clinical Medicine,

Trinity 2024

Abstract

The intersection of synthetic biology and human induced Pluripotent Stem cells (iPSC) technology presents a promising avenue for the exploration and treatment of multifactorial disorders. Synthetic biology, through the integration of exogenous gene networks known as gene circuits, enables cellular reprogramming, allowing control over cell behaviour. In parallel, iPSC technology has enabled the development of research models for human cell types that are challenging to obtain, such as neurons, providing valuable insights into the genetic, molecular, and intercellular mechanisms underlying complex disorders.

Neurodegenerative disorders, particularly Alzheimer's disease (AD), are garnering attention due to their rising prevalence in the aging population as well as their complex pathology. Despite the association of molecular drivers, such as Amyloid β ($A\beta$) and tau, a comprehensive understanding of its aetiology and pathology remains elusive, impacting the efficacy of current therapies and the identification of new targets.

The combination of synthetic biology and iPSC technology could offer a promising avenue to investigate distinct cellular mechanisms involved in AD neurodegeneration. However, this is a relatively novel approach with considerable limitations to overcome.

Firstly, the available synthetic biology tools for eukaryotic cells are still limited, resulting in less optimized gene circuits. Secondly, the epigenetic modification of the iPSC genome makes it difficult to design and predict the behaviour of gene circuits at different cell states, resulting in often ineffective outputs.

To address these challenges, I designed a gene circuit activated by extracellular $A\beta$, inducing the expression of potentially neuroprotective genes. This circuit served as both a testing tool to troubleshoot and optimize the above limitations and as a proof-of-concept for the integration of a gene circuit within iPSCs.

In the initial phase, I explored various components of a eukaryotic gene circuit. Through this investigation, I selected an optimal GPCR-based mechanism for sensing extracellular $A\beta$, which

triggers gene expression changes induced via CRISPR activation (CRISPRa). This necessitated an investigation of several inducible promoter systems for regulated sgRNA expression, a method less well characterised compared to inducible gene expression of mRNA. Validating a viable inducible system provided valuable insights on how CRISPR systems can be effectively integrated into larger gene circuits within eukaryotic cells.

Additionally, I investigated the use of the human synapsin promoter to ensure targeted gene expression in iPSC-derived neurons when integrated into the adeno-associated virus integration site 1 (AAVS1) safe harbour locus. The results demonstrated that the epigenetic instability of AAVS1 throughout differentiation can be mitigated when combined with cell-type-specific promoters.

While the final circuit was ultimately ineffective, confirming the complexity of establishing such regulated gene networks within eukaryotic cells, I developed an iPSC line able to induce the expression of CRISPRa only in the neuronal state. As a proof-of-concept experiment, I used this cell line to investigate the impact of Msh Homeobox 2 (*MSX2*) upregulation on the transcriptome of iPSC-derived neurons. The results suggest a potential neuroprotective function, warranting further investigations into possible downstream molecular pathways that could play a therapeutic role against AD.

The development of this cell line provides a new tool for inducing gene expression changes in iPSC-derived neurons and describes important steps in optimizing the intersection of iPSC technology and synthetic biology for the study of complex disorders such as neurodegeneration.

Statement of Authorship

The data and their interpretation are the product of my own work unless stated otherwise.

Experiments were performed at the Oxford Centre for Human Genetics (CHG) under the supervision of Dr. Ben Davies. Funding for the project was provided by the Medical Research Council and the Nuffield Department of Medicine as part of the DPhil in Clinical Medicine.

Karyotyping of genome-engineered iPSC lines was performed by Dr Daniela Moralli at the Chromosome Dynamics Core (CHG) (Chapters 2 and 5).

Sequencing data, including RNA-Seq libraries (Chapters 2 and 6), were generated with the Technology Platforms Group (CHG). Dr Benjamin Wright provided support with the RNA-Seq data analysis and carried out the gene set enrichment analysis, while Dr Matthieu Mossieć carried out the GeneFeast analysis.

HHEX and NT lentiviruses have been generated by Dr Ping Zhang (CHG) (Chapter 2 and 5).

None of the material included in this thesis was submitted as part of any other degree or professional application.

This thesis consists of approximately 40,000 words.

Acknowledgements

First and foremost, I would like to thank my supervisor Dr Ben Davies, for inviting me to his laboratory and allowing me to pursue the scientific questions that I found most fascinating. His scientific rigour and management have helped me grow as a person and a scientist. I will always be thankful for his patience and support throughout this journey.

I am equally grateful to Professor Noel Buckley and Professor John Todd, for sharing their expertise and insights, helping me shape the trajectory of this project.

For creating a work environment that supported me through the tumultuous times of my PhD, I would like to express my gratitude to all the members of the Davies lab: Dr Phalguni Rath, for his steadfast support of my ideas and goals; Polinka Hernandez, Daniel Biggs, Dr Emma Jaeger, and Chris Preece for treating me as a member of the team and sharing their valuable knowledge; Katie Holmes, Cherie Lai and Greg Holt for all the conversations over coffee during the “ups” but especially the “downs” of the PhD.

Additionally, I would like to thank the Nucleate UK team for showing me that the value of impactful work is even more meaningful when it is carried out to the benefit of other. While there are too many names to thank here, I am grateful to have had the pleasure to work with each one of you. A special thank you to James Bayne, for inviting me onto this memorable project with him, through which I have grown more than I thought I ever could, opening up new exciting prospects for my career going forward.

I would like to finish by thanking my parents Valeria and Luigi for our weekly calls but constant encouragement, my brothers for providing levity when I needed it most and my flatmates – Rob Kerr, Alexis McGivern, Raffy Maristela and Emre Yildirim– for providing a home away from home. Above all, I owe my resilience and gratitude to Rebecca Shaw for being my supporter and inspiration for almost a decade. Thank you for accompanying me in this journey.

List of Abbreviations

AAVS1	Adeno-associated Virus site 1
AD	Alzheimer's disease
APP	A β precursor protein
Aβ	Amyloid β
BMP	Bone morphogenetic protein
BP	Biological Processes
CC	Cellular Component
CLU	Clusterin
CLYBL	Citramalyl-CoA lyase
CMKLR1	Chemerin chemokine-like receptor 1
CMV	Cytomegalovirus promoter
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRISPRa	CRISPR activation
CRISPRa-SAM	CRISPRa - Synergistic Activation Mediator
CRISPRa-VPR	CRISPRa - VP64 -p65 -Rta
CRISPRi	CRISPR inactivation
crRNA	CRISPR RNA
dCas9	dead Cas9
FAD	Familial Alzheimer's disease
GO	Gene Ontology
GPCR	G-Protein Coupled receptor
GSEA	Gene Set Enrichment Analysis
HDR	Homology Directed Recombination
HEK293T	Human Embryonic Kidney cells 293T
HHEX	Haematopoietically expressed homeobox 2
hSyn	Human synapsin promoter
HTLA	HEK293T with tTA luciferase activation
iPSC	induced Pluripotent Stem Cells
KRAB	Kruppel associated box
LOAD	Late Onset Alzheimer's disease
MF	Molecular Function
MSX2	Msh Homeobox 2
NFT	neurofibrillary tangles
NHEJ	Non-Homologous End Joining
PSEN	Presenilin
RLU	Relative luciferase units
rtTA	Reverse tetracycline transactivator
sgRNA	single-guide RNA
TetO	Tetracycline Operator
TetR	Tetracycline Repressor
TEV	Tobacco etched virus protease
TEVcs	TEV cutting site
TGF-β	Transforming growth factor β receptor

tracrRNA	trans-activating CRISPR RNA
TRE	Tetracycline responsive element promoter
tTA	Tetracycline transactivator
βA-TEV	β-arrestin2-TEV

Table of Contents

Abstract	2
Statement of Authorship.....	4
Acknowledgements	5
List of Abbreviations	6
Table of Contents	8
List of Figures	12
List of Tables	13
Chapter 1 Introduction	14
Synthetic Biology – A new cell engineering approach	14
From bacterial toggle switches to mammalian gene circuits.....	14
Tetracycline repressor a core component of mammalian circuits	18
GPCR-based sensors in mammalian gene circuits	20
CRISPR as a DNA Regulator	24
CRISPR Discovery and first applications	24
CRISPRi and CRISPRa.....	27
Inducible CRISPR activation	30
Alzheimer’s disease – health challenge of the decade	34
AD aetiology and modelling	34
iPSC-derived neurons	38
iPSC-derived neurons as model for Alzheimer’s disease	40
CRISPRa/i and iPSC-derived neurons	42
Thesis objectives and rationale	45
A feedback-controlled gene circuit for the study of AD	45
Design of the gene circuit.....	45
Thesis Outline	47
Chapter 2 Materials and Methods	49
Cell Lines.....	49
Bacteria	49
Cell lines	49
iPSC line.....	49
Guide RNA design and Molecular Cloning	49
In vitro gRNA design.....	49

Golden Gate Assembly	49
Ligation-Dependent Cloning.....	50
Bacterial Cloning.....	51
DNA extraction.....	51
PCR and Gel Electrophoresis.....	51
DNA Purification and Sequencing	52
RNA Extraction, cDNA synthesis and Quantitative PCR	53
DNA and RNA quantification	53
Immunostaining.....	53
Lentiviral Production	54
RNA sequencing and statistical analysis	55
RNA Library preparation and Sequencing.....	55
RNA Sequencing Analysis	55
Differential Gene Expression Analysis	55
Gene Set Enrichment Analysis.....	56
Experimental Statistical Analysis	58
Cell Culture and Manipulation	58
HEK293T and HTLA Cells Maintenance.....	58
iPSC Maintenance.....	58
Cell cryopreservation	59
Cell thawing.....	59
iPSC flow cytometry for pluripotency assay.	59
iPSC Karyotyping.....	59
Plasmid DNA Lipofection	60
RNA Transfection	60
Luciferase Assay	60
GFP Assay	60
RNP Electroporation.....	61
Low seeding and colony picking.....	61
Cortical Neurons Differentiation	61
Viral Transduction	62
Chapter 3 Validation of the core circuit components	64
Introduction	64
Validation of the Tango system	65
Validation of CRISPRa-SAM sgRNAs	68
Approaches for inducible guide expression	70

Pol III – TetO system	71
tRNA-TetO system.....	74
Csy4 – TetO system	77
Closing Remarks	80
Chapter 4 Design and validation of the final circuit.....	84
Introduction	84
Design of the BATV2 plasmid	85
Validation of BATV2 in HEK293T cells.....	92
Design of the CLICK4 plasmid.....	93
Validation of the CLICK4 plasmid	101
Closing Remarks	104
Chapter 5 Validation of hSYN-driven CRISPRa-SAM in iPSC-derived neurons	106
Introduction	106
Targeting of SAM-BATV2 into AAVS1 Safe Harbor Locus	107
Quality Control of KOLFBAT Hom and KOLFBAT Het.....	110
Validation of the Integrated CRISPRa-SAM in Pluripotent Stem Cells.....	111
Differentiation of KOLFBAT/SAM into Cortical Neurons	113
Closing Remarks	121
Chapter 6 Effect of MSX2 upregulation on the transcriptome of iPSC-derived neurons ...	122
Introduction	122
MSX2 sgRNA transduction, RNA sequencing and quality control	123
RNA sequencing quality control	124
Differential analysis of 48 hours and 10 days samples	126
Gene Set Enrichment Analysis	132
GeneFeast to visualise gene enrichment analysis	135
Discussion on differential analysis	137
Closing Remarks	140
Chapter 7 Discussion	141
Results Summary	142
CMKLR1-Tango can be a suitable sensing mechanism for A β	144
Limitations of Pol III-tTA and tRNA-tTA inducible systems	146
tTA-Csy4 inducible system is effective but flawed in the context of CLICK4	149
Human synapsin promoter is suitable to overcome AAVS1 silencing in iPSC-derived neurons.	151
MSX2 upregulation KOLFBAT-derived neurons.....	154
The circuit as a model for Alzheimer's and further therapeutic applications	158

Conclusion and Future directions	162
Bibliography	164

List of Figures

Figure 1.1 Diagram of original gene circuits	16
Figure 1.2 Diagram of logic gates used in gene circuit engineering..	18
Figure 1.3 Diagram of Tet systems for circuit engineering.	19
Figure 1.4 Diagram of the PRESTO-Tango system.....	22
Figure 1.5 Diagram of GPCR-inducible CRISPR systems..	23
Figure 1.6 Diagram of the natural CRISPR defence mechanism	25
Figure 1.7 Diagram of Non-homologous end joining (NHEJ) and homology directed recombination (HDR) caused by CRISPR /Cas9 cleavage of the DNA.	26
Figure 1.8 Diagram of CRISPRi and CRISPRa systems.....	29
Figure 1.9 Diagram of APP processing.	35
Figure 1.10 Diagram of gene circuit design for the sensing of A β and the expression of neuroprotective gene.....	47
Figure 2.1 Schematic of gene enrichment score analysis.	57
Figure 2.2 Schematic of iPSC derived differentiation of cortical neurons.	62
Figure 3.1 Diagram of the core components of the circuit..	64
Figure 3.2 CMKLR1-Tango analysis..	67
Figure 4.1 Diagram of stepwise circuit activation..	84
Figure 4.2 Diagram illustrating the two genetic sequences composing the complete circuit.. ..	85
Figure 4.3 Diagram illustrated the inserted exogenous sequence into the first intronic region of AAVS1.....	86
Figure 4.4 Assembly 1 of BATV2.....	87
Figure 5.1 Integration of BATV2 in KOLF2-C1 iPSC.	108
Figure 5.2 Gel electrophoresis of KOLFBAT..	109
Figure 5.3 Quality control of KOLFBAT.	111
Figure 5.4 Investigation of Cas9 and β A expression in KOLFBAT..	112
Figure 5.5 Investigation of CRISPRa-SAM expression in pluripotent KOLFSAM and KOLFBAT. .	113
Figure 5.6 Neuronal differentiation of iPSC. Timeline of iPSC neuronal differentiation.....	114
Figure 5.7 Investigation of neuronal markers across differentiation.	116
Figure 5.8 Lentiviral transduction in pluripotent iPSC..	117
Figure 5.9 Lentiviral transduction in iPSC-derived neurons.....	118
Figure 5.10 Comparison of CRISPRa-SAM driven gene expression between pluripotent and neuronal cell state.....	119
Figure 6.1 MSX2 upregulation at 48 hours and 10 days timepoints.	124
Figure 6.2 Quality Control analyses of RNA sequencing results..	125
Figure 6.3 Dispersion plots for the 48 hours and 10 days datasets.	126
Figure 6.4. Volcano plots of differentially expressed genes 48 hours and 10 days after MSX2 transduction.	128
Figure 6.5 Heatmap of the top and bottom 25 differentially expressed genes in the 48 hours timepoint analysis.....	130
Figure 6.6 Heatmap of the top and bottom 25 differentially expressed genes in the 10 days timepoint analysis.....	131
Figure 7.1 Diagram representing the components of the A β -activated gene circuit..	142
Figure 7.2 Design of gene circuit with Csy4 tethered to CMKLR1-Tango.....	150

List of Tables

Table 2.1 E.coli cell lines	49
Table 2.2 List of sgRNA for CRISPR and CRISPR-SAM. The list shows the sgRNA sequences utilised in this project. Below the oligonucleotides sequences with the appropriate BbsI overhangs are detailed. Lowercase base indicates the addition of a guanine at the 5' end and respective cytosine at the 3' end. Added to enhance efficiency (Hsu et al., 2013)	50
Table 2.3 General Immolase PCR reaction.....	52
Table 2.4 General Thermal Cycling Program.....	52
Table 2.5 Exol-SAP PCR purification protocol.....	53
Table 2.6 List of Primary and Secondary antibodies.....	54
Table 2.7 Reagents of Neuronal Maintenance Medium	62
Table 6.1 Top 20 positively enriched terms in the 48-hours dataset (top) and 10-days dataset (bottom). The “Source” column records the Gene Ontology (GO) ID for each term. The “Ont” column records the gene ontology category: cellular component (CC), molecular function (MF), biological process (BP). The estimation score (ES) indicates the maximum deviation from 0 and it represents the degree to which a gene set is represented at the extreme of the rank list generated from the 48 hours or 10-days dataset. The normalised estimation score (NES) is equal to the value of each ES divided by the average of the enrichment score of all the dataset permutations performed by the software. Adjusted p-value was < 0.01. Adjusted p-value was based on the Benjamini-Hochberg false discovery method to each NES value, generating an adjusted p-value (padj). In the GSEA report, a padj value of zero (0.0) indicates an actual p value of less than 1/number-of-permutations.....	132
Table 6.2 Top 20 negatively enriched terms in the 48-hours dataset (top) and 10-days dataset (bottom). The “Source” column records the Gene Ontology (GO) ID for each term. The “Ont” column records the gene ontology category: cellular component (CC), molecular function (MF), biological process (BP). The estimation score (ES) indicates the maximum deviation from 0 and it represents the degree to which a gene set is represented at the extreme of the rank list generated from the 48-hours or 10-days dataset. The normalised estimation score (NES) is equal to the value of each ES divided by the average of the enrichment score of all the dataset permutations performed by the software. Adjusted p-value was < 0.01. Adjusted p-value was based on the Benjamini-Hochberg false discovery method to each NES value, generating an adjusted p-value (padj). In the GSEA report, a padj value of zero (0.0) indicates an actual p value of less than 1/number-of-permutations.....	134
Table 6.3 GeneFeast metacommunities. Metacommunities for the positively (01 and 02) and negatively (03 and 04) terms in the 10-days datasets. Metacommunities were grouped based on the GeneFeast algorithm through analysis of similar genes associated to the terms. The “Ont” column records the gene ontology category: cellular component (CC), molecular function (MF), biological process (BP).	136

Chapter 1 Introduction

Synthetic Biology – A new cell engineering approach

From bacterial toggle switches to mammalian gene circuits

The term “synthetic biology” was first conceptualized by Hobom (1980) as a synonym for genetic engineering in biology. With the advent of molecular cloning and the establishment of techniques such as transcriptomics and proteomics, molecular biology shifted toward utilizing gene engineering as a core tool in its scientific discovery process. As the understanding of genome, transcriptome, and proteome interactions progressed, genetic engineering applications increased in complexity, transitioning from the perturbation of single genes to the simultaneous manipulation of several genes (Cuccato et al., 2009).

The escalating complexity of these interactions led to a paradigm shift in the genetic engineering field, moving from a "top-down" to a "bottom-up" approach. While the former involved small genetic modifications to observe changes in the cell, the latter combined isolated and minimal elements, such as genes, to create specific networks, aiming to recapitulate natural behaviours. To better understand how synthetic biology developed into its current form, it is worth revisiting the key milestones that defined it.

The development of the first genetic toggle switch in *E.coli*, is considered the first example of logic design in synthetic biology (Gardner et al., 2000). The study proposed the design for an artificial toggle switch where the activation or repression of a gene could be dictated by a controllable input. The switch was composed of two repressors and two promoters. One promoter expressed the repressor for the second promoter and vice versa. With this design the expression of the gene of interest was dictated by the two opposing states, P1 over P2 or P2 over P1. Specifically, a *LacI* repressor was put under the control of a P_{LS1con} promoter – inhibited by the lambda repressor. On the other hand, the lambda repressor was put under the control of the P_{trc-2} promoter – inhibited by *LacI*. The switch between the two promoters was achievable through exogenous Isopropyl β-D-1-thiogalactopyranoside (IPTG), which triggers *LacI* detachment from P_{trc-2} allowing for the expression of the lambda repressor which in turns inhibits the P_{LS1con} promoter and the expression of *LacI*. Alternatively, a thermal pulse causing the lambda repressor to detach from P_{LS1con} results in *LacI* expression, which represses the P_{trc-2} promoter (**Fig. 1.1A**). The toggle switch proposed a novel approach to gene expression by responding to external chemical and temperature stimuli, demonstrating that cellular behaviour could be controlled through targeted inputs.

Subsequent advancements included designing more complex behaviours, such as periodic oscillation of the toggle switch (Elowitz and Leibler, 2000) and autoregulation through a negative feedback loop (Becskei and Serrano, 2000) (**Fig. 1.1B/C**). As these synthetic gene networks developed, the field borrowed terminology and concepts from mathematics and electronic engineering to improve their modelling and prediction. Eventually, the term "genetic circuits," initially coined by Monod and Jacob (1961) to describe gene networks, was adopted to characterize the synthetic interactions between genes.

Shortly after the development of gene circuits in bacteria, the field demonstrated the ability to create gene circuits in mammalian cells as more genetic tools were adapted for eukaryotic cells. Kramer et al. (2004) introduced the first toggle-switch in Chinese hamster ovary cells (CHO), utilizing antibiotic-inducible repressors created by combining bacterial repressors (pristinamycin and erythromycin repressors) with a eukaryotic repressor domain (KRAB) – PIP-KRAB and E-KRAB, respectively. In this design, E-KRAB inhibits the expression of PIP-KRAB and the glycoprotein SEAP (secreted alkaline phosphatase). Addition of erythromycin prevents E-KRAB from binding to the promoter allowing for the expression of SEAP and PIP-KRAB, which in turn, is able to inhibit E-KRAB expression. The addition of pristinamycin reverse the process, demonstrating the generation of a toggle switch in a eukaryotic system, regulated by chemical stimuli (**Fig. 1.1D**).

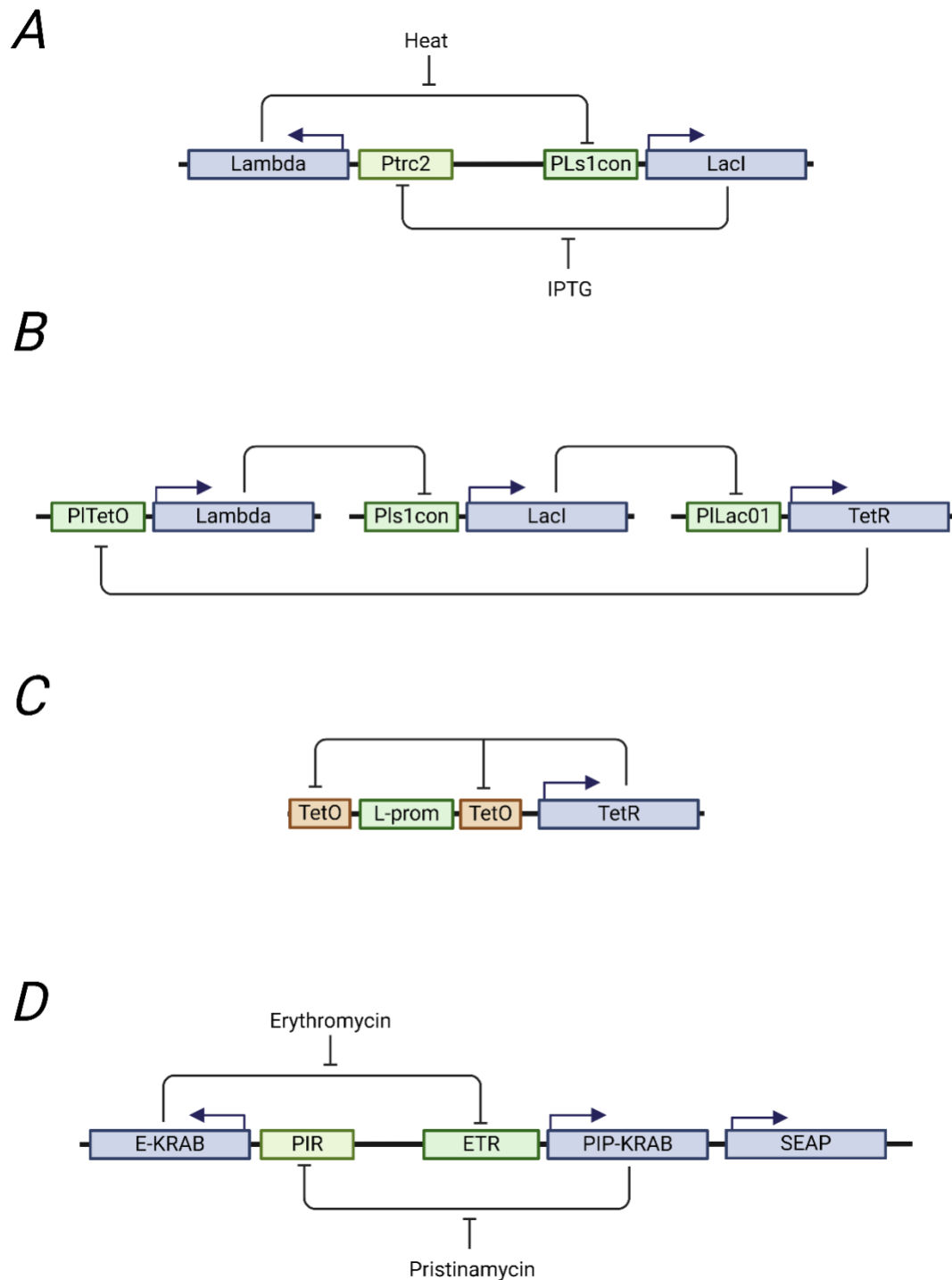


Figure 1.1 Diagram of original gene circuits **A)** Bacterial toggle switch. The switch is regulated by two repressors: *LacI* and *Lambda*. The repressors are expressed by promoters (*P_{trc-2}* and *P_{Ls1con}*) regulated by the opposite repressors. Expression of *LacI* inhibits the *P_{trc-2}* promoter silencing the expression of *Lambda*. Addition of *IPTG*, prevents *LacI* inhibition of *P_{trc-2}* which activates *Lambda* expression. *Lambda* inhibits *P_{Ls1con}* silencing *LacI*. The process can be reversed by exposure to heat that prevents *Lambda* inhibition of *P_{Ls1con}*. **B)** Oscillation toggle switch. The design follows the original toggle switch, however the *TetR* repressor and *P_{ITetO}* promoters are added. As before the three repressors are expressed by promoters inhibited by another repressor, resulting in unstable and oscillating system where the expression of each repressor can impact the rest of the system cyclically. **C)** Autoregulatory circuit with negative feedback. A *TetR* is expressed from a *L-prom* containing two *TetO* sequences. Higher rates of *TetR* expression result in a stronger silencing of the promoter, reducing the *TetR* expression which in turn results in lower silencing of the promoter. **D)** Eukaryotic toggle switch. The design follows the original bacterial switch; however, the

bacterial repressors are swapped for antibiotic-inducible repressors. Erythromycin is used to control the repression of the ETR promoter driving the expression of the pristinamycin repressor and pristinamycin to control the repression of the PIR promoter driving the expression of the erythromycin repressor. This design was also used to control the expression of SEAP, under the ETR promoter demonstrating that the toggle switch can be used to regulate the expression of genes outside the core circuit components

The second wave of synthetic biology was defined by an expansion of the available tools, facilitating increased modularity alongside the growing complexity of circuits (Purnick and Weiss, 2009). Beyond the initial ON/OFF logic gate, also referred to as Boolean, new design rules were devised to enhance control over intended cellular behaviours. For instance, OR gates could be activated by either of two different inputs, while AND gates required both inputs for circuit activation (Kis et al., 2015) (**Fig. 1.2**). These designs progressed into more intricate configurations resembling coding commands, including NAND (NOT + AND), NOR (NOT + OR), and various others (Ausländer et al., 2012, Miyamoto et al., 2013) (**Fig. 1.2**).

Eventually, the iteration of these gene circuits in mammalian cells led to the establishment of common components to regulate cell behaviour, such as (TetR) for inducible gene expression (Gossen and Bujard, 1992), synthetic GPCR pathways to sense external inputs (Kroeze et al., 2015), and CRISPR for precise gene editing and manipulation (Cong et al., 2013, Mali et al., 2013).

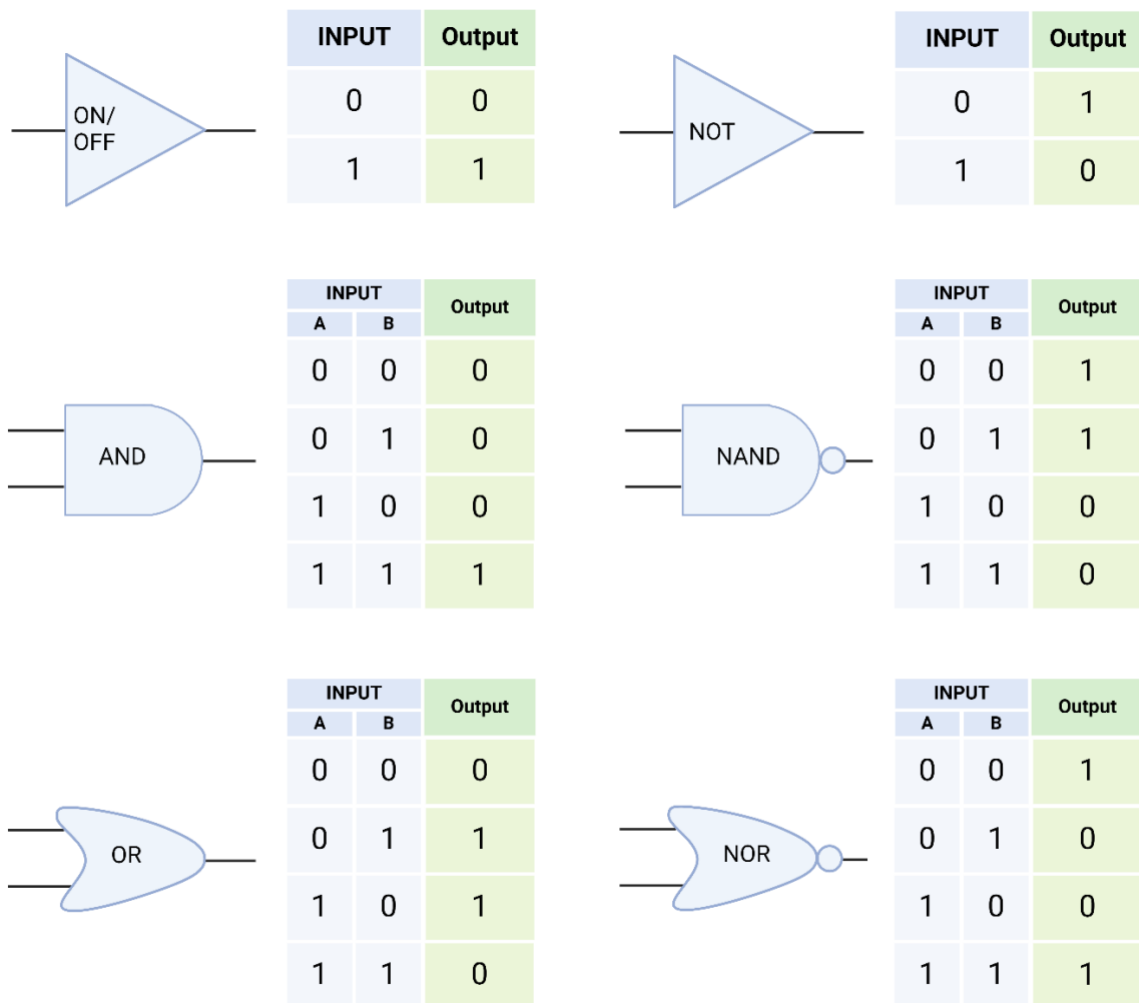


Figure 1.2 Diagram of logic gates used in gene circuit engineering. The inputs (blue column) are shown in binary 1=presence of input, 0= absence of input. Output (green column) represents the expected output based on the logic gate, also in binary. 1 = presence of output, 0 = absence of output.

Tetracycline repressor a core component of mammalian circuits

Among the components that contributed to the high degree of modularity in synthetic biology, the tetracycline transactivator system has emerged as one of the most reliable and flexible regulators in mammalian cells (Gossen and Bujard, 1992, Gossen et al., 1995, Gossen and Bujard, 2001, Kitada et al., 2018).

This system is naturally expressed in *E.coli* to confer selective resistance against the tetracycline antibiotic (Gossen and Bujard, 1992). In the absence of tetracycline, the gene is suppressed by the tetracycline repressor molecule (TetR) which binds to the tetracycline operator sequence (TetO), located within the promoter. TetR binding to the promoter sequence prevents transcription, effectively silencing the gene. The presence of tetracycline in the cell triggers a

conformational change of TetR which induces its release from TetO, allowing for the expression of the resistance gene (Ryding et al., 2001) (**Fig. 1.3A**).

In the seminal study by Gossen and Bujard (1992), the tetracycline responsive element was successfully integrated into HeLa cells. This was achieved by fusing TetR to the C-terminal domain from the VP16 transactivating protein, found in the herpes simplex virus (HSV), to induce gene transcription in mammalian cells. When combined with TetR, the resulting tetracycline transactivator (tTA) interacts with the TetO and recruits the polymerase complex. This system was combined with a TetO embedded into a minimal cytomegalovirus promoter (CMV), which remains virtually silent without the support of tTA. Thus, Gossen and Bujard demonstrated the ability to induce gene expression downstream of the CMV+TetO promoter and silence it with the addition of tetracycline, later referred to as Tet-Off system (**Fig. 1.3B**).

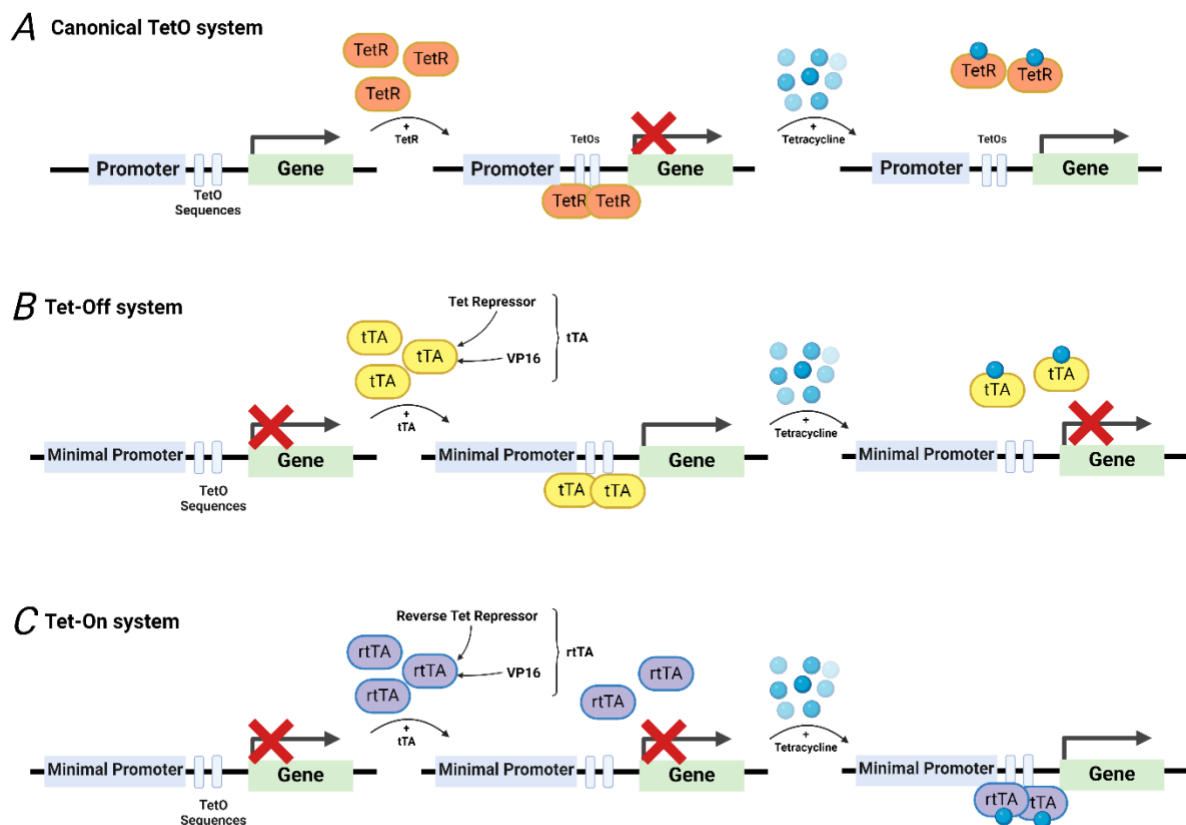


Figure 1.3 Diagram of Tet systems for circuit engineering. **A)** Original Tet system. This system is naturally expressed in *E.coli*. Tet repressor proteins (TetR) bind to Tet operator sequences (TetO) embedded within the promoter expressing the tetracycline resistance gene. Presence of tetracycline induces a conformational change that prevents TetR from binding to the TetO allowing the expression of the resistance gene. **B)** Tet-Off system. TetR is fused to VP16 (tTA) and, when bound at its operon, is able to recruit and activate the eukaryotic polymerase. TetO sequences are embedded within a minimal CMV promoter which is silent when tTA remains unbound. Expression of tTA allows binding to the TetO sequences, thus activating gene expression. The presence of tetracycline induces a conformational change in the tTA releasing it from the TetO sequence thus silencing gene expression. **C)** TetOn system. The TetR was modified to require tetracycline to bind to the TetO sequence (reverse TetR). VP16 is fused to reverse TetR (rtTA). Without tetracycline rtTA

does not interact with the TetO sequence hence the minimal CMV promoter is unable to induce gene expression. Addition of tetracycline induces the conformational change allowing rtTA to interact with the TetO, recruiting the polymerase system and thus inducing the expression of the gene.

Since the inception of the tetracycline inducible system, numerous variations have been developed. Gossen et al. (1995) later created a reverse tetracycline transactivator (rtTA) by mutating amino acids in the TetR component involved in tetracycline binding. This reverse system (later named Tet-On) only activates gene expression in the presence of both rtTA and tetracycline (**Fig. 1.3C**). While the original Tet-On and Tet-Off served as key proof-of-concept for regulatory systems in mammalian cells, they were also associated with a degree of leakiness that impacted overall efficiency (Gossen and Bujard, 1992, Gossen et al., 1995, Ryding et al., 2001).

Since their development, these systems have undergone optimization through various methods. Urlinger et al. (2000) used *S. cerevisiae* as a screening platform for rtTA mutations, selecting those that improved signal amplification, decreased background noise, and enhanced transcriptional activity. Subsequently, viral evolution screening was applied to select for further mutations (Das et al., 2004, Zhou et al., 2006, Das et al., 2016), which enhanced transcriptional activity of the rtTA or through systematic mutations of the CMV promoter to minimize background expression (Loew et al., 2010).

Takara Bio has further optimized Tet-On and Tet-Off through subsequent iterations, focusing on improved expression in human cells. These advancements involve removing cryptic splice sites from the tTA mRNA sequence, incorporating a human codon sequence, and optimizing the CMV+TetO combination by adding seven TetO sequences within the CMV promoter, creating an inducible promoter, known as the Tetracycline Response Element (TRE). The latest generation, Tet-On 3G, includes additional mutations to the rtTA, increasing sensitivity to doxycycline (a version of tetracycline), and further optimizing the TRE promoter to reduce background noise (pTRE3G).

The high modularity of the tetracycline-responsive systems, combined with their extensive optimization for mammalian cells, has made them core components in the design of mammalian gene circuits.

GPCR-based sensors in mammalian gene circuits

As circuit complexity increased, several circuit designs incorporated the use of synthetic receptors to sense the intended input in the environment, moving away from direct interaction with the genes as seen in the original toggle switches (Becskei and Serrano, 2000, Gardner et al., 2000, Elowitz and Leibler, 2000).

Mammalian cells inherently utilize a diverse array of receptors and signalling pathways to interact with their environment, providing a rich pool of components that can be redesigned for gene circuit applications (Auslander and Fussenegger, 2016). The G-protein-coupled receptors (GPCRs) represent the largest receptor family in humans, playing a pivotal role in various cellular responses by interacting with inputs ranging from small molecules to light (Teixeira and Fussenegger, 2023).

Despite the vast diversity in ligand binding within the GPCR family, they share a relatively conserved signalling pathway through interaction with the $\alpha\beta\gamma$ G-protein subunits. Ligand binding induces a conformational change that prompts the exchange of GDP to GTP on the α subunit, causing its dissociation from the $\beta\gamma$ subunits and subsequent interaction with the target signalling pathway (Wolfe and Trejo, 2007). Signal termination is regulated by the arrestin protein family, inducing dissociation between the G protein and the ligand through cytoplasmic interaction with the phosphorylated GPCR, triggering clathrin-dependent endocytosis (Wolfe and Trejo, 2007). The small and relatively conserved arrestin family comprises four molecules, arrestin-1 and 4 are confined to rods and retinal cones, while arrestin-2 and 3, also referred to as β -arrestin-1 and 2, are ubiquitously expressed and functionally overlapping as reflected in their highly homologous amino acid sequences (Jean-Charles et al., 2017). The consistency of the GPCR signal termination pathway, the large number of receptors and the similarity of β -arrestin-1 and 2, made the GPCR signalling pathway a highly suitable component as a synthetic tool for sensing circuits.

Initially established by Barnea et al. (2008) and further developed by Kroeze et al. (2015), the PRESTO-Tango system was created by modifying the GPCR termination pathway (**Fig. 1.4**). The GPCR of choice is fused to a tTA via a Tobacco Etched Virus protease cleavage site (TEVcs). Between the GPCR and the TEVcs the C-terminus of the V_2 vasopressin receptor was added to promote the recruitment of β -arrestin-2. Activation of the receptor through the correct ligand-binding interaction leads to the recruitment of β -arrestin-2, which is fused to a Tobacco Etched Virus protease (TEV) (β A-TEV). The proximity of the TEV protease to its cleavage site enables the release of tTA, triggering the modulation of target gene expression from a synthetic downstream pathway. The original PRESTO-Tango system incorporated a luciferase reporter gene regulated by a TRE promoter, as a cellular read-out. In this platform, different compounds could be tested on

any known GPCRs, with luciferase expression serving as the signal indicating the interaction with the Tango GPCR and the compound of interest (Kroeze et al., 2015).

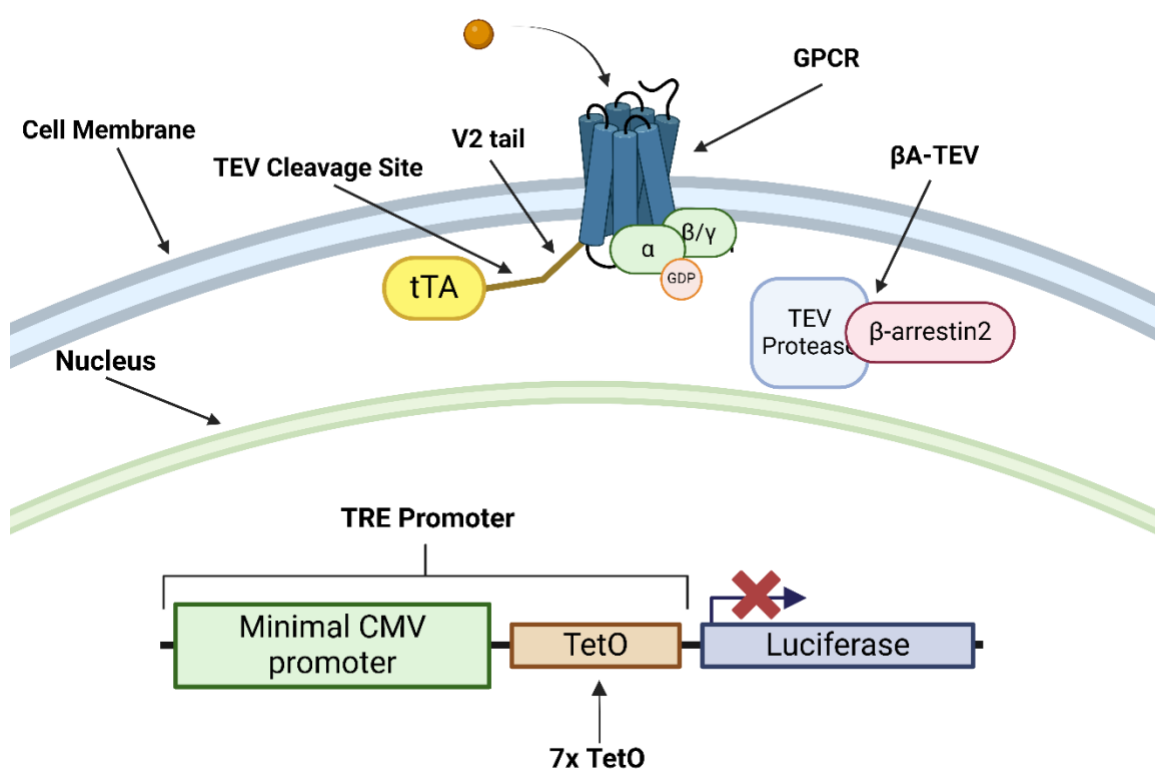


Figure 1.4 Diagram of the PRESTO-Tango system. A tTA is fused to a GPCR via a TEV cleavage site. Activation of the GPCR eventually recruits β -Arrestin2 fused to the TEV protease. Proximity between the TEV protease and the TEV cleavage site induces the release of tTA. tTA is then able to travel into the nucleus and interact with the TetO sequence of a TRE promoter allowing the expression of a reporter gene such as luciferase

The success of the PRESTO-Tango allowed for consistent optimization, leading to the recent development of the Tango-Trio. This latest version combines the recruitment of β -arrestin-1 and 2 as well as the inclusion of the internalization domain FYVE (named after the four proteins in which it was initially discovered: Fab1, YOTB, Vac1, and EEA1). The combination of these 3 components allowed for the detection of constitutive activation of the GPCR, variation in the recruitment of β -arrestin-1 and 2 as well as independent endocytosis (Zeghal et al., 2023).

Finding the appropriate ligand for an orphan GPCR can significantly expand the range of druggable receptors (Roth, 2013). With these additional measurements, the Tango-Trio platform allowed further investigation of orphan GPCRs, accounting for independent activation and inactivation profiling the ligands for 50 orphan GPCRs, demonstrating the value of this system in deorphanizing these receptors (Zeghal et al., 2023).

While the Tango system is valuable for discovering novel GPCR-ligand interactions, it can also enhance gene circuits by serving as the input sensor module within larger constructs. In two studies, the tTA of the original Tango system was replaced with a nuclease-dead Cas9 (dCas9)

component of a CRISPR activation machinery (Baeumler et al., 2017, Kipniss et al., 2017). They further incorporated logic gates by modifying the original system in order to increase control over the activation of CRISPRa. Kipniss et al. (2017), swapped the tTA for the TEV protease and fused dCas9 for the CRISPRa-VPR system to the β A molecule. Activation of the GPCR recruited β A-dCas9 which was released by the TEV protease tethered to the receptor (**Fig. 1.5A**). Baeumler et al. (2017), split the dCas9 into its N and C-terminus and tethered them to the GPCR instead of tTA. This method demonstrated the ability to control CRISPRa activation through an AND gate where both receptors had to be activated to release each half of the dCas9 and thus induce the upregulation of the gene of interest (**Fig. 1.5B**). The specificity of GPCRs combined with the logic gates permits the development of more sophisticated mechanisms demonstrating that the Tango system can serve as a suitable tool for circuit activation in mammalian cells.

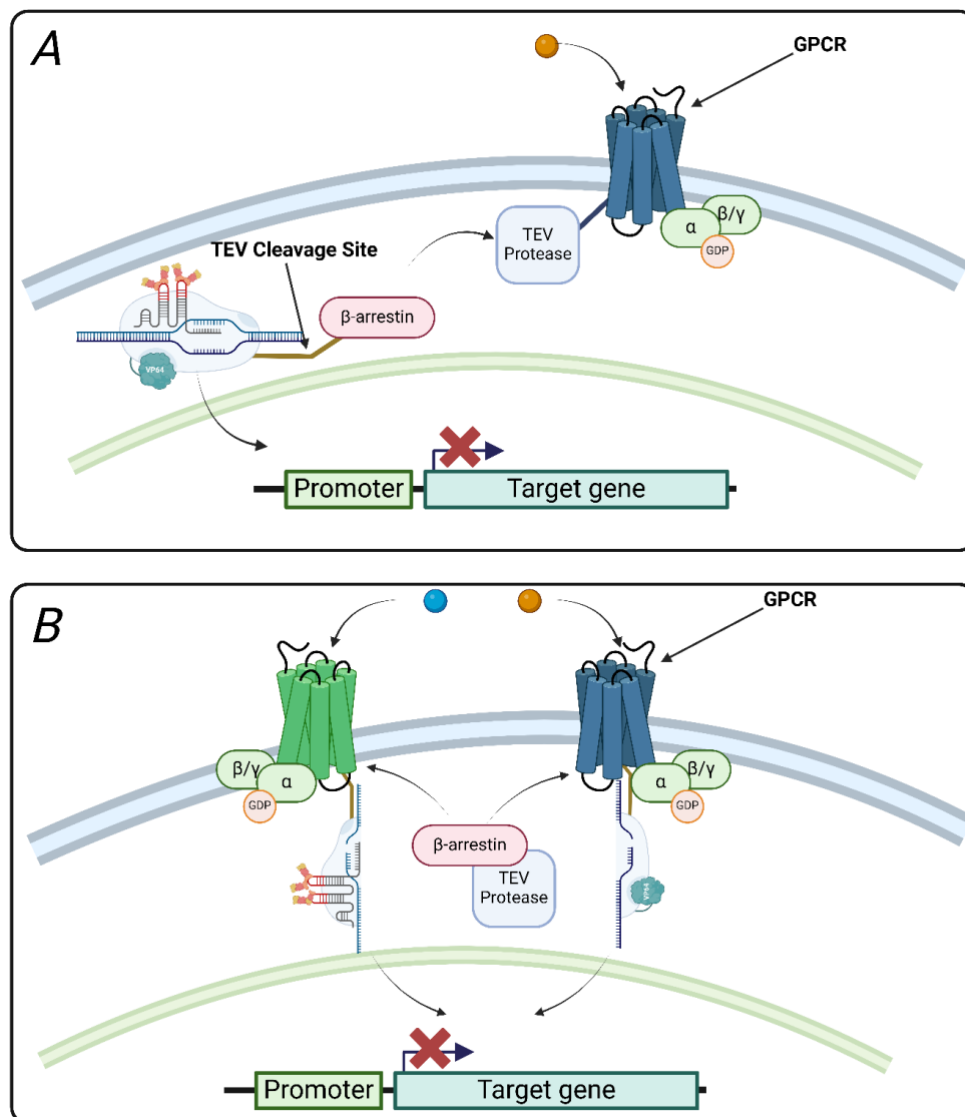


Figure 1.5 Diagram of GPCR-inducible CRISPR systems. **A)** Cha Cha system by Kipniss et al. (2017). CRISPRa-machinery is tethered to β A via a TEVcs, while the TEV protease is fused to the GPCR. In this state, CRISPRa is unable to interact with the genome and induce the expression of the gene of interest. Ligand binding interaction between the GPCR and the intended ligand triggers the recruitment of β A-CRISPRa. Proximity of the TEV to the TEVcs releases the

6A from CRISPRa machinery which is able to induce gene expression. B) GPCR-CRISPRa system by Baeumler et al. (2017). The CRISPRa machinery is split between the N and C-terminus. Each end is tethered to a different Tango GPCR. Activation of both GPCRs (AND gate) releases both halves of the CRISPRa machinery. Once reconstituted the CRISPRa-machinery can induce the expression of the gene of interest.

CRISPR as a DNA Regulator

CRISPR Discovery and first applications

The Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) system is a natural immune defence mechanism evolved in bacteria and archaea to counter bacteriophages and invading plasmids. This system relies on a CRISPR array of genome targeting sequences, known as spacers acting as recognition sites for invading targets, interspersed with identical repeats. These work together with a CRISPR-associated (Cas) effector to modify the genomic target sequence. Bacteria surviving a viral or plasmid infection integrate short fragments of the invading sequence (protospacers) into the CRISPR array. On repeat infection, transcription and processing of the spacer sequence into CRISPR-RNA (crRNA) when combined with a trans-activating CRISPR RNA (tracrRNA) and the Cas effector, allows the complementary binding of the crRNA to the invading DNA. Target recognition triggers the cleavage activity of Cas effectors such as Cas9, which distinguishes between the crRNA sequence and the target sequence through the recognition of a protospacer adjacent motif sequence (PAM). The PAM sequence, usually characterized by two guanines and one other base (NGG), enables the Cas effector to target the invading sequence. There the Cas effector induces an enzymatic double-strand break (DSB) in the invading DNA, rendering the sequence inactive (Deltcheva et al., 2011, Gostimskaya, 2022)(**Fig. 1.6**).

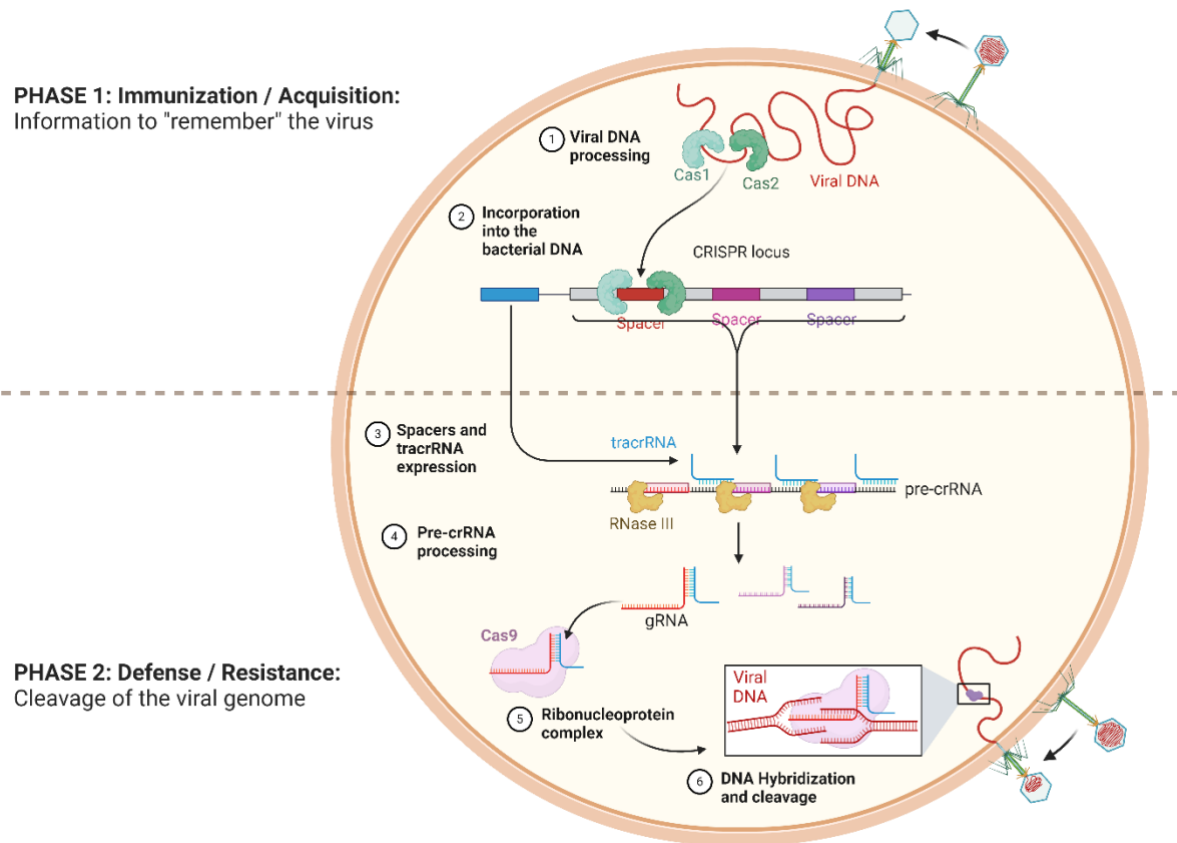


Figure 1.6 Diagram of the natural CRISPR defence mechanism. Phase 1: bacteriophages insert their genome into the bacteria. Surviving bacteria integrate a portion of the viral DNA into their own genome. Different viral DNA sequences are integrated, interspersed by spacer sequences into the CRISPR locus. Phase 2: upon infection by the same bacteriophage, RNase III processes the integrated viral sequence in the form of a crRNA, which is combined with a tracrRNA to form a guide RNA which directs the Cas enzyme to the matching inoculated sequence. The Cas distinguishes the invading sequence from the integrated one by the presence of two guanines and another nucleotide (NGG) adjacent to the targeted sequence and cleaves the two strands of the invading DNA, terminating the infection. Image from BioRender (Le Rhun et al., 2019).

The seminal study by Jinek et al. (2012) demonstrated that only the Cas9 enzyme and the guide RNA (gRNA) were required to achieve double-strand breaks in the sequence of interest *in vitro*, making it a viable gene engineering method. The study also showed that tracrRNA and crRNA could be combined into an individual RNA sequence, later known as single-guide RNA (sgRNA). Subsequently, various teams optimized aspects of the CRISPR system to enhance its effectiveness in eukaryotic cells, focusing on codon optimization of the Cas9 sequence, incorporation of a larger part of the tracrRNA into the sgRNA for improved efficiency, and addition of nuclear localization signals to the Cas9 gene (Cong et al., 2013, Cho et al., 2013, Hwang et al., 2013, Jinek et al., 2013, Mali et al., 2013). Mali et al. (2013) and Cong et al. (2013) were the first to successfully demonstrate that the CRISPR system could be reprogrammed to modify endogenous genes in the genome. The mechanism of modifying the genome takes advantage of the endogenous DNA repair pathways in eukaryotic cells that respond to the presence of an induced double-strand break: non-homologous end-joining (NHEJ) or homology-directed

recombination (HDR). NHEJ occurs throughout the cell cycle, repairing double-strand breaks by ligating the two ends without a homologous template. Following the CRISPR/Cas9-induced DSB, the two ends are held together by the Ku70-Ku80 heterodimer. The two ends are repaired, usually without unexpected mutations. However, repetitive cleavage of the same site by the Cas9 endonuclease can eventually lead to the introduction of mutations like indels or substitutions (Stinson and Loparo, 2021). In contrast, HDR usually takes place during DNA replication, using the sister chromatid as a template. In this process, the two ends of the DSB are resected to create 3' single-strands which invade the undamaged DNA of the sister chromatid, creating a D-loop. Within the D-loop, a high fidelity polymerase uses the undamaged DNA template and synthesizes a new sequences, which is then re-inserted into the damaged DNA (Porteus, 2016) (**Fig. 1.7**). Mali et al. (2013) and Cong et al. (2013) demonstrated that the presence of an exogenous DNA sequence (donor sequence) with homology to the target during Cas9-induced DSB could serve as a repair template in the HDR pathway, integrating the exogenous sequence into the host's genome.

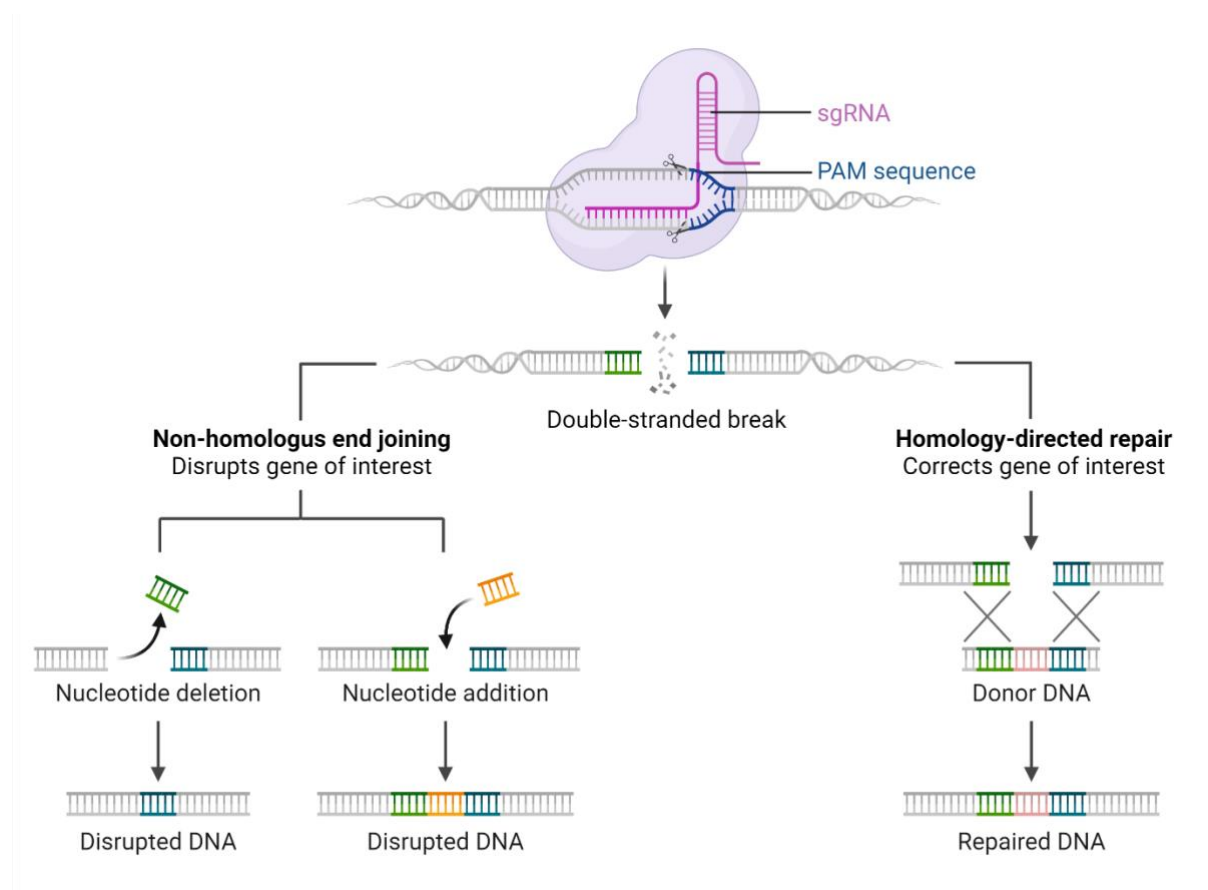


Figure 1.7 Diagram of Non-homologous end joining (NHEJ) and homology directed recombination (HDR) caused by CRISPR /Cas9 cleavage of the DNA. In NHEJ, following the introduction of a double-strand break, the DNA is repaired, sometimes with the addition or removal of random nucleotides to connect the two broken ends. In HDR, following double-strand breaks a donor sequence with homology to the ends of the broken strands is used as a template to repair the broken DNA. Addgene. (2017) Chapter 3: Using CRISPR in Your Experiments. CRISPR 101: A Desktop Resource

CRISPRi and CRISPRa

Since its initial application as a genome engineering tool, CRISPR has undergone significant advancements, expanding its range of tasks. Notably, variations of CRISPR have been developed to function as transcription factors, providing a powerful tool for controlling gene expression.

A study from Qi et al. (2013) demonstrated that mutations in the RuvC1 and HNH nuclease domains of the Cas9 sequence (D10A and H841A) led to the creation of a catalytically inactive molecule with preserved sgRNA-directed DNA recognition behaviour, termed dead-Cas9 (dCas9). They further demonstrated that dCas9 targeting promoter regions was able to silence gene expression without causing any double-strand breaks leading to the concept of CRISPR interference (CRISPRi) (**Fig. 1.8A**). This inhibition occurred as the repressor physically blocked the RNA polymerase during transcription elongation. Importantly, CRISPRi demonstrated reversibility, offering temporal control absent in the original CRISPR/Cas9 editing modifications.

CRISPRi was later optimized for eukaryotic editing by Gilbert et al. (2013) through the fusion of dCas9 to the Kruppel associated box (KRAB) domain of the Kox1 gene. KRAB is a binding dependent transcriptional repressor through its ability to recruit H3K9me3 and heterochromatin protein 1 β (HP1 β) between itself and the promoter, halting transcription (Groner et al., 2010). The dCas9-KRAB was shown to be effective in inhibiting the transcription of target genes in eukaryotic cells (**Fig. 1.8B**).

Expanding on CRISPRi, Gilbert et al. (2013) developed the first version of CRISPR activation (CRISPRa) by fusing dCas9 to four copies of the VP16 transcription factor (VP64) and the NF-kappa p65 activation domain. This system upregulated transcription when targeted to the gene promoter, recruiting RNA polymerase effectively (**Fig. 1.8C**). While CRISPRa was successful in activating the gene of interest, its upregulation was modest at best (Kampmann, 2018). To enhance upregulation signals, variations such as CRISPRa-VPR, Cas9-SunTag, and Synergistic Activation Mediator (SAM) were introduced.

CRISPRa-VPR was developed by fusing VP64, the p65 subunit and the Epstein-Barr virus R transactivator (Rta) to the dCas9 (Chavez et al., 2015) (**Fig. 1.8D**). CRISPRa-VPR was shown to induce significantly stronger upregulation than the VP64-only. Interestingly, they also highlighted two CRISPRa-related issues that remain relevant: gene-dependent variation in gene upregulation and an inverse correlation between the basal gene expression level and the relative expression that can be induced by CRISPRa-VPR. A similar iteration, named Cas9-SunTag, was developed by Tanenbaum et al. (2014). The SunTag consists of a single polypeptide chain derived from single chain variable fragment antibodies (scFv) as a tool to recruit further activation factors.

Specifically, the Cas9-SunTag is composed of the canonical dCas9 fused to a peptide chain composed of 10 x GCN4 epitopes. Simultaneously a VP64 molecule is fused to a scFV targeting the GCN4 epitope. With this system, once the dCas9 has bound to the promoter of interest, it recruits a maximum of 10 x VP64-GFP-GCN4_{scFV} which in turn triggers strong transcriptional upregulation (Tanenbaum et al., 2014) (**Fig. 1.8E**). While the Cas9-SunTag proved relatively effective, it also demonstrated a few significant drawbacks. The most notable of these was the size of the construct, especially the tag itself, which could affect transfection, expression efficiency, and protein half-life. Finally, Konermann et al. (2015) developed a RNA-scaffold version of CRISPRa, known as Synergistic Activation Mediator (SAM). The CRISPRa-SAM strategy still employed the canonical fusion of dCas9 and VP64, however the two hairpins of the sgRNA were engineered to include MS2 aptamer sequences to which the coat protein of the MS2 bacteriophage binds (Peabody, 1993). The system thus includes the MS2 coat proteins fused to transcriptional activators, namely the p65 and the activation domain of the human heat-shock factor (HSF1). CRISPRa-SAM demonstrated some of the most robust and specific upregulation signals while maintaining a smaller construct size (Konermann et al., 2015) (**Fig. 1.8F**). In comparison to the original CRISPRa, the VPR, and SunTag version, SAM was demonstrated to induce a stronger fold increase in gene expression using only a single sgRNA (Chavez et al., 2016). Recently, minor modifications to the CRISPRa-SAM system have demonstrated further enhancements to its efficiency. Specifically, the removal of the polyU tract in the MS2 targeted tetraloop, the addition of a GC-rich stem extension and the removal of the bulge sequence adjacent to the MS2 aptamer were all shown to improve the system's upregulation by stabilizing the RNA structure (Heidersbach et al., 2023).

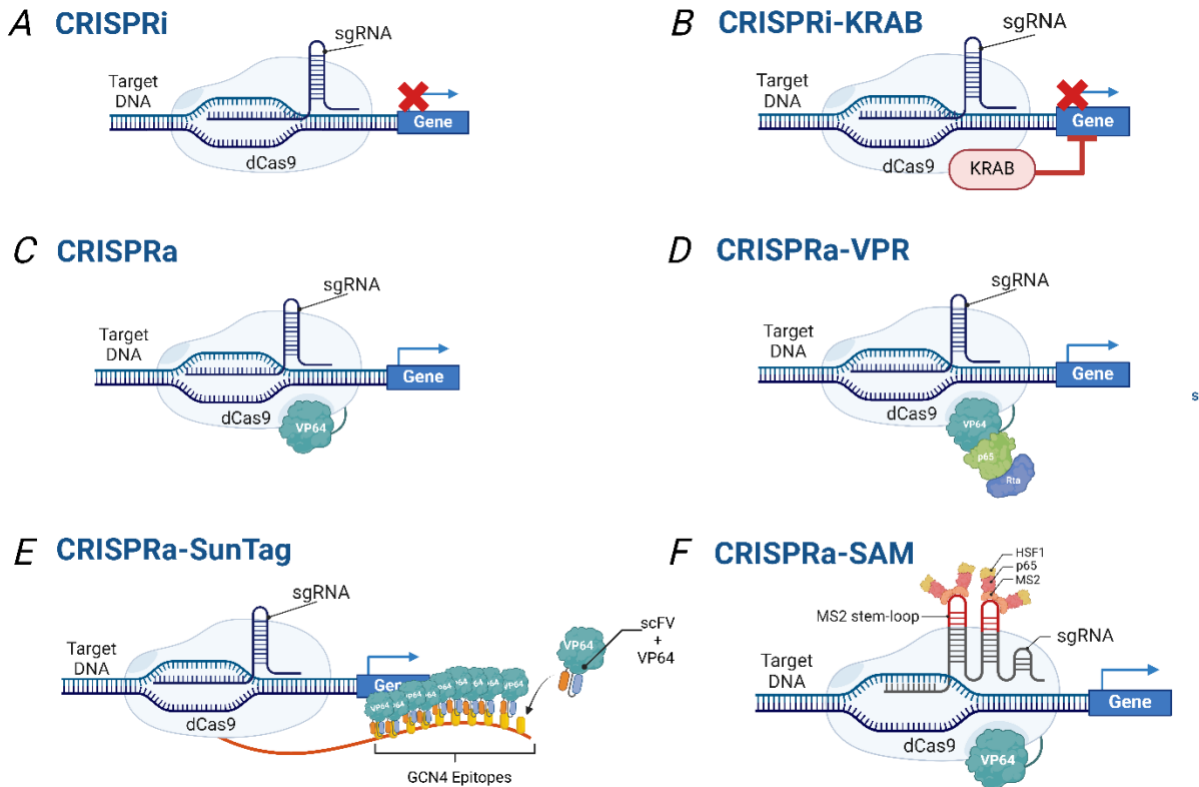


Figure 1.8 Diagram of CRISPRi and CRISPRa systems. **A)** CRISPRi, composed of a dCas9 obstructing the binding of the polymerase. **B)** CRISPRi-KRAB, addition of the KRAB repressor induces transcriptional repression of the site. **C)** CRISPRa, combination of dCas9 and VP64 allows targeting of the promoter and recruitment of the polymerase machinery. **D)** CRISPRa-VPR, in addition of VP64, the p65 and Rta domains are added to strengthen the recruitment of the polymerase. **E)** CRISPRa-SunTag, several GCN4 epitopes are connected to the dCas9 to allow the binding of VP64 fused to scFV targeting GCN4 allowing for the binding of multiple VP64+scFVs. **F)** CRISPRa-SAM, containing a modified sgRNA with two MS2 loop for the binding of the MS2 proteins fused to p65 and HSF1 transcription factors.

Further modifications have been made to both CRISPRi and CRISPRa. Li et al. (2020) developed enCRISPRa and enCRISPRi, tools specifically designed for epigenetic modulation. enCRISPRa was composed of CRISPRa-SAM system where the VP64 fused to the dCas9 was replaced by the core domain of the histone acetyltransferase p300 (Hilton et al., 2015), and the VP64 was fused directly to the MS2 proteins recruited to the sgRNA hairpins. With this strategy, the dCas9-p300 is able to catalyse the acetylation of Lys27 on Histone H3 (H3K27a) within the target promoter and the VP64 is able to trigger transcription. enCRISPRi was composed of the original KRAB system combined with the MS2 protein fused to the lysine-specific demethylase (LSD1) (Kearns et al., 2015). These tools allowed for the catalysis of H3K27a and the removal of enhancer-associated H3-Lys4 mono and di-methylation (H3K4me1/2) modifications, respectively (Li et al., 2020).

The ability to regulate target gene expression using CRISPRa and CRISPRi has proven valuable in studying various disorders (Kanafi and Tavallaei, 2022). For example, Wang et al. (2018) successfully employed a number of CRISPRi systems to inhibit the expression of *granulin* (*GRN*),

a growth factor shown to promote cancer progression by inducing self-renewal of hepatic stem cells. The study used a combination of CRISPR fused to DNA methyltransferase (DNMT3a), histone 3 lysine 27 methyltransferase (EZH2) and KRAB to suppress the *GRN* promoter, resulting in reduced cancer cell proliferation and invasion.

Similarly, CRISPRa has been demonstrated to be a useful tool for the treatment of a number of disorders (Becirovic, 2022). Muscular dystrophy type 1, is caused by mutations in the laminin- α 2 (*LAMA2*) gene resulting in compromised muscle fibres, among other deleterious phenotypes. Overexpression of *LAMA1* in a pre-symptomatic muscular dystrophy mouse model through classic CRISPRa targeting demonstrated the prevention of fibrosis and paralysis (Kemaladewi et al., 2019). Similarly, retinal degeneration was found to be attenuated in mouse models inoculated with adeno-associated vectors expressing CRISPRa-VPR, which upregulated the cone photoreceptor-specific M-opsin (*Opn1mw*) gene, rescuing the degeneration of the cones (Böhm et al., 2020). As shown, the various pre-clinical treatments developed based on CRISPRi and CRISPRa demonstrated their adaptability to modulate gene expression, targeting both gain and loss of function without fully modifying the genome which has made them suitable therapeutic tools (Bendixen et al., 2023).

Inducible CRISPR activation

Exploration of strategies for regulating Cas9 expression offers an added layer of control for genetic engineering and synthetic biology.

Efforts have been made to regulate Cas9 through inducible Polymerase II (Pol II) TetO expression. González et al. (2014) integrated a TRE-regulated Cas9 into the *AAVS1* safe harbour locus of iPSCs, which was inhibited by a rtTA. The addition of doxycycline in combination with the sgRNA was able to control gene editing with low background noise in the absence of dox. A similar approach was also used to induce Cas9 expression for the development of transgenic mice (Dow et al., 2015). A Tet-On system also relied on doxycycline induction of Cas9 expression to increase the specificity of multiple gene targets (Cao et al., 2016, Wang et al., 2017). A more updated version was developed by Lundin et al. (2020) where a Cas9 cassette regulated by a tetracycline controlled element promoter was integrated into the adeno-associated virus integration site1 (*AAVS1*) human safe harbour locus and the Gt(*ROSA26*)Sor (*Rosa26*) safe harbour locus in mice. The system recorded no Cas9 leakiness or activity outside of the dox-based induction. Recently, Park et al. (2022) developed an AND gate Cas9 expression system relying on both Tet-On and a protein stabilising domain. However, TetO-regulated Cas9 systems have recorded suboptimal

temporal control as they required several days to be activated and deactivated, making these strategies less effective when time is an important variable (Nguyen et al., 2016).

Achieving inducible sgRNA expression could provide a faster and more controlled method to regulate CRISPR systems. However, while fine-tuning inducible gene expression for Pol II transcription is well-established, achieving inducible sgRNA expression through regulated Polymerase III (Pol III) promoters, crucial for sgRNA transcription, has presented challenges (Kor et al., 2023). In fact, Pol III promoters are characterised by a number of differences from Pol II promoters, such as a defined transcription start position, poly-T sequences as termination signals and the expressed small RNAs do not go through further modifications such as capping or polyadenylation (Gao et al., 2018). These properties make the Pol III promoter system ideal for the expression of small functional RNA molecules, such as shRNA and sgRNAs.

One strategy to inhibit gene expression involved combining Pol III promoters and TetO. Originally H1 and U6 Pol III promoters combined with TetO were demonstrated to regulate shRNA expression through the addition of doxycycline (Henriksen et al., 2007, Herold et al., 2008). This strategy was successfully extended to sgRNA expression, which was found to be effectively inhibited by TetR and activated by the addition of doxycycline for *in vivo* mouse editing (Aubrey et al., 2015). Similarly, both de Solis et al. (2016) and Kumar et al. (2018) used both U6 TetO and H1 TetO to regulate sgRNA expression for *in vivo* gene editing while Sun et al. (2019) successfully used a U6 promoter containing two TetO sequences to decrease background expression and increase control of sgRNA expression. Similarly, Kelkar et al. (2020) employed a H1 TetO promoter to regulate sgRNA expression for self-inactivation of CRISPR. Interestingly, the U6₂xTetO promoter was shown to be characterised by extensive leakiness compared to other inducible systems, a problem common to TetO regulators (Chylinski et al., 2019).

As the combination of Pol III promoters and TetO was shown not to yield the same efficiency as the more optimized Pol II – TetO combination, several studies investigated the possibility of expressing the sgRNA through a Pol II promoter. Here, an additional cleavage of the Pol II-generated mRNA is necessary to free the sgRNA molecule and ensure that the sequence is not affected by the canonical mRNA modifications, such as the addition of a 5' cap or a polyadenylation tail (McCarty et al., 2020).

One possible strategy is the use of the endogenous tRNA-processing mechanism (Xie et al., 2015). The first systems were described in plants and utilised polycistronic sgRNAs interspaced by precursor tRNA (pre-tRNA) sequences. These proved effective for the generation of multiplexed sequences driven by Pol III promoters. In this design, the pre-tRNAs were effectively

processed by the naturally expressed ribonucleases P and Z, allowing for the release of the interspersed sgRNAs post-transcriptionally (Xie et al., 2015, Qi et al., 2016). Shortly after, the same multiplexing strategy was demonstrated to be effective in mammalian cells through both Pol III and Pol II promoters (Xu et al., 2016, Dong et al., 2017). Xu et al. (2016) demonstrated the possibility of generating a multiplex tRNA:sgRNA system driven by both U6 and CMV promoters as well as an endogenous (muscle-specific) *MHCK7* promoter. Similar results were also corroborated by Port and Bullock (2016) in *Drosophila*, where they demonstrated a conditional expression of Cas9 through the addition of a Gal4-based inducible system. While these multiplexed systems worked efficiently, it is worth noting that the tRNA sequence contains fully functional and naturally integrated Pol III promoters (Knapp et al., 2019). The intrinsic ability of tRNA sequences to drive gene expression can be detrimental in the context of gene circuits where regulation of gene expression is paramount. In fact, tRNAs themselves have been successfully employed to drive the expression of sgRNA albeit with mild efficiency (Mefferd et al., 2015, Wei et al., 2017). For this reason, Knapp et al. (2019) introduced systematic mutations to the tRNA sequence in order to identify the minimally required structure that still retained the RNase P and Z recruitment ability without the internal Pol III promoter activity. With this strategy, the tRNA promoter was de-coupled from the tRNA cleaving scaffold ensuring that the tRNA:sgRNA could be regulated by a CMV promoter only. The same study successfully employed the de-coupled tRNA:sgRNA scaffold to regulate a CRISPRa-SAM machinery, with results which were only second to a Csy4 system, which showed the most effective sgRNA induction.

The Csy4 endonuclease, derived from *P. aeruginosa*, was one of the first tools employed for the excision of sgRNA from primary transcripts driven by Pol II promoters (Nissim et al., 2014). Identified for its role in recognizing and cleaving a 28nt repetitive RNA sequence in processing CRISPR transcripts (Haurwitz et al., 2010), the Csy4 system became an integral component of the CRISPR toolkit, as a CRISPR-based RNA screening platform (Qi et al., 2012).

Nissim et al. (2014) pioneered the development of a Pol II inducible sgRNA expression system using the Csy4 endoribonuclease. In their study, they drove the expression of a sgRNA sequence flanked by two 28nt Csy4 recognition sites and controlled by a CMV promoter. Upon transcription, the Csy4 enzyme successfully released the sgRNA by cleaving at the intended sites, showcasing the feasibility of this strategy for achieving sgRNA expression from Pol II promoters. The study also highlighted that while using a non-native system like Csy4 allowed for tighter regulation, concerns about toxicity arose, as high concentrations of Csy4 negatively impacted cell survival.

Since its inception, the Csy4:sgRNA system has been successfully implemented for sgRNA multiplexing platforms (Kurata et al., 2018, Ferreira et al., 2018). Moreover, it has been a central component of a more complex platform that utilized both Csy4 and TetO to regulate the co-expression of up to 24 sgRNAs for CRISPRa/i screening (Shaw et al., 2022). The combination of the cleavage precision of Csy4 and the inducibility of the Tet operon enhances the versatility and efficiency of the system for various CRISPR applications. Interestingly, Csy4 was also adapted for more complex regulatory systems such as the generation of an inducible hairpin sgRNA (Ferry et al., 2017). In the study, a constitutively expressed sgRNA containing a double-stranded sequence was unable to interact with the Cas9. The Csy4 cleaving site was integrated between the two strands so that Csy4 cleavage allowed for the elimination of the superfluous strand, allowing in the interaction between the sgRNA and Cas9.

TetO sequences, GPCR, CRISPR systems are all very effective tools for the manipulation of mammalian cell behaviour. While their use has often been demonstrated independently, only a few complex circuits have been developed making use of these components together. However, more sophisticated circuits combining the sensitivity of GPCRs and modularity of TetO and CRISPR systems can be effective for multiple applications, especially in the study of complex disorders such as neurodegenerative diseases.

Alzheimer's disease – health challenge of the decade

AD aetiology and modelling

Over the past two decades, advancements in medical and healthcare practices have led to significant improvements in addressing major causes of death, such as stroke and heart disease. Despite these strides, the aging population has witnessed a substantial increase in mortality associated with neurodegenerative diseases. Latest studies reveal that 55 million people are currently affected by dementia worldwide with an expected threefold increase by 2050. In the UK, dementia cases are expected to reach 1.2 million by 2040 (Chen et al., 2023). The profound impact of Alzheimer's disease (AD) on the aging population, coupled with the absence of proven treatments capable of fully eliminating the disease, has spurred intense research efforts to develop more effective therapies.

AD is a synaptic dysfunction disorder encompassing genetic, cellular, and neurological failures; it is characterised by a progressive decline in cognitive function, affecting memory, language, and physiological functions, ultimately leading to death (Knopman et al., 2021). Pathological features of AD include extracellular aggregates of A β peptides (A β plaques), neurofibrillary tangles (NFT) composed of hyperphosphorylated tau, neuronal loss, and gliosis. The exact contribution of plaques or NFTs to neurodegeneration remains unclear, but both are linked to Alzheimer's neuropathology. In addition to these features, AD is characterised by inflammation, disruption of the blood-brain barrier, metabolic disturbances, altered cellular degradation, and endocytic pathways (De Strooper and Karran, 2016). Moreover, neurons are not the sole cell types affected by AD degeneration. Various cell types, including oligodendrocytes, astrocytes, microglia, and the cells of the blood-brain barrier (endothelial cells, pericytes, and astrocytes), have been shown not only to be impacted by the degeneration observed in AD but also potentially play an active role in the disease progression (De Strooper and Karran, 2016). The intricate network of molecular, cellular, and dynamic mechanisms involved in AD complicates the identification of key elements driving the pathology, adding to the complexity of understanding and developing effective treatments for this neurodegenerative disorder.

The pathogenesis of AD involves several genetic factors, and specific gene variants have been identified that increase the risk of AD development. Notably, variants of the A β precursor protein (APP), as well Presenilin 1 or 2 (*PSEN1/2*), Apolipoprotein E (*APOE*), and Clusterin (*CLU*) have been linked to AD (Golde, 2022).

APP is a transmembrane protein expressed in several tissues but mostly concentrated in the brain, where it is implicated in several neuronal processes, such as synaptic formation, synaptic

repair, anterograde neuronal transport, and iron export (Chen et al., 2017). Due to its involvement in these neuronal pathways, APP is rapidly turned over. At the time of replacement, old APP undergoes proteolytic cleavage through two distinct pathways (O'Brien and Wong, 2011). Cleavage by α -secretase results in a large soluble ectodomain (sAPP α) and an 83-residue long C-terminal fragment (**Fig. 1.9**). On the other hand, cleavage by β -secretase produces the soluble β ectodomain (sAPP β) and a 99-residue long C-terminal fragment. In both cases, the 83 and 99-residue fragments are subsequently cleaved by γ -secretase. While the processing of sAPP α results in the non-toxic P3 fragment, the processing of sAPP β results in A β (Haass, 2004, Chen et al., 2017). The cleaving site of γ -secretase can vary slightly, producing A β peptides of different lengths (Tcw and Goate, 2017). Of the various isoforms of A β the most common ones are composed of 40 and 42 residues (A β_{40} and A β_{42}). A β , and specifically A β_{42} , is characterised by hydrophobic residues in its C-terminal, and tends to interact with other A β_{42} peptides, forming oligomers that eventually combine into plaques (Walker, 2020). It is worth noting that γ -secretase is composed of four subunits: *Presenilin 1* and 2 (*PSEN1-2*), nicastrin, the transmembrane protein anterior pharynx defective 1 (APH-1), and presenilin enhancer 2 (*PEN-2*). (O'Brien and Wong, 2011, Zhang et al., 2014). Mutations of the *PSEN1* and *PSEN2* genes have been strongly associated with FAD risks as they have been shown to increase the production of the A β_{40} and A β_{42} variants (Vetrivel et al., 2006).

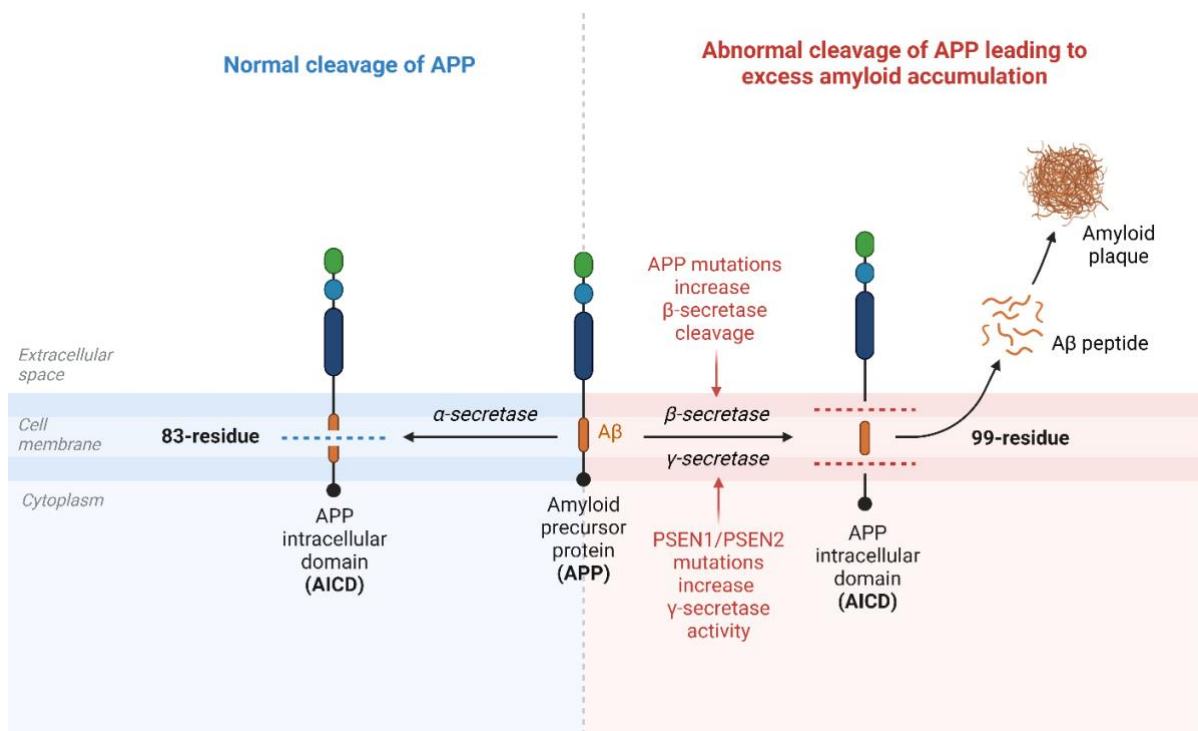


Figure 1.9 Diagram of APP processing. (Left) APP processing by α -secretase generates an APP intracellular domain (AICD) and an 83-long residue that is further processed. (Right) APP processing by β -secretase results in AICD and a 99-residue fragment that is further processed by γ -secretase, resulting in A β fragments (Foster et al., 2019).

The apolipoprotein E (APOE) gene has also been associated with AD (Corder et al., 1993, Saunders et al., 1993). The effect of APOE on AD is isoform dependent, with the APOE2 variant showing increased neuroprotection and the APOE4 variant acting as the main risk factor (McNamara et al., 1998, Reiman et al., 2020). APOE has been shown to play a role in A β deposition (Kloske and Wilcock, 2020). APOE4 is associated with increased A β binding activity but decreased proteolytic ability (de Chaves et al., 2008), possibly hindering A β oligomer clearance and increasing its accumulation rate, potentially involving microglia (Kloske and Wilcock, 2020).

Similarly, *clusterin* (*CLU*) has been identified as the third most significant risk factor for late onset AD development (Lambert et al., 2013). Clusterin is a secretory protein found in a large number of tissues. While neurons have been found to express this gene, astrocytes are the main clusterin producers in the brain (Foster et al., 2019). Increased levels of clusterin have been observed in the brain of AD patients, especially near A β plaques (Miners et al., 2017, May et al., 1990). However, the role that clusterin plays in AD aetiology is still unclear (Foster et al., 2019). A number of studies have suggested that clusterin plays an amyloidogenic inhibiting role, by preferentially binding to A β oligomers preventing further aggregation (Yerbury and Wilson, 2010, Narayan et al., 2012), triggering clearance (Bell et al., 2007, Nuutinen et al., 2007, Merino-Zamorano et al., 2016, Zandl-Lang et al., 2018) or generally inhibiting neuronal degeneration (Cascella et al., 2013). Similarly, knockout of *CLU* in APP/PS1 mice resulted in increased A β but lower inflammation (Wojtas et al., 2017). A more recent study proposed that clusterin secreted by astrocytes aided excitatory synaptic transmission and decreased amyloidogenic stress in 5xFAD mouse models (Chen et al., 2021). On the other hand, other studies have suggested that clusterin could have amyloidogenic properties by promoting plaque formation and preventing the uptake of A β by astrocytes and microglia (DeMattos et al., 2002, Nielsen et al., 2010, Mulder et al., 2014). Interestingly, knockout of both *CLU* and *APOE4* in mice overexpressing human APP (PDAPP) resulted in increased A β production and amyloid deposition, suggesting that the two genes might be working cooperatively to reduce A β generation and deposition (DeMattos et al., 2004). More recently, Liu et al. (2022b) suggested that *CLU* upregulation inhibits A β degradation and promotes A β generation.

Due to the strong correlation between A β and AD development, the amyloid cascade hypothesis (ACH) has become the dominant theory for the pathogenesis of AD. It describes the development of A β plaques as the main pathogenic cause of AD, leading to tau hyperphosphorylation and eventually cell death (Hardy and Higgins, 1992). However, over the past 30 years, several studies have demonstrated that amyloid plaques did not correlate with AD degeneration

(Giannakopoulos et al., 2003, Bennett et al., 2006, Aizenstein et al., 2008). Based on these new data, the focus of the hypothesis shifted from plaques to A β oligomers as the most pathogenic elements, and added the role of penetrance into the hypothesis, ultimately involving APOE and the interaction between tau and A β (Hardy, 2006, Selkoe and Hardy, 2016). However, even this revisited ACH was found to overly emphasize A β as the single most significant driver of AD degeneration, overlooking the complex pathology influenced by numerous genetic, molecular, vascular, and lifestyle factors (Kepp et al., 2023).

The complexity of AD pathology has recently been further emphasised by the limited efficiency of the latest A β antibodies, donanemab, aduanemab, lecanemab, solanezumab and bapinezumab. Aducanumab, a monoclonal antibody targeting A β oligomers, produced contradictory results during its phase 3 trials, resulting in weak evidence of its ability to reduce A β oligomers (Schneider, 2022, Kepp et al., 2023). On the other hand, lecanemab and donanemab showed a modest slowdown of neurodegeneration in patients (van Dyck et al., 2023, Sims et al., 2023)

The limited success of therapies focused on A β in treating AD has shifted attention towards phosphorylated tau, a key feature found not only in AD but also in various neurodegenerative disorders collectively known as tauopathies (Braak et al., 2006). Tau is a member of the microtubule-associated proteins (MAPs) family and plays a crucial role in stabilizing microtubules, contributing to cytoskeleton stability (Pirscoveanu et al., 2017).

In neurons, tau is essential for processes such as axonal transport, neurite growth, and microtubule binding (Johnson and Stoothoff, 2004). Tau phosphorylation is critical for its normal function, as it increases its binding affinity to cytoskeletal components, particularly microtubules. However, inappropriate phosphorylation can occur, leading to the formation of hyperphosphorylated tau.

The actual relationship between A β and tau is still very much unclear. However, a number of studies have demonstrated that tau hyperphosphorylation is promoted by A β oligomers, consistent with the ACH (Hardy, 2003, Johnson and Stoothoff, 2004, De Felice et al., 2008, Bloom, 2014, Pooler et al., 2015). When tau is hyperphosphorylated, it loses its ability to stabilize microtubules and instead tends to aggregate into fibrils known as neurofibrillary tangles. These NFTs are believed to impact cellular homeostasis in several ways. Firstly, they sequester normal tau and other MAPs, leading to microtubule disassembly (Alonso et al., 1996). Secondly, NFTs have been shown to cause blockages in the axonal transport system, resulting in neuronal degeneration (Salehi et al., 2003). Interestingly, the impact of A β and NFTs on neurons differs. A β

leads to fast neuronal "apoptosis," a process completed within 24 hours, while neurons bearing NFTs are characterized by a longer, chronic degeneration (Coleman and Yao, 2003, Raina et al., 2003). This has led to the suggestion that the accumulation of hyperphosphorylated tau could be a symptom of neurodegeneration in response to apoptosis rather than its driving cause (Raina et al., 2003, Zhang et al., 2009).

Given the complexity of the genetic and molecular interplays that characterize AD aetiology, more sophisticated approaches and models are required to explore the pathways and interactions that drive the disorder. This understanding is crucial for developing effective therapeutic strategies that target the multiple facets of AD pathology.

iPSC-derived neurons

The advent of iPSC technology, pioneered by Takahashi and Yamanaka in 2006, has allowed the reprogramming of somatic cells into pluripotent stem cells, providing a powerful tool to model diseases in a human-specific context. Takahashi and Yamanaka (2006) proved that induction of four transcription factors, *OCT3/4*, *SOX2*, *c-MYC*, and *KLF4* (later named Yamanaka factors) in mouse fibroblast was sufficient to result in reprogramming to a pluripotent stem cell state. Shortly after, the same reprogramming was achieved in human fibroblast (Takahashi et al., 2007) and again with a slightly different cocktail of factors - *OCT4*, *SOX2*, *NANOG*, and *LIN28* (Yu et al., 2007). While reprogramming represents a pivotal scientific advancement, one of the most significant strengths of iPS cells lies in their capacity for differentiation into a wide array of human cell types. Over the decade since the development of the first iPSCs, several protocols have been established for their differentiation into numerous cell types (Penney et al., 2020). This versatility confers several fundamental advantages in disease modelling and drug screening. Firstly, the ability to differentiate previously difficult-to-access cell types like neuronal and cardiac cells, facilitating the study of cell-specific disorders such as neurodegeneration using more physiologically relevant models. Secondly, the uniform genetic background shared by differentiated cell types derived from the same iPSC line reduces variability, which is crucial for investigating cell-specific differences, particularly in drug screening studies. Finally, the accessibility and ability to manipulate iPS cells enable effective genetic techniques, further broadening their scope of applications (Ebert et al., 2012, Chang et al., 2019).

The initial protocols for neuronal differentiation from iPSC drew inspiration from methodologies established for isolated neurons (Reynolds and Weiss, 1992). However, the field quickly progressed towards more standardized protocols involving the differentiation of iPSC into

neuronal progenitor cells, followed by subsequent differentiation into specific neuronal types (Chambers et al., 2009, Shi et al., 2012, Nicholas et al., 2013).

A notable development in iPSC-based neuronal differentiation was the protocol proposed by Chambers et al. (2009), which focused on the modulation of the SMAD signalling pathway as a crucial element for differentiation into the neuronal lineage. SMAD proteins are part of the bone morphogenetic proteins (BMP), which are part of a strongly conserved family of the transforming growth factor β receptor (TGF- β) (Blank and Karlsson, 2011). While the specific role of BMP and SMAD during the early stage of neuronal differentiation has not yet been fully elucidated (Jovanovic et al., 2018), it was demonstrated that inhibition of this signalling pathway during early stage differentiation of pluripotent stem cells drove the cells towards the neuronal lineage (Chambers et al., 2009). Originally, the signalling protein, Noggin, was identified as a key inhibitor of several BMPs resulting in neuronal induction (Valenzuela et al., 1995). More recently, the drug SB431542 (SB) was shown to also induce neuronal differentiation by inhibiting the ALK4,5 and 7 proteins, also part of the TGF- β signalling family (Smith et al., 2008). When combined effectively, these compounds blocked the SMAD signalling pathway, promoting iPSC differentiation into neurons (Chambers et al., 2009). This protocol also emphasized the importance of iPSC distribution and introduced cell confluency as a critical step for achieving pure neuronal differentiation. Subsequent refinements to this protocol included the replacement of Noggin with a similar BMP inhibitor known as dorsomorphin as well as the inclusion of vitamin A as its metabolite, retinoic acid, is a central component of cortical neurogenesis (Kim et al., 2010, Shi et al., 2012).

While "chemical" differentiation (through exposure to SB and dorsomorphin) has been a prevalent approach, "genetic" differentiation protocols have also emerged. In these protocols, specific genes are exogenously induced in iPSC to trigger the desired differentiation. For instance, the overexpression of neurogenin-2 (*NGN2*) was demonstrated to induce neuronal differentiation in iPSC over a two-week period (Zhang et al., 2013). This genetic differentiation protocol has been optimized over the years, leading to the establishment of "induced neurons" (iN) (Hulme et al., 2022). Despite providing a faster and standardized method for generating pure neuronal populations, the use of iN neurons as a model might come with some limitations. A comparison between neurons derived from SMAD inhibition and iN neurons revealed that iN neurons exhibited less maturity, suggesting that SMAD-inhibited neurons better represented the desired neuronal model of a more adult state (Rosa et al., 2020). Furthermore, rapid iN differentiation resulted in a combination of peripheral and central nervous system neurons, challenging the initial argument that was used in favour of iN regarding the purity of the neuronal cell populations.

Short exposures to SB and dorsomorphin before *NGN2* overexpression were shown to produce a more homogeneous population of forebrain neurons, emphasizing the higher purity of the SMAD inhibitor differentiation protocol (Chen et al., 2020, Hulme et al., 2022).

iPSC-derived neurons as model for Alzheimer's disease

With the establishment of robust differentiation protocols, several studies have used iPSC-derived neurons as models for AD aetiology and pathology research, offering advantages over traditional mouse models. While mouse models have contributed significantly to AD research, there are limitations, especially in replicating certain aspects of human AD pathology. (Young et al., 2018, Penney et al., 2020). Mouse models such as the APP/PS1, carrying the human APP^{K670_M671} and the PS1 gene sequence with missing exon 9, and the 5xFAD model, carrying the human APP^{I716V}, and ^{V717A} mutations as well as the PS1 gene with the M146L and L286V mutations, have been shown to effectively recapitulate some of the key phenotypes associated with familial Alzheimer's Disease (FAD). New models have been developed, such as the LOAD1 model carrying human APOE4 and TREM2 R47H variants as well as the LOAD2 model carrying APOE4/TREM2 R47H/hAPP Knock-in alleles. However, these models failed to recapitulate late-onset AD equally well (Zhong et al., 2024). Moreover, human-specific factors have been shown to alter the AD phenotype, making mouse models less reliable. For example, Balusu et al. (2023) demonstrated that the long-non coding RNA, maternally expressed gene 3 (*MEG3*), induced necrosis of neurons in response to amyloid plaques, demonstrating the importance of human-based models for the study of AD (Penney et al., 2020).

Neuronal differentiation from AD patient cells yielded several valuable results. Neurons carrying unique APP mutations provided insights into how this gene is associated with neurodegeneration (Penney et al., 2020). iPSC-derived neurons with an APP^{V717I} mutation resulted in increased APP expression as well as enhanced cleavage by both β - and γ -secretase, resulting in an increased production of A β ₄₂ and A β ₃₈ (Muratore et al., 2014). On the other hand, iPSC-derived neurons with an APP^{E693A} mutation, were characterised by endoplasmic reticulum stress as well as oxidative stress (Kondo et al., 2013). Moreover, iPSC-derived neurons with a APP^{A673T} mutation were characterised by lower A β production (Maloney et al., 2014). *PSEN*-related mutations were also successfully investigated using iPSC-neurons derived from patients samples (Penney et al., 2020). Generally, iPSC-derived neurons carrying *PSEN1* mutations were shown to develop defects in several pathways such as mitophagy, autophagy, ubiquitin proteasome, and endolysosomal degradation (Penney et al., 2020). *APOE* was also extensively modelled using iPSC-derived neurons, demonstrating higher levels of A β and tau phosphorylation, as well as

characteristic autophagy and mitophagy defects (Lin et al., 2018, Fang et al., 2019, Wadhwani et al., 2019).

Patient-derived iPSC neurons provide a crucial model for Alzheimer's research, especially when combined with genetic modification, particularly when examining multiple genetic effects simultaneously. The planned integration of *PSEN1* mutations into human iPSCs enables a more comprehensive understanding of the specific role of each variant in AD (Woodruff et al., 2013). This approach has *been* validated by studies like that of Paquet et al. (2016), which replaced healthy *PSEN1* and *APP* genes with mutated versions using CRISPR-directed HDR. Similarly, Meyer et al. (2019) generated iPSC from AD patients and compared them with an iPSC line overexpressing *APOE4* and a healthy control. This comparison showcased that sporadic AD SAD-derived neurons and *APOE4*-overexpressing neurons were characterised by severe gene network dysregulation, with a specific impact on the repressor element-1 silencing transcription factor (*REST*), which is strongly involved in neuronal differentiation and aging. Finally, *CLU*-KO iPSC-derived neurons have demonstrated an increased resistance to $A\beta$ stressors, further corroborating the hypothesis that *CLU* plays an amyloidogenic role in AD (Robbins et al., 2018).

iPSCs also present an appealing option for AD cell therapies. Several studies have explored the use of iPSCs in regenerative medicine for AD without complex gene editing. For instance, transplantation of iPSC-derived neuronal cholinergic progenitor cells into dementia model mice resulted in successful differentiation into cholinergic neurons, restoring cholinergic function and rescuing cognitive deficits (Yue et al., 2015). Neuronal progenitor cells also restored spatial memory in *APP*^{-/-} mice (Fujiwara et al., 2013). Additionally, research by Cha et al. (2016) demonstrated that exposure to iPSC-derived glial cells reduced $A\beta$ plaque deposition in FAD mouse models, aligning with decreased cognitive dysfunction. Takamatsu et al. (2014) illustrated that genetically edited iPSC-derived macrophages, modified to overproduce *neprilysin 2*—an enzyme involved in $A\beta$ degradation—significantly reduced $A\beta$ levels *in vivo*. In a more recent development, Mo et al. (2023) showed that intranasal administration of iPSC-derived cortical neurons led to memory restoration and a decrease in $A\beta$ plaques.

While iPSCs hold promise as models and potential regenerative therapies for various disorders, including AD, they are still in the early stages of development and come with significant limitations (Penney et al., 2020). A primary concern is their genomic instability, a consequence of the reprogramming process that induces pluripotency. This instability manifests as chromosomal aberrations, copy number variations, and single nucleotide variations, during reprogramming and over prolonged culture periods (Rebuzzini et al., 2016, Yoshihara et al., 2017). Moreover, the

tumorigenicity of iPSCs is considered a major hurdle for cell therapies, as their inherent pluripotency and tendency to differentiate can lead to tumorigenic phenotypes, even resulting in teratoma formation (Lee et al., 2013)

CRISPRa/i and iPSC-derived neurons

CRISPRa and CRISPRi, with their ability to induce gene expression changes, are equally valuable for the development of screening platforms. These platforms have significantly contributed to the understanding of cellular mechanisms, particularly in the context of neurodegeneration (Gilbert et al., 2014).

Kampmann (2017) proposed a genetic screening platform focusing on the cellular mechanisms of neurodegeneration. He highlighted the limitations of current methods, which primarily rely on transcriptomic and proteomic analyses to characterize phenotypes distinguishing vulnerable from resilient neurons in neurodegenerative disorders. While these methods identify differentially expressed and translated genes, they often fail to establish causality. Kampmann proposed a more functional approach utilizing induced pluripotent stem cell iPSC-derived neurons as model to study the effects of both neurodegenerative and neuroprotective genes using CRISPRa and CRISPRi. This platform enabled the investigation of various iPSC lines associated with degeneration, providing valuable insights into the underlying genetic and molecular mechanisms.

The Kampmann lab has successfully generated two iPSC cell lines constitutively expressing CRISPRa and CRISPRi by integrating them into the citramalyl-CoA lyase (CLYBL) safe harbor locus (Tian et al., 2019, Tian et al., 2021). These lines could then be differentiated into neurons to investigate the effects of experimental perturbations in gene expression and to serve as a platform for genome-wide screening. Experiments using both CRISPRa and CRISPRi iPSC-derived neurons revealed that the prosaposin lysosomal protein, encoded by the *PSAP* gene, was associated with oxidative stress and ferroptosis, specifically in neurons, demonstrating the effectiveness of this type of platform to reveal novel pathways associated with neurodegeneration (Tian et al., 2021).

A strength of these platforms is the iPSC pluripotency, which allows the same screenings to be carried out on different cell types that share the same original genome.

In a study by Leng et al. (2022), iPSC lines constitutively expressing CRISPRi-KRAB, integrated into the CLYBL locus, were differentiated into astrocytes. These astrocytes were then subjected to a large-scale pooled sgRNA library screening to interrogate their inflammatory reactivity to the

combined effect of IL-1 α , TNF, and C1q. The screening identified two novel inflammatory reactive signatures, enhancing our understanding of astrocyte responses in the context of inflammation.

An analogous system was developed for both CRISPRa and CRISPRi, as described by Dräger et al. (2022). The combined screening of this CRISPRa/i platform on microglia uncovered several neurodegeneration-associated genes, including *PFN1* and *INPP5D*. Additionally, the study demonstrated the effects of pharmacological perturbation on neurodegeneration-related genes in microglia, providing valuable information for potential therapeutic interventions. Similarly, CRISPRa was effectively utilized to increase APP expression in iPSC-derived neurons, showcasing the broader impact of heightened APP expression on the transcriptome (Marei et al., 2023, Ovchinnikov et al., 2018).

A more complex approach was taken by Cooper et al. (2022) that combined iPSC-derived neurons expressing CRISPRi, with massive parallel reporter assays (MPRAs) to functionally investigate several gene variants associated with Alzheimer's diseases. With this screening approach, a GWAS was initially conducted to identify non-coding variants associated with AD or progressive supranuclear palsy. Subsequently, MPRAs were employed in HEK293T cells to identify functional variants that affected gene expression. Finally, the CRISPRi screening pool was used to test the effect of each regulatory variant individually. This comprehensive strategy implicated specific genes, including *C4A*, *PVRL2*, *APOC1*, *PLEKHM1*, and *KANSL1*, in AD and supranuclear palsy. Yin et al. (2023) further expanded this screening method by implementing a genome-wide arrayed CRISPR library. This approach allowed a systematic investigation of the function of individual genes across the entire genome. The combination of GWAS candidate selection, MPRA, and whole-genome CRISPRi screening provides a powerful tool for uncovering the genetic determinants of complex disorders and their associated molecular pathways.

While the combination of CRISPR and iPSC can result in valuable tools for the recapitulation of gene mechanisms involved in the development of neurodegeneration, it is worth noting some of the current limitations of this approach.

As mentioned before, several stem cells have been engineered to overexpress specific genes to speed up differentiation. This process has been achieved for a number of stem cell differentiations, such as myeloid differentiation, through the overexpression of the *myogenic differentiation 1* gene (Shoji et al., 2016), osteogenic differentiation through *HSPA1A* gene overexpression (Zhang et al., 2016b) or overexpression of *SOX2*, *NGN1*, *NGN2* for neuronal differentiation (Zhang et al., 2013, Klajn et al., 2014, Lu et al., 2019). While this strategy has proven effective in accelerating specific lineage differentiation, it has also demonstrated that

constitutive overexpression of genes outside of the intended lineage can derail differentiation and make the cells less relevant as a model.

Moreover, there are criticisms of these models due to the lack of environment-dependent maturation of various neuronal cell types in iPSC models, given that cell types are often differentiated in isolation, potentially impacting the accuracy of the model (Bohlen et al., 2017). To address this, 3D iPSC models, such as organoids, are believed to provide a maturation environment more closely resembling *in vivo* conditions (Sloan et al., 2017).

Finally, as iPSC undergo differentiation, their epigenetic landscape transforms to activate or repress required genomic regions. This can be especially detrimental when integrating transgenes in those regions. Constitutive integration of exogenous sequences into the adeno-associated virus integration site 1 (AAVS1), a common safe harbour locus, has been reported to result in variable expression, possibly associated with specific lineage differentiation (Bhagwan et al., 2020, Ordovás et al., 2015). Transgenes integrated into AAVS1 have been found to be strongly silenced across differentiation while other integration sites, such as *CLYBL* and the human *ROSA26* safe harbour loci, showed decreased expression of CAG-driven CRISPRa as the cells differentiated into cardiomyocytes (Karbassi et al., 2024). Finally, a study by Klatt et al. (2020) demonstrated that AAVS1 undergoes methylation during differentiation into specific lineages, suggesting modifications to the iPSC genome during differentiation that can impact the reliability of exogenous genes and more complex gene circuits.

Overall, the current iPSC-CRISPRi/a platforms have demonstrated the utility of exploring the genetics of neurodegenerative disorders. Optimization of these platforms, by targeting the limitations discussed above can further improve the reliability of these experimental systems and thus help expand our understanding of complex disorders.

Thesis objectives and rationale

A feedback-controlled gene circuit for the study of AD

Gene circuits designed to sense and respond to specific disease biomarkers, such as extracellular A β , through synthetic signalling pathways could serve as effective tools for regulating therapeutic functions. This involves inducing the desired therapeutic behaviour within specific limits of time and intensity of the intended input. This feature becomes particularly crucial for disorders where the homeostatic fragility of the system necessitates controlled exposure to the therapeutic agent (Kitada et al., 2018).

An illustrative example of the utility of such gene circuits is evident in a study by Rössger et al. (2013). In this research, a gene circuit was engineered for the regulated production of pramlintide, a potential long-term treatment for type 1 and type 2 diabetes. The gene circuit was activated through a lipid-sensing receptor (LSR) comprised of the ligand-binding domain from peroxisome proliferator-activated receptor α (PPAR α) and the phloretin-response repressor, TtgR. This unique receptor, incorporating both PPAR α and TtgR repressor elements, enabled the circuit to respond to ligand binding interactions, specifically fatty acids. When the PPAR α was activated by ligand binding, coactivators were recruited, inducing the expression of pramlintide. In contrast, when the PPAR α was not activated, the TtgR repressor effectively blocked the expression of pramlintide. This design allowed the circuit to produce pramlintide only when required, adding a sophisticated layer of control to the delivery of the therapeutic. Similarly, Ausländer et al. (2014), developed a gene circuit for the sensing of histamine, based on the activation of the HRH2 GPCR. In this design, the binding of histamine to HRH2, activates adenylate cyclase which converts ATP to cyclic AMP (cAMP), triggering its downstream signalling pathway. At the end of the cAMP-signalling pathway, the cAMP-responsive binding protein 1 (CREB1) interacts with its promoter pCREB1 to induce the expression of the gene of interest, in this case, SEAP. Finally, as previously mentioned, Baeumler et al. (2017), Kipniss et al. (2017) fused the Tango receptor to a dCas9 for CRISPRa activation, producing GPCR-regulated CRISPR systems.

Design of the gene circuit

Based on these concepts, I aimed to develop a gene circuit activated by the sensing of extracellular A β to induce the expression of neuroprotective genes. In order to do, so three main components were required: a sensing mechanism to generate a signal upon interaction with A β , a regulating mechanism to maintain action potential until the input signal was generated and a gene expression mechanism to generate the output signal (**Fig. 1.10**).

The Tango system served as the sensing mechanism. Specifically, I decided to use chemerin chemokine-like receptor 1 (CMKLR1), as recent studies suggested its interaction with A β (Peng et al., 2015, Zhang et al., 2020, Yu and Ye, 2015). Moreover, the CMKLR1 was one of the first successful GPCRs integrated into the Tango system providing an efficient basis from which to develop the circuit (Kroeze et al., 2015).

CRISPRa-SAM was chosen as the tool for inducing the circuit's output. This output involved activating an endogenous gene of interest, necessitating a system capable of targeting the endogenous genome. The modularity of CRISPRa-SAM allows for the investigation of various putatively neuroprotective genes by simply changing the sgRNA (Konermann et al., 2015, Chavez et al., 2016, Heidersbach et al., 2023).

To control the activation of the CRISPRa-SAM system until the interaction of the CMKLR1-Tango receptor with extracellular A β , I explored ways to regulate the expression of the sgRNA under the control of the transcriptional transactivator (tTA), released from the Tango receptor.

Finally, I intended for the circuit to be constitutively integrated into the genome of human iPSC as previously demonstrated by Tian et al. (2019), Tian et al. (2021). I divided the circuit into two constructs, intended for integration into the *AAVS1* and *CLYBL* safe harbour loci. In addition, I explored the use of neuronal specific promoters to allow the expression in iPSC-derived neurons.

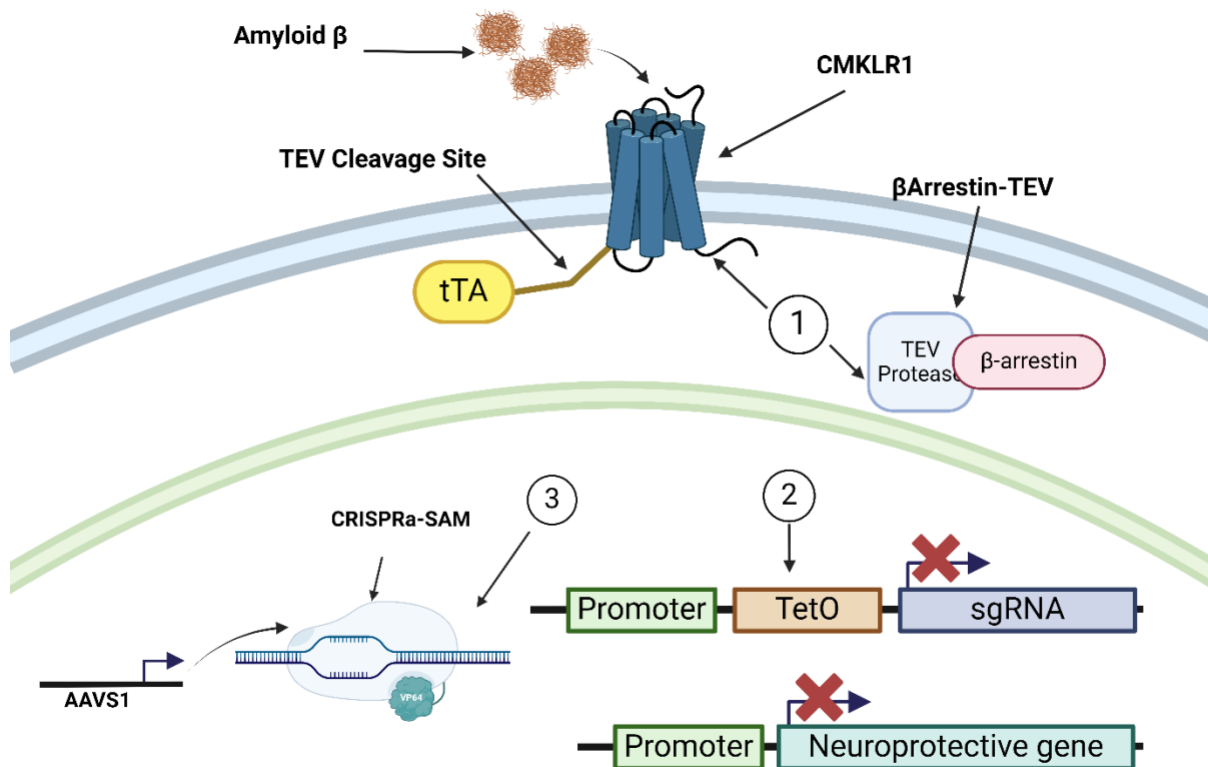


Figure 1.10 Diagram of gene circuit design for the sensing of A β and the expression of neuroprotective gene. The TANGO GPCR is fused to a tTA via a TEV cleavage site. Expression of a sgRNA targeting the endogenous neuroprotective gene is inhibited by the lack of tTA as the TRE promoter driving its expression requires interaction with tTA to induce sgRNA expression. A CRISPRa-SAM machinery is constitutively expressed but inactive until the sgRNA is expressed.

Thesis Outline

The objective of the thesis was to explore genetically engineered iPSC-derived neurons, and its circuit components, as a viable tool for Alzheimer's Disease (AD) research.

Chapter 3 investigates the choice and validation of individual components of the gene circuit. This includes confirming that CMKLR1-Tango effectively functions as a synthetic receptor for sensing extracellular exogenous A β . Simultaneously, three TetO-inducible systems for sgRNA expression are examined.

Chapter 4 details the cloning steps necessary for generating two constructs housing the complete circuit, an AAVS1 targeting vector, BATV2, containing the CRISPR-SAM and β A-TEV and a CLYBL targeting vector, CLICK4, containing CMKLR1-Tango and inducible sgRNA components. The chapter also explores the validation of both plasmids in non-neuronal cells through transient transfections and explores their behaviour and limitations when working together in a single circuit.

Chapter 5 details the integration of BATV2 into iPSC and the differentiation of the CRISPRa-SAM iPSC line into neurons. Throughout differentiation, the use of an endogenous neuronal promoter

to drive the CRISPRa-SAM machinery is investigated, as is the utility of the cell line to study potentially neuroprotective genes through gene upregulation.

Chapter 6 presents a proof-of-concept study that explores the induced expression of the muscle segment homeobox 2 (*MSX2*) gene in CRISPRa-iPSC-derived neurons to investigate its effects on the transcriptome. *MSX2* was chosen as a target gene due to its potential relationship with clusterin (*CLU*), a gene associated with an increased risk of Alzheimer's disease (AD) (Lambert et al., 2013). Notably, a recent study by the Buckley lab demonstrated that iPSC-derived neurons. Additionally, unpublished RNA-sequencing data from these *CLU*-KO neurons revealed that *MSX2* was the most differentially expressed gene, suggesting a possible involvement in the neuroprotective phenotype observed in this cell line. Given *MSX2*'s potential involvement in various neuronal and neurogenesis pathways, along with emerging evidence of its association with *CLU*, a hypothesis arises regarding *MSX2*'s possible neuroprotective role in neuronal cells. Manipulating *MSX2* expression in neurons could provide valuable insights into the pathways regulated by this transcription factor and its potential neuroprotective effects. However, considering *MSX2*'s role in stem cell differentiation, its premature expression could also disrupt neuronal differentiation, underscoring the importance of precise regulation. This makes *MSX2* an ideal candidate for investigating its upregulation in the CRISPRa-iPSC-derived neurons developed in this study.

The results are briefly discussed within each chapter and are further analysed in a concluding discussion, presented as Chapter 7.

Chapter 2 Materials and Methods

Cell Lines

Bacteria

The *Escherichia coli* cell lines shown in (**Table 2.1**) were used for molecular cloning via heat shock transformation.

Table 2.1 *E.coli* cell lines

Cell Line	Genotype
DH5 α	fhuA2 lac(del)U169 phoA glnV44 Φ 80' lacZ(del)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17
ElectroSHOX	F- mcrA Δ (mrr-hsdRMS-mcrBC) Φ 80lacZ Δ M15 Δ lacX74 recA1 endA1 ara Δ 139 Δ (ara, leu)7697 galU galK λ - rpsL (StrR) nupG

Cell lines

Human Embryonic Kidney cells (HEK 293T) were provided by the laboratory of Shoumo Bhattacharya (Wellcome Centre for Human Genetics). The HTLA cell line, a HEK293T line that stably expresses β -Arrestin-TEV (β A-TEV) protease and a luciferase reporter regulated by a tetracycline transactivator (tTA) (Barnea et al., 2008), were provided by the laboratory of Tudor Fulga (Weatherall Institute).

iPSC line

The study utilized the KOLF2-C1 (also known as WTSli018-B-1) human induced stem cell line, derived from a healthy male's skin fibroblasts, and provided by the laboratory of Bill Skarnes at the Wellcome Sanger Institute.

Guide RNA design and Molecular Cloning

In vitro gRNA design

Target spacer sequences for CRISPR and CRISPRa were designed using CRISPR-ERA (<http://crispr-era.stanford.edu/>) and cross-checked on CRISPOR (<http://crispor.tefor.net/>) and NCBI primer-blast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>).

Golden Gate Assembly

The Golden Gate cloning protocol, originally devised by (Engler et al., 2008) was used to insert CRISPR/Cas9 target guides into the appropriate CRISPR or CRISPR SAM sgRNA vectors (**Table 2.2**)

Table 2.2 List of sgRNA for CRISPR and CRISPR-SAM. The list shows the sgRNA sequences utilised in this project. Below the oligonucleotides sequences with the appropriate BbsI overhangs are detailed. Lowercase base indicates the addition of a guanine at the 5' end and respective cytosine at the 3' end. Added to enhance efficiency (Hsu et al., 2013)

CRISPR sgRNA sequences	
gRNA target	Sequences
AAVS1	GGGGCCACTAGGGACAGGAT
CLYBL	TTTCCTCATCCTTCCAACAT
hFLNC	CCCTTGACCACACAGGAAAC
CD2	ACTGTAAAAGATGTAAAGAG
MSX2	CGTCTCCCACTTCCCCTCGG
NT Control	TGAGTTAGCTCTGGTAGTGC
CRISPR sgRNA oligonucleotides	
Oligo	Sequences
CD2-T	CACCgACTGTAAAAGATGTAAAGAG
CD2-B	AAACCTCTTTACATCTTTTACAGTc
HHEX-T	CACCGCCTGGTAGCAAGCGCGTCCC
HHEX-B	AAACGGGACGCGCTTGCTACCAGGC
MSX2-T	CACCgCGTCTCCCACTTCCCCTCGG
MSX2-B	AAACCCGAGGGGAAGTGGGAGACGc
NT-Control-T	AAACTGAGTTAGCTCTGGTAGTGC
NT-Control-B	CACCGCACTACCAGAGCTAACTCA

2µL of 100µM of complementary oligos (T and B; **Table 2.2**) were combined with 16µL Tris-EDTA buffer (Invitrogen). The mix was heated at 100°C for 5 minutes followed by slow cooling to room temperature to allow the annealing of the two oligo strands. 1µL of the annealed oligos were combined with 150ng of the relevant vector plasmid, 1µL of the relevant restriction enzyme, 1µL T4 DNA ligase, 1µL 10x T4 DNA ligase buffer in a total of 20µL reaction. The mix underwent thermocycling according to the following parameters: 10 cycles for 5 minutes at 37°C or the preferred temperature of the restriction enzyme used, followed by 10 minutes at 16°C, 1 cycle of 5 minutes at 50°C followed by 5 minutes at 80°C. Following thermocycling, 1µL ATP-Dependent Plasmid Safe DNase 10U/µL (LGC Biosearch) and 1µL of 1mM ATP (LGC Biosearch) was added to the mix, which was subsequently incubated at 37°C for 1 hour to digest the noncircular DNA.

Ligation-Dependent Cloning

The ligation-dependent cloning protocol was used to generate all the targeting vectors and expression constructs used in this study. The protocol involves the ligation of a minimum of two DNA fragments. 100ng of the insert and vector fragments were prepared by appropriate restriction digestion; the vector was also treated with phosphatase (NEB). The resulting fragments were separated by gel electrophoresis on 0.5-2% agarose, depending upon fragment size, prepared in 1x TAE Buffer (40mM Tris, 20mM Acetic Acid, and 1mM Na₂EDTA). Separated

fragments were then extracted from the agarose using the GeneJet PCR Purification Kit (ThermoFisher) for fragments <10 Kb, or the QIAquick Gel Extraction Kit (QIAGEN) for fragments >10 Kb. One to two insert fragments, depending on the cloning, were ligated with the vector at a 1:3 ratio, using T4 Ligase (NEB) as per manufacturer's instructions.

Bacterial Cloning

5µL of the ligation mix was added to 100µL of ice-thawed chemically competent *E.coli* cells (see *Table 2.1*) and incubated on ice for 30 minutes. Cells were then heat-shocked at 42° for 90 seconds before being incubated on ice for another 5 minutes. 50µL of the heat-shocked cells were spread evenly over an LB agar plate (autoclaved 3.4% LB broth with agar (Sigma) and 96.6% MilliQ supplemented with 100 µg/mL carbenicillin) and incubated at 37°C overnight. Colonies were recovered the following day and grown in an LB medium supplemented with 100µg/mL ampicillin and incubated in a 225 rpm shaking platform at 37°C overnight. The plasmid DNA was purified through GeneJET Plasmid Miniprep Kit (ThermoFisher). The plasmid was verified through restriction digests and Sanger sequencing by Genewiz Inc.

DNA extraction

Cells were lysed with an appropriate volume of lysis buffer (0.1% SDS, 0.4mg/mL Proteinase K, 40µg/mL RNase A (ThermoFisher)), and incubated at 37°C overnight. The cell lysis was mixed with 1:10 volume 8M LiCl and 1 volume 100% isopropanol to precipitate the genomic DNA. The DNA was pelleted at 13,000 rpm for 10 minutes. The DNA was washed twice with 70% Ethanol and centrifuged at 13,000 rpm for 10 minutes. The DNA was dried for 10 minutes at 37°C before being resuspended in DNase-Free Water at a final concentration of approximately 100ng/µL.

PCR and Gel Electrophoresis

1µL of DNA (between 5 and 50ng) was mixed with Immolase PCR Mix (Meridian Bioscience) and the relevant primers, as shown in (*Table 2.3*).

Table 2.3 General Immolase PCR reaction

PCR Reaction		
Reagent	Concentration	Volume (μL)
ImmoBuffer	10x	2
MgCl ₂	50mM	0.6
Primer 1	10μM	1
Primer 2	10μM	1
dNTPs	100mM	0.4
Immolase Polymerase	50x	0.1
Q Solution	5x	4
DNA	5 to 50ng/μL	1
H ₂ O		9.9

The mix was then treated in a thermocycler according to the protocol described in (**Table 2.4**).

Table 2.4 General Thermal Cycling Program

Thermal Cycling Program			
Step	Temperature	Time	Number of cycles
Initial Denaturation	95°C	10:00	1
Denaturation	95°C	0:10	34
Annealing	58° to 64°C	0:30	34
Extension	72°C	0:30 x kb	34
Final Elongation	72°C	10:00	1

PCR samples were mixed with Orange-G loading buffer (25μL glycerol, 5μL Tris HCL (1M), 2μL EDTA (0.5M), Orange G (SLS) (11mM)) and loaded into a 1-2 % agarose gel (Sigma) prepared with 1x TAE (40mM Tris, 20mM Acetic Acid and 1mM Na₂EDTA), submerged in 1x TAE buffer. The electrophoresis was carried out at a constant 130V for a period of between 30 and 60 minutes. Amplification products were imaged under ultraviolet light with a ChemiDoc™ MP Imaging System (BioRad).

DNA Purification and Sequencing

PCR amplification products were treated with Exonuclease I (ExoI) (NEB) and Shrimp Alkaline Phosphatase (SAP) (ThermoFisher) as shown in (**Table 2.5**)

Table 2.5 Exo-SAP PCR purification protocol

Exo-Sap DNA purification		
Reagent	Concentration	Volume (μL)
ExoI	20,000 units/ml	0.075
SAP	1 unit/μL	0.75
SAP Reaction Buffer	10x	0.75
PCR product		15
H ₂ O		0.675

and incubated at 37°C for 1 hour to remove any unincorporated primers and dNTPs. 500ng of purified DNA was combined with 2.5μL of the relevant primer (10μM) and DNase-Free water for a total volume of 10μL. The sample was Sanger sequenced by Genewiz Inc.

RNA Extraction, cDNA synthesis and Quantitative PCR

RNA was extracted through RNAeasy Mini Kit columns (QIAGEN) in an RNase-Free benchtop fume hood. RNA was treated with the RNase-Free DNase I Set (QIAGEN) according to the manufacturer's instructions. 100-400ng of RNA was synthesized into cDNA using the qPCRBIO cDNA synthesis kit (PCR Biosystems) according to manufacturer's instructions. Following incubation, cDNA was diluted 1:3 using RNase-Free water and stored at -20°C. Quantitative PCR was carried out using PowerUp™ SYBR™ Green Master Mix (Applied Biosystems) according to manufacturer instructions. qPCR parameters were quantified using CFX96 Touch Real-Time PCR Detection System (BioRad). Expression fold-change was calculated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001) using *MAPK9* as a reference transcript for any iPSC related qPCR analysis and *TBP* as a reference transcript for any other samples.

DNA and RNA quantification

DNA and RNA samples were quantified by absorbance measurement at 260nm on a Nanodrop spectrophotometer 1000 (ThermoFisher). Sample purity was also determined on the same machine through the A260/A280 and A260/A230 ratios.

Immunostaining

5 x 10⁴ cells were seeded on a 12mm round glass cover slips in a 24 well-plate. The cells were washed once with PBS and fixed in 4% paraformaldehyde in PBS for 15 minutes at room temperature. The cells were then washed three times in PBS and permeabilised with permeabilisation/blocking solution (0.1% TritonX-100 + 5% Fish Gelatin + 0.3M Glycine, in PBS) for 30 minutes at room temperature. The cells were incubated with primary antibodies (**Table 2.6**)

in the same permeabilisation/blocking solution at 4°C overnight. The following day, the cells were washed with PBS, 3 times and stained with fluorescent labelled secondary antibodies diluted in PBS (**Table 2.6**) at room temperature and in the dark for 1 hour. After 3 more washes in PBS, the coverslips were mounted on a clean glass slide with DAPI-containing fluoroshield (Sigma). The coverslip was dried for 1 in the dark at room temperature and stored at 4°C. SP8 LIGHTNING confocal microscope (Leica) was used to take pictures of the cells. Targets with conjugated antibodies containing Alexa Fluor-488 and Alexa Fluor-568 were imaged using Texas Red, Fluorescein and Cy5 (for DAPI), respectively.

Table 2.6 List of Primary and Secondary antibodies

Primary Antibodies				
Target Specificity	Host	Working Concentration	Cat. No.	Source
TBR2	Rabbit	1:600	81493	Cell Signaling Technologies
Nestin	Mouse	1:1800	33475	Cell Signaling Technologies
Tuj1	Mouse	1:200	4466S	Cell Signaling Technologies
BRN1	Rabbit	1:200	12137	Cell Signaling Technologies
Secondary Antibodies				
Specificity	Host	Working Concentration	Cat. No.	Source
Alexa Fluor 488-conjugated anti rabbit IgG	Goat	1:1000	A11034	Invitrogen
Alexa Fluor 568-conjugated anti rabbit IgG	Goat	1:1000	A11036	Invitrogen
Alexa Fluor 488-conjugated anti mouse IgG	Goat	1:1000	A11001	Invitrogen
Alexa Fluor 568-conjugated anti mouse IgG	Goat	1:1000	A11004	Invitrogen

Lentiviral Production

Lentiviral production was carried out by Dr Ping Zhang, in the Knight Lab. 8×10^6 HEK293FT were co-transfected with psPAX2 (Addgene, #12260), pMD2.G (Addgene, #12259), pLenti_sgRNA(MS2)_zeo backbone (Addgene, #61427) integrated with the appropriate sgRNA. 72 hours after transfection, media was transferred into a 50mL falcon tube, centrifuged at 400xg for 5 minutes and filtered through a 0.45µm filter to eliminate any cell debris. The filtered supernatant was centrifuged at 29000 rpm for 2 hours at 4°C to collect the virus. Following centrifugation, the supernatant was aspirated into 0.1% Virkon and the pellet was resuspended in 200µL 1.5% BSA/PBS and aliquoted in 0.2mL PCR tubes before storing in -80°C.

RNA sequencing and statistical analysis

RNA Library preparation and Sequencing

RNA was extracted in triplicates from iPSC-derived neurons, 2 and 10 days after viral transduction, according to the protocol described above. RNA quality was confirmed via RNA ScreenTape® on the Agilent 4200 TapeStation and quantified using a Qubit fluorometer. The library preparation and sequencing were performed under the supervision of James Docker in the Technology Platforms group. Polyadenylated mRNA was isolated via magnetic beads using the NEBNext® PolyA mRNA Magnetic Isolation Module (NEB). cDNA synthesis and library preparation were performed using the NEBNext® Ultra II Directional RNA Library Prep Kit for Illumina® (NEB) according to the manufacturer's protocol, with the addition of custom 8bp unique dual indexes. The library was multiplexed and sequenced using the Illumina NextSeq 500/550 High Output Kit v2.5 (75 Cycles) as single-read.

RNA Sequencing Analysis

BCL files generated from the Illumina sequencer were uploaded onto the BaseSpace platform (Illumina®) and converted into FASTQ files. Subsequently, the Wellcome Trust Centre Bioinformatics team processed the data using an in-house RNA sequencing pipeline. This pipeline involved trimming and aligning reads to the UCSC hg19 human genome using the STAR-aligner v2.5 (Dobin et al., 2013), resulting in binary alignment files (BAM). GENCODE v19 served as the reference transcriptome (Harrow et al., 2012). The FeatureCount from the Subread package v1.22.2 was used to assign genomic features to the mapped reads of the RNA sequencing data, producing a count table where each raw sequencing read was mapped to a sample (columns) and an Ensemble gene code (row) (Liao et al., 2014). Additionally, FeatureCount categorized reads into various categories such as correctly mapped, ambiguous, multimapped, no features, and unmapped.

Differential Gene Expression Analysis

Differential gene expression was carried out together with Dr Benjamin Wright (Computational Genomics Group at CHG) using the Bioconductor DESeq2 package on R (Love et al., 2014). Two data tables were created comparing the unnormalized counts (*MSX2* vs Control at 48 hours and *MSX2* vs Control at 10 days). DESeq2 generates a model of the observed counts using three parameters: a normalization parameter to avoid biases between counts across different libraries, a dispersion parameter to account for the variance between reads, statistical modelling through fitting of a negative binomial model. Following the generation of the model, DESeq2 performs hypothesis testing, comparing the treatments with a Wald statistical test. DESeq2 outputs the

following results: 1) “baseMean”, the average of the normalised count, divided by the size factor, taken over all the samples. 2) “Log₂FoldChange”, the effect of size estimate describing the difference in expression of each gene between the sample (MSX2) and the control (NT). 3) “lfcSE”, the standard error of the Log₂ fold change estimate. 4) “stat”, Wald test result. 5) “pvalue” p-value generated from the Wald test. 6) “padj” p-value adjusted for multiple testing through Benjamini-Hochberg method with a false discovery rate set at a <1% threshold.

Gene Set Enrichment Analysis

Gene set enrichment analysis (GSEA) was carried out by Dr Matthieu Mossiee (Computational Genomics Group at CHG) to assess if the gene sets in the MSX2 and NT datasets showed statistically significant differences between phenotypes. Unlike traditional methods relying on thresholds for analysing overexpressed genes, GSEA ranks the entire list of differentially expressed genes based on Log₂ fold change and adjusted p-values (Subramanian et al., 2005). The ranked list is characterized by the most significantly upregulated genes at the top and the most significantly downregulated at the bottom. The dataset is then compared to databases such as the Gene Ontology (Ashburner et al., 2000), where pathways, molecular functions and general phenotypes are associated to gene sets (**Fig. 2.1**). Next, GSEA generates a table a list of terms which are enriched in the dataset. The table produces: 1) enrichment score (ES) indicating the extent of gene set overrepresentation in the inputted database; achieved by determining by what degree the genes in the gene set are overrepresented at the extremes of the list where genes show the largest degree of statistically significant differential expression. The score is calculated using an adapted version of the non-parametric Komogorov-Sminorv test. 2) A normalised enrichment score (NES) by dividing each ES by the average of the null distribution generated over several gene set permutations. 3) A False discovery rate (FDR) percentage by applying the Benjamini-Hochberg method to each NES value, generating an adjusted p-value (padj). 4) A nominal p-value, calculated by permutations, the probability of the ES score being generated over random permutations.

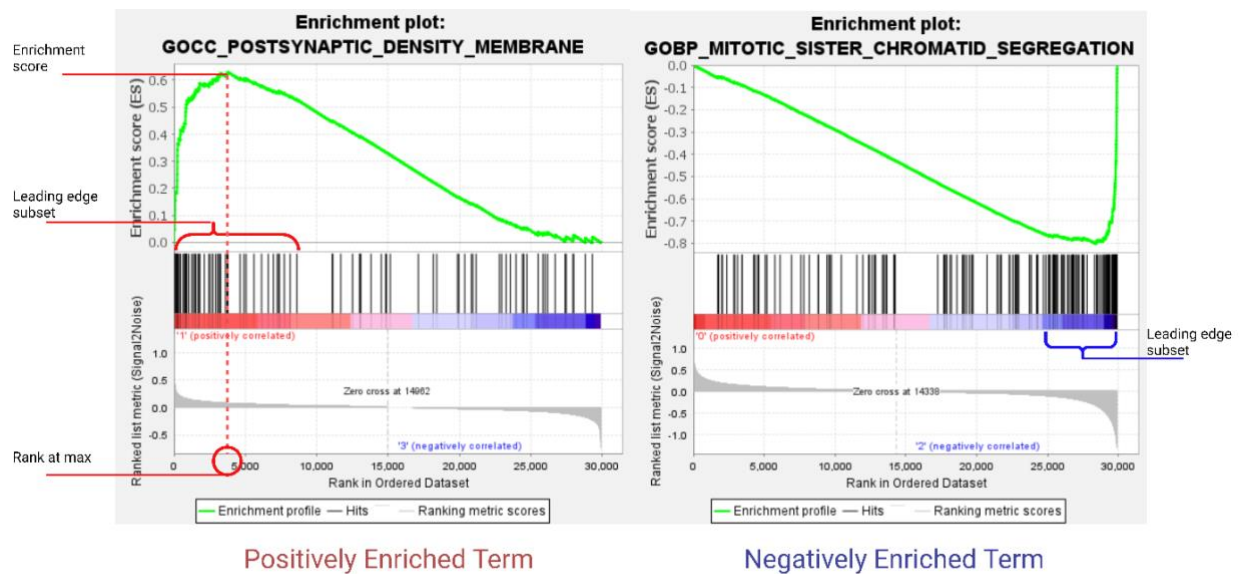


Figure 2.1 Schematic of gene enrichment score analysis. The enrichment score is calculated by comparing each gene in the dataset, from most differentially expressed to least differentially expressed to each gene set associated to a term. A running-sum statistic increases when a gene overlaps in the gene set and decreases when it is not. The ES is equal to the maximum deviation from zero encountered by going down the gene list. A peak (left) indicates a positively enriched (upregulated) term. A valley (right) indicates negatively enriched (downregulated) terms. The black bars in the middle portion of the graph indicate where the overlapping genes appear in the list. The bottom portion of the graph indicates the value that each gene has on the ranking. The further away from 0, the more the gene contributes to the term.

A list of significant terms was extracted from GSEA based on a Familywise Error Rate (FWER) p -value < 0.5 . For each of these terms, genes that were core-enriched were added to a list of genes of interest. Both the list of terms and associated genes of interest were submitted to GeneFeast for further analysis. GeneFeast, a tool developed by Taylor *et al.* (<https://arxiv.org/abs/2309.00061>) was employed to categorise the most common enriched pathways into metacommunities based on the most overrepresented genes shared across the metacommunity. Briefly, the algorithm reduces each term's gene list to a subset of genes of interest. Next, overlapping genes are calculated between each pair of terms through the overlap coefficient or the Jaccard index. A network of terms is built using pair of terms which exceed a defined overlap threshold. In this instance, the default 0.5 was used. Communities are identified within the network through 'greedy modularity maximisation' which is attenuated by an adaptive algorithm that limits the maximum community size. Finally, communities are grouped into meta communities when there is a strong gene-set overlap between communities. Based on this analysis GeneFeast generates a summary of all the metacommunities identified in the dataset represented by a heatmap describing the overlapping genes in the communities, ordered by Log2 fold change and presence in the communities. The software also allows to investigate the individual terms making each community within the metacommunity with the same visualization tools.

Experimental Statistical Analysis

Statistical analyses were performed using GraphPad Prism (version 10.1.0.316). Data from multiple experiments are presented as mean $\pm$ SD. For groups with a sample size of three or more, appropriate statistical tests were conducted. Two-tailed unpaired t-tests were used to compare two unmatched groups when data followed a Gaussian distribution. For comparisons involving a single variable across multiple groups, one-way ANOVA was employed, followed by Tukey's multiple comparisons test. When comparing multiple groups across two variables, two-way ANOVA was used, followed by either Sidak's or Tukey's multiple comparisons test. If data did not follow a normal distribution, log₁₀ transformation was applied to normalize it. Statistical significance was determined at $p < 0.05$. The number of biologically independent experiments is indicated by "n". Further details on the specific statistical tests used are provided in the corresponding experimental sections and figure legends.

Cell Culture and Manipulation

HEK293T and HTLA Cells Maintenance

HEK293T cells were maintained in Dulbecco's Modified Eagle Medium/F12 (DMEM) supplemented with 100mg/mL Penicillin, 100mg/mL Streptomycin, 10% Foetal Bovine Serum (FBS) (Gibco) in the appropriate culture vessels kept at 37°C with 5% CO₂. The medium was changed every 48 or 72 hours depending on confluency. Once the cells reached 80% confluency, they were washed once with PBS before being dissociated with 1 x TryPLE™ (Gibco) according to the manufacturer's protocol and resuspended in 10mL total volume using PBS. Cells were counted using a haemocytometer before being pelleted (100 x g for 5 minutes), resuspended in an appropriate volume of medium and transferred into a new vessel.

HTLA cells were maintained in the same manner as the HEK293T cells. The Dulbecco Medium was further supplemented with 2µg/mL Puromycin and 100µg/mL Hygromycin (Gibco).

iPSC Maintenance

KOLF2-C1 cells were maintained in Essential 8™ (Gibco) Medium without any antibiotics. Unless otherwise stated, the cells were seeded on a culture vessel previously coated with 10µg/mL vitronectin (Gibco) and maintained at 37°C with 5% CO₂. Medium was changed daily. Once the cells reached ~60% confluency, they were dissociated using 1 x TryPLE™ (Gibco) according to manufacturer guidance. The cells were then resuspended in a 10mL total volume using PBS and centrifuged at 100 x g for 5 minutes. The pellet was resuspended in E8 Medium, and the appropriate number of cells seeded in a new culture vessel, using E8 Medium supplemented with

10 μ M Rho-kinase inhibitor Y-27632 (Rock-In) (STEMCELL). Medium was replaced the following day with E8 without Rock-In.

Cell cryopreservation

For cryopreservation, cells were resuspended at a concentration of 1 x 10⁶ cells in 500 μ L medium, DMEM for HEK293T and HTLA cells, E8 for pluripotent iPSC, Neuronal Maintenance medium for iPSC-derived neurons. The cells were combined with equal volume 2 x freezing medium (20% DMSO and FBS (Gibco)). 1 mL total volume was pipetted into a 1-mL cryovial. The cryovials were then stored in a freezing container which was left in a -80°C freezer overnight to achieve a slow rate of freezing. The cryovials were moved in liquid nitrogen for long term storage.

Cell thawing

For thawing, cryovials were partially defrosted at 37°C until a small ice crystal was still present in the middle of the medium. The cells were slowly resuspended in 5 mL culture medium previously warmed to 37°C. The cells were pelleted at 100 x g for 5 minutes. HEK293T and HTLA cells were resuspended in DMEM and transferred to a new vessel. Pluripotent iPSC and Day 20 iPSC neuron precursors were resuspended in their respective medium supplemented with Rock-In (10 μ M), in a pre-coated 10 μ g/mL vitronectin vessel. Medium was replaced the following day with E8 without Rock-In.

iPSC flow cytometry for pluripotency assay.

Stem cells were dissociated with TrypLE™ (Gibco) and 1 x 10⁶ were centrifuged at 100 x g for 5 minutes. The pellet was washed once with PBS and centrifuged at 100 x g for 5 minutes. The cells were fixed in 4% paraformaldehyde in PBS at room temperature for 20 minutes. The staining process was carried out using Stemflow™ Human Pluripotent Stem Cell Transcription Factor Analysis Kit (BD) according to the manufacturer's protocol. The pluripotent markers included in the assay were: OCT4, NANOG, SOX2 and SSEA-4. The cells were analysed through LSRFORTESSA™ X20 CELL ANALYZER (BD Biosciences) and analysed by FlowJo software (version 10.6.2).

iPSC Karyotyping

Approximately 1 x 10⁴ iPSC were plated on to 13 mm round cover glass for 24 hours. The cells were karyotyped by the Chromosome Dynamics Core Facility at the Wellcome Centre for Human Genetics using DAPI staining of metaphase spreads.

Plasmid DNA Lipofection

Approximately 5×10^4 cells were plated in a 24 well-plate. The following day, 2 μ L FuGENE6® (Promega) was mixed with 25 μ L Opti-MEM™ (Gibco) and incubated for 5 minutes at room temperature. Meanwhile, 550ng of total plasmid DNA was combined in a separate tube. The lipofection mix was combined with the DNA and incubated at room temperature for a further 5 minutes. Finally, 200 μ L of medium was combined with the lipofection mix and the total volume of the reaction was added to the cells. Unless explicitly stated, the transfection mix was removed from the cells after 24 hours. Cells were harvested for DNA or RNA, 48 hours after transfection.

RNA Transfection

Synthetic sgRNA was obtained from Synthego using the target sequences described in **Table 2.2**. Approximately 5×10^4 were plated in a well of a 24 well plate. The following day, 3 μ L Lipofectamine™ RNAiMAX (ThermoFisher) was combined with 22 μ L Opti-MEM™ (Gibco) and incubated for 5 minutes at room temperature. Simultaneously, approximately 1 μ g synthetic sgRNA was combined with 24 μ L Opti-MEM and incubated for 5 minutes at room temperature. The RNAiMAX and RNA mix were then combined and incubated for a further 15 minutes at room temperature. Finally, 200 μ L of medium was added to the mix and the transfection mix was added to the cells. The transfection mix was removed 24 hours after transfection. RNA or DNA was harvested 48 hours after transfection.

Luciferase Assay

HTLA cells were seeded into 0.1% gelatine-coated 24 well-plates at 5×10^4 cells per well. On day 1, DNA was transfected in the cells via lipofection (see above). On day 2, cells were exposed to 1mM Chemerin (R&D Systems) in serum-free medium. On day 3, luciferase assay was carried out using Rapid Detection of Firefly Assay Kit (Promega) according to manufacturer guidance. Relative luciferase expression was detected on a ClarioStar microplate reader (BMG Labtech) using the following programme: 10 second measurement read for luciferase activity, 1 second delay between reads, 2 read cycles.

GFP Assay

HEK293T and HTLA cells were seeded into 0.1% gelatine-coated 24 well plate at 5×10^4 cells per well. The cells were transfected the following day. After 12 hours, GFP expression was visualized under the Zoe microscope (Bio-Rad). The cells were dissociated with TrypLE™ (Gibco) and fixed with 4% paraformaldehyde in PBS for 15 minutes at room temperature. After fixing, cells were washed 3

times with PBS and stored at 4°C for up to a week. The percentage of GFP expressing cells was calculated through the Attune NxT Flow Cytometer (ThermoFisher).

RNP Electroporation

A transfection mix was generated by combining 1µg synthetic sgRNA , 0.6 µL of Alt-R™ S.p. HiFi Cas9 Nuclease V3 (20 µM) (IDT) and 4.4 µL Buffer R from the Neon™ Transfection System Kit (Invitrogen). The mix was incubated for 15 minutes at room temperature. After incubation, 1µL of plasmid DNA at a concentration of 500ng/µL was added to the mix. Finally, 10⁵ cells in a volume of 7µL was added to the RNP mix. Cells were suctioned into the P10 tip of the Neon™ and received 2 pulses of 1200 volts for 30ms each. The cells were plated in a 24 well-plate in E8 containing RevitaCell™ (Gibco).

After 24 hours recovery, cells were exposed to 350ng/ml Puromycin for 48hrs. After selection, cells were split, and DNA was harvested for genotyping. In case of successful integration, the cell the cells were low-seeded for colony picking.

Low seeding and colony picking

Following antibiotic selection, 10⁶ were seeded on a 6-cm dish in E8 containing 1X RevitaCell™ (Gibco). After 2 days, RevitaCell was removed from the medium and the cells were left to grow until single colonies were visible under a 4x magnification stereoscope, which took approximately 10 days. Colonies were picked using a sterile p20 Gilson pipette and transferred into a 96wp previously treated with 10µg/mL vitronectin and containing E8. After 7 days, colonies were split into two 96wps. One plate contained 1:3 of the cells, the second plate contained the remaining 2:3. Once the plate containing 2:3 of the cells reached confluency, DNA was harvested for genotyping. Based on the genotyping results the remaining plate was used to expand the colonies carrying the correct integration.

Cortical Neurons Differentiation

Differentiation of KOLF2-C1 iPSC into cortical neurons was based on the protocol described by Shi et al. (2012) with some minor modifications.

5 x 10⁴ cells were plated on a 6-well-plate coated with growth factor reduced, phenol red-free Matrigel (Corning) in E8 medium supplemented with Rock-In (10µM). Once the cells reached 100% confluency (Day 0), the medium was changed to Neuronal maintenance medium (NMM) (Table 1:6) supplemented with SB431542 (10mM) (Tocris) and Dorsomorphin dihydrochloride (1 mM) (Sigma), also known as Neuronal Induction medium (NIM) (**Table 2.7**).

Table 2.7 Reagents of Neuronal Maintenance Medium

Reagents
50% DMEM/F12 + GlutaMAX
Insulin (5µg/ml)
2-mercaptoethanol (100µM)
NEAA (100µM)
Sodium pyruvate (100mM)
1X N2 supplement
1X B27 supplement
1% GlutaMAX
50% Neurobasal

Between day 0 and day 12, NIM was changed daily, and pluripotent stem cells differentiated into neuroepithelial cells, forming a uniform sheet. On day 12, cells were dissociated into single cells with 1x Accutase™ (Stemcell Technologies) and split 1:2 into Matrigel coated 6 well-plates in NMM supplemented with Rock-In (10µM). Between day 13 and day 16, medium was replaced with fresh NMM supplemented with 20 ng/mL of basic Fibroblast Growth Factor (R&D Systems), which induced the formation of neuronal rosettes (**Fig.2.2**). At day 16, cells were dissociated into single cells with Accutase and split 1:3 into Matrigel coated 6-well-plates with NMM supplemented with Rock-In (10µM). Cells were grown in NMM between day 17 and 29, during this period cells were split or cryopreserved anytime they reached 80% confluency. At day 30, cells were dissociated using Accutase and seeded in 24-well plate coated with 0.01% Poly-L-Ornithine (Sigma) and Matrigel in E8 medium supplemented with Rock-In (10uM). Cells were seeded at a concentration of 5×10^4 per well. From day 31 onwards, cells were maintained in NMM medium changed every two days (**Fig.2.2**).

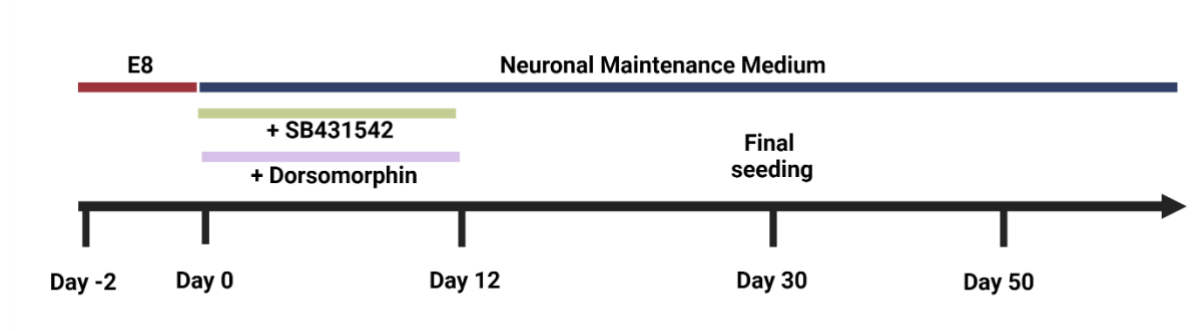


Figure 2.2 Schematic of iPSC derived differentiation of cortical neurons.

Viral Transduction

5×10^5 pluripotent stem cells were plated in a vitronectin-coated 24wp. 5×10^5 iPSC-derived neurons were plated at day 30 at a concentration of 5×10^5 in a 0.01% Poly-L-Ornithine (Sigma)

and Matrigel coated 24-well-plates and grown until the desired maturation. Cells were exposed to E8 or Neuronal medium supplemented with 5 µg/mL Polybrene (VectorBuilder™). The lentivirus containing the desired sgRNA was added to the cells at 1 MOI for pluripotent stem cells and 60 MOI for neurons and the cells were incubated overnight. 24 hours after transduction, the medium was replaced with fresh E8 or Neuronal maintenance medium. RNA was extracted 48 hours after transduction

Chapter 3 Validation of the core circuit components

Introduction

Before designing the complete circuit, I aimed to confirm the efficiency of its core components. Namely, a sensing mechanism to interact with extracellular A β (Tango receptor), a regulatory system to keep the circuit off until activation of the receptor (tTA-activated promoter for sgRNA expression), and a mechanism to upregulate the gene of interest (CRISPRa-SAM) (**Fig. 3.1**).

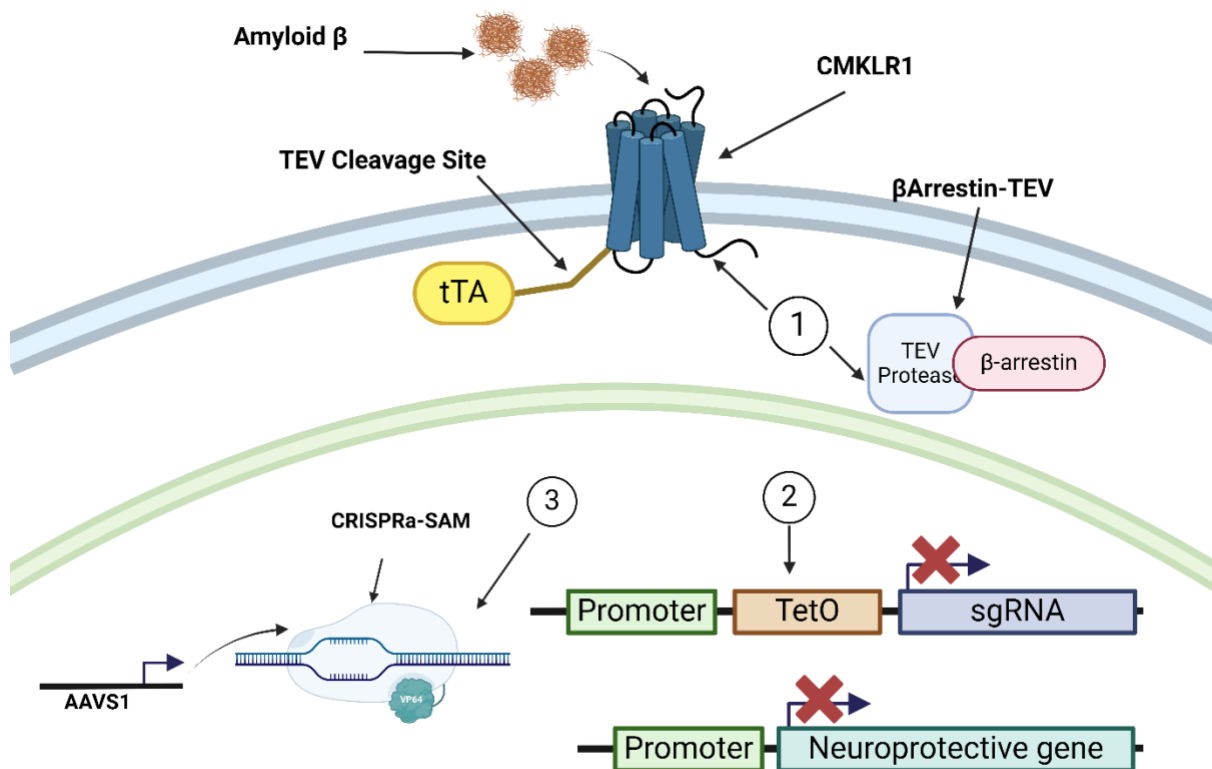


Figure 3.1 Diagram of the core components of the circuit. 1) The Tango system is composed of a GPCR receptor fused to a tetracycline transactivator (tTA) via a TEV cleavage site as well as a β -Arrestin molecule fused to a TEV protease. 2) A tTA-inducible promoter to regulate the expression of MS2 sgRNA required for the activation of the CRISPR-SAM machinery. 3) CRISPRa-SAM machinery for the expression of genes of interest.

For the sensing mechanism, I investigated the use of the chemerin chemokine-like receptor 1 (CMKLR1) in conjunction with the Tango system (Barnea et al., 2008, Kroeze et al., 2015). The CMKLR1 receptor is a member of the GPCR family, which is expressed in the brain, and is known to recognize the adipokine molecule, chemerin (Wittamer et al., 2003, Peng et al., 2015). With respect to AD's pathology, CMKLR1 was found to colocalize with A β_{42} in the hippocampus and frontal cortex of *APP/PS1* mice (Peng et al., 2015). Furthermore, A β_{42} was found bound to CMKLR1 suggesting interaction of the receptor with this molecule. Finally, a recent study demonstrated that CMKLR1 deficiency increased A β_{42} deposits in the brain of the same *APP/PS1* mice, which resulted in a decrease in mortality as well as improved cognitive functions (Zhang et al., 2020).

Further studies from Chen et al. (2022) demonstrated that the CMKLR1 receptor is implicated in the activation and migration of microglia to A β plaques upon CMKLR1/ A β interaction suggesting that the receptor might be involved in neuroinflammation.

For the circuit's output, the CRISPRa-SAM machinery was chosen, specifically for its efficiency in driving gene upregulation (Konermann et al., 2015, Chavez et al., 2016). The modularity of CRISPRa-SAM would allow the investigation of various putatively neuroprotective genes by simply changing the sgRNA. However, before integrating the sgRNAs into the final circuit, validation of the efficiency of the selected guides in inducing the expression of the targeted genes was required.

Finally, the Tango system signals via the release of the tetracycline transactivator (tTA), induced by the cleavage of the TEVcs by TEV protease when the GPCR is bound to its ligand (**Fig. 3.2A**). In order to transduce the signal from the receptor activation to CRISPRa-SAM, regulation of sgRNA expression by the tTA is required. As discussed before, a number of regulatory systems have been invented for the expression of sgRNAs. In this chapter I also investigate a series of tTA-activated Pol III and Pol II promoters for the expression of sgRNA to identify the most efficient one to be integrated into the final circuit.

Validation of the Tango system

To validate the use of the Tango system in the laboratory, I sought to confirm the functionality of the CMKLR1 receptor, initially investigating the ligand binding interaction between CMKLR1 and its known ligand, chemerin (Peng et al., 2015, Zhao et al., 2016). In fact, CMKLR1 was one of the first GPCRs successfully integrated into the Tango system (Barnea et al., 2008, Kroeze et al., 2015). In this study, the CMKLR1-Tango receptor was used to validate its affinity for chemerin, providing a solid reference point for comparing the ligand-binding interaction between A β and CMKLR1 (Barnea et al., 2008). Given recent studies demonstrating the affinity between A β and CMKLR1 (Peng et al., 2015, Chen et al., 2022), along with the extensive optimization of the CMKLR1-Tango system, I deemed this GPCR a suitable candidate for incorporation into the gene circuit.

HTLA cells are engineered from HEK293T cells to constitutively express the β A-TEV and harbour a luciferase reporter gene under the control of the TetO (TetO-luciferase) (Barnea et al., 2008) (**Fig. 3.2A**). These cells were transfected with the CMKLR1-Tango plasmid (Addgene, #66253), which expresses a CMKLR1-tTA fusion receptor. The cells were then exposed to different concentrations of chemerin, 24 hours post-transfection. After 24 hours of incubation, a luciferase reporter assay

was performed to assess the degree of receptor activation which would result in the release of tTA and its binding to the TetO-luciferase reporter gene. A CMKLR1-chemerin exposure curve was generated to determine the optimal chemerin concentration for inducing a robust luciferase signal. As shown in **Fig. 3.2B**, a concentration of 100mM of chemerin induced the strongest luciferase signal $>20 \times 10^6$ relative luciferase units (RLU). However, due to the observation of significant levels of cell cytotoxicity, a lower concentration of 1mM was chosen for subsequent experiments. This concentration still produced a strong signal of approximately 1.5×10^5 RLU

which was significantly higher than the signal from unexposed cells while showing minimal cytotoxicity (**Fig. 3.2B**).

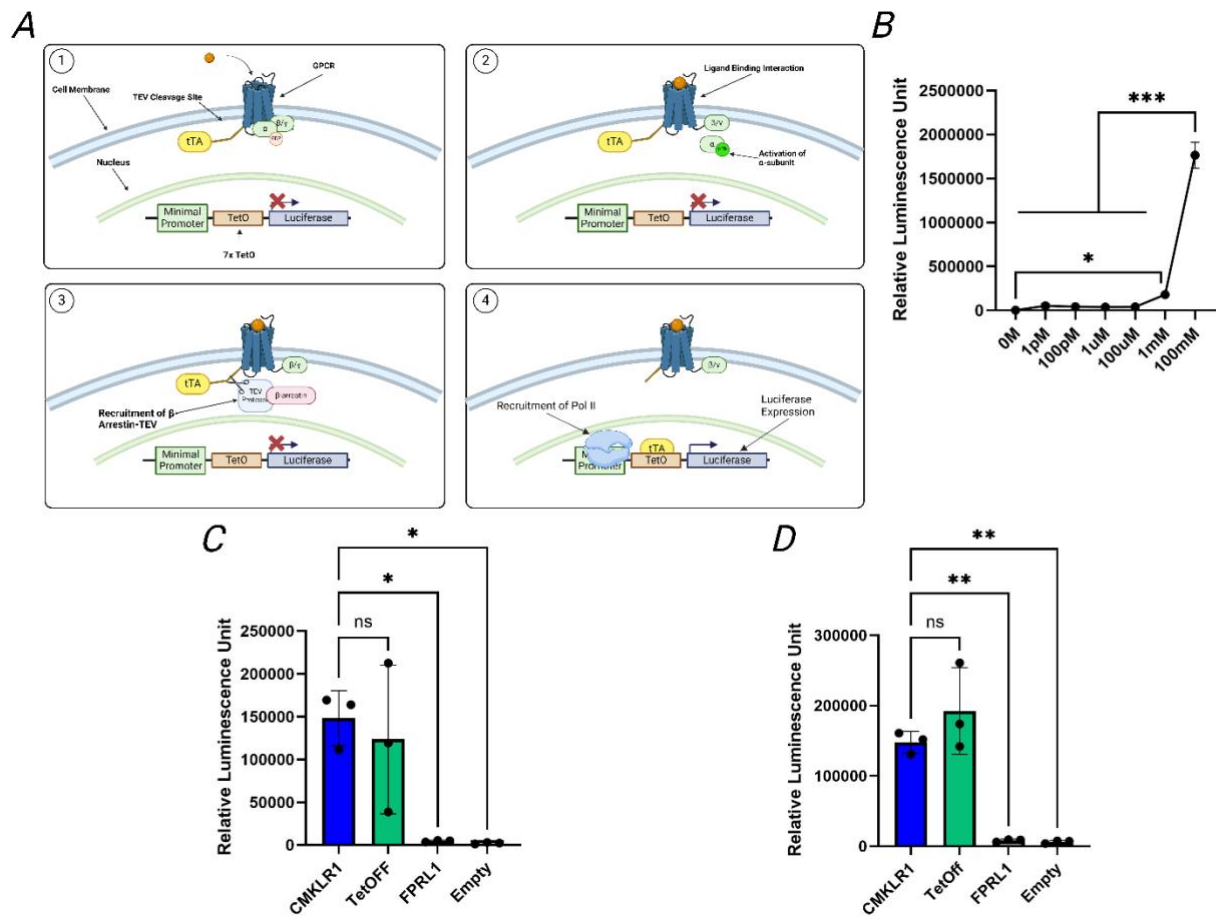


Figure 3.2 CMKLR1-Tango analysis. **A**). Schematic of the Tango system (Kroeze et al., 2015). 1. The GPCR of choice is fused to tTA via a TEV-cleavage site. A luciferase reporter gene driven from a minimal promoter is inactive without the presence of tTA. 2. Interaction of the ligand to the GPCR, activates the receptor by allowing for the release of the Gα-subunit through the phosphorylation of associated GDP. 3. βA-TEV is recruited to terminate the activation of the receptor. Proximity of the TEV to the TEV cleavage site releases the tTA. 4. tTA migrates to the nucleus where it interacts with TetO recruiting the Polymerase machinery to induce the expression of luciferase. **B**) Concentration-response curve for the CMKLR1-Tango receptor from 0 to 100mM chemerin. A one-way ANOVA analysis confirmed a significant difference between concentrations [$F(6, 14) = 383.10, p < 0.0001$]. A Tukey HSD post hoc test confirmed that the 100mM concentration is significantly different from all other concentrations (* $p < 0.05$). The 1mM concentration was shown to induce a significantly higher luminescent value than the 0M control ($p = 0.0263$). Error bars represent standard deviation and sample size = 3. **C**) Relative Luciferase Expression (RLU) of CMKLR1/FPRL1 following 1mM chemerin exposure. A one-way ANOVA demonstrated that chemerin exposure significantly influenced luciferase expression [$F(3,8) = 8.298, p = 0.0077$]. A Tukey's multiple comparisons test demonstrated that CMKLR1-transfected cells produced significantly more relative luminescent units (RLU) than the FPRL1 transfected cells ($p = 0.0221$) and the control transfected with an empty vector ($p = 0.0202$). The RLU of CMKLR1-transfected cells was not significantly different from the positive control – where the Tango system is bypassed via addition of tTA ($p = 0.9120$). Error bars represent standard deviation and sample size = 3. **D**) Relative Luciferase Expression (RLU) of CMKLR1/FPRL1 following 1mM Aβ₂₅ exposure. A one-way ANOVA demonstrated that Aβ₂₅ exposure significantly influenced luciferase expression [$F(3,8) = 27.20, p = 0.00022$]. A Tukey's multiple comparisons test demonstrated that CMKLR1-transfected cells produced significantly more relative luminescent units (RLU) than the FPRL1 transfected cells ($p = 0.0029$) and the control transfected with an empty vector ($p = 0.027$). The RLU of CMKLR1-transfected cells was not significantly different from the positive control – where the Tango system is bypassed via addition of tTA ($p = 0.3801$). Error bars represent standard deviation and sample size = 3.

The experiment was repeated, comparing the CMKLR1-chemerin interaction with another GPCR Tango, FPRL1 (Addgene, #66283), which was not expected to interact with chemerin and thus served as a negative control (He and Ye, 2017). To generate a robust positive control, the Tango system was circumvented by co-transfecting TetOFF, a plasmid encoding a CMV-driven tTA expression cassette (Takara Bio™), inducing luciferase expression by the direct binding of tTA at the TetO within the HTLA cells (from now on referred to simply as tTA) (**Fig. 3.2A**) The results, as shown in **Fig. 3.2C**, indicated that CMKLR1-chemerin (1mM) resulted in an average of 1.5×10^5 RLU, with a similar range seen in the positive control. In contrast, FPRL1-transfected cells, exposed to chemerin (1mM), exhibited no discernible luciferase expression, suggesting that the Tango system is a reliable indicator of specific CMKLR1-ligand interactions.

Next, the interaction between CMKLR1 and A β was explored. HTLA cells were transfected with CMKLR1 or FPRL1 and exposed to 20 μ M of A β_{25} , chosen for its ability to form large β -sheet fibrils and retain the toxicity of full-length A $\beta_{1-40/42}$ peptides (Naldi et al., 2012, Smith and Klimov, 2019). 24 hours after exposure, CMKLR1 transfected cells induced a strong luciferase signal (approximately 1.5×10^5 RLU), which was consistent with the relative luciferase expression resulting from chemerin exposure, confirming the ligand binding interaction between CMKLR1 and A β (**Fig. 3.2D**). Based on this, the Tango system was confirmed to be a reliable mechanism for initiating the activation of the circuit and that CMKLR1 can serve as an A β sensing receptor.

Validation of CRISPRa-SAM sgRNAs

The CRISPRa-SAM relies upon a unique sgRNA sequence scaffold comprising two MS2 aptamers which allow for the recruitment of an HSF-p65-MS2 transcriptional activator complex that, together with the dCas9-VP64 transcriptional activator, triggers robust upregulation of gene expression from bound promoters (Konermann et al., 2015). sgRNAs targeting the proximal promoters of the genes investigated in this project had to be designed and validated.

Furthermore, as the ultimate objective of the project is to integrate the gene circuit into stem cells, it was imperative for the sgRNAs to induce gene upregulation in this cell line. Therefore, the testing of the sgRNAs was conducted using a KOLFSAM iPSC line. KOLFSAM has been engineered by Dr. Phalguni Rath, in the Davies lab, and stably expresses the CAG-driven CRISPRa-SAM machinery (Hitoshi et al., 1991, Niwa et al., 1991) which was integrated into the *AAVS1* safe harbour locus.

Two target genes—*CD2* and *HHEX*—were chosen based on their low expression levels in HEK293T cells, undifferentiated iPS cells, and neuronal cells. *CD2*, encodes a cell surface glycoprotein involved in T-cell adhesion (Binder et al., 2020). The haematopoietically expressed

homeobox 2 gene (*HHEX*) is a transcriptional repressor involved in haematopoietic and hepatic differentiation (Oh and Jang, 2019, Jackson et al., 2023). A further target gene, Msh homeobox 2 (*MSX2*) is expressed at higher levels in pluripotent stem cells and is involved in several developmental pathways such as mesendodermal differentiation (Wu et al., 2015). It was selected as a potential neuroprotective gene based on unpublished results from the Buckley lab (see Chapter 6).

SgRNAs targeting the promoters of *CD2*, *HHEX*, and *MSX2* were designed using the CRISPOR algorithm and were generated as nucleotide sequences containing overhangs for the *BbsI* restriction site, allowing cloning into the U6 promoter-driven SAM sgRNA vector (Addgene, #61424) via Golden Gate cloning. A previously validated non-targeting sequence (NT) was employed as a negative control (Killian et al., 2017, Kazi et al., 2018, Luo et al., 2019).

Following lipofection of the sgRNA vectors into pluripotent KOLFSAM cells, RNA was extracted after 48 hours of incubation. The RNA was reverse transcribed into cDNA, and the expression of the genes of interest was quantified through qPCR.

As illustrated in **Fig. 3.3**, the relative activation of *HHEX* and *CD2* were considerably higher than the signal observed for *MSX2*. This difference in the gene upregulation can be attributed to differences in the basal expression of each gene, with *MSX2* showing higher expression in iPS cells. Previous studies have reported a negative correlation between the basal expression of a gene and the fold change in gene expression caused by CRISPRa-SAM (Konermann et al., 2015, Chavez et al., 2015, Chavez et al., 2016). Nevertheless, a significant upregulation of all three genes compared to the NT control, confirmed that the sgRNAs could be effectively employed for the rest of the study.

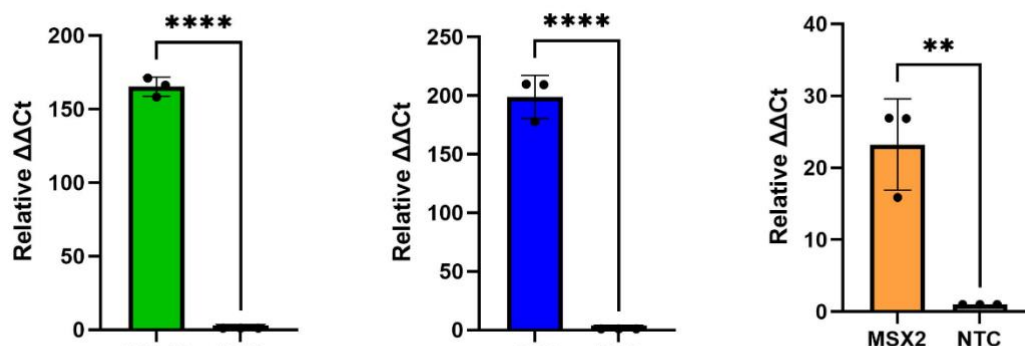


Figure 3.3 Analysis of sgRNA sequences employed in the study. A) Relative expression of *HHEX*, *CD2* and *MSX2* compared to non-targeting control (NTC), following transfection of the designed SAM sgRNAs in KOLFSAM iPS cells. Gene expression analysis revealed that all three primers induced statistically significant upregulation compared to the negative control (NTC). The p-values for the t-tests comparing each primer to the NTC were: > 0.001 , > 0.001 , and 0.0037 , indicating strong evidence of upregulation across all primers tested. *HHEX* (green) and *CD2* (blue) show a stronger gene relative expression than *MSX2* (orange). Error bars represent standard deviation and sample size = 3.

Approaches for inducible guide expression

The next component required by the gene circuit is the controlled expression of the MS2 sgRNA via the tTA inducible system, enabling the targeted upregulation of a gene of interest through CRISPRa-SAM.

I explored three possible methods: firstly, investigating previously reported tTA-inducible Pol III promoters, secondly relying on more conventional tetracycline-inducible Pol II transcripts in combination with methods to cleave out a functional sgRNA from within the transcript either via

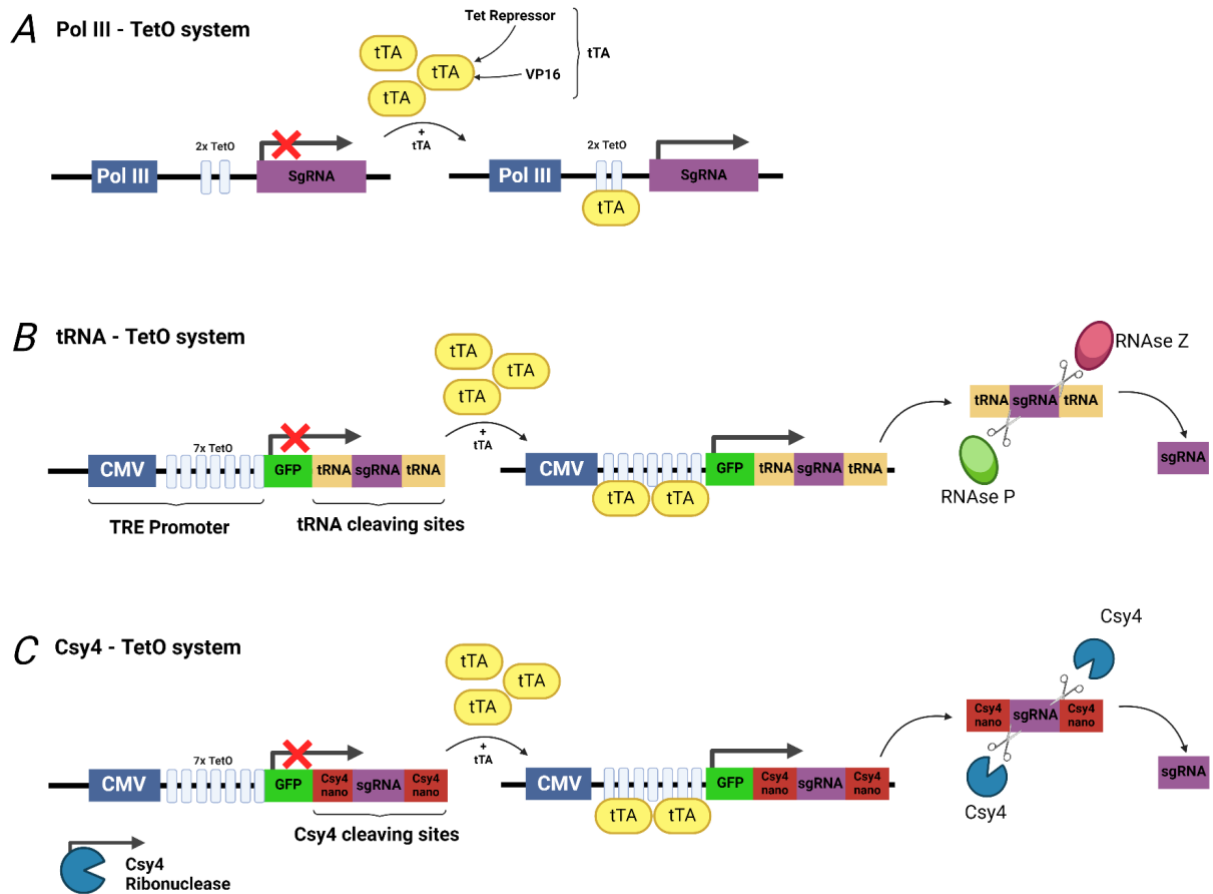


Figure 3.4 Illustration of the tTA-inducible systems to express sgRNA investigated in this study. **A)** Illustration of the Pol III-TetO system. Pol III promoter combined with TetO for tTA-inducible activation of a gene of interest. Two TetO sequences integrated downstream of the Pol III promoter serve as anchoring points for tTA, recruiting the polymerase system. **Addendum:** The logic of the Pol III + TetO system was incorrect, please refer to the explanation in the results below. **B)** Illustration of the tRNA – TetO system. A TRE promoter, consisting of a CMV minimal promoter and seven TetO sequences, used to drive the expression of GFP and an sgRNA embedded between two tRNA sequences. Addition of tTA recruits the polymerase machinery, resulting in the expression of the GFP-tRNA:sgRNA construct. The sgRNA is liberated by the P and Z ribozymes, naturally expressed to cleave tRNA transcripts. **C)** Illustration of the Csy4 – TetO system. This system operated with the same logic as the tRNA system; however, the sgRNA was flanked by two sequences cleaved by the Csy4 endonuclease, liberating the sgRNA post-transcriptionally. Unlike the P and Z ribozymes, Csy4 is not endogenously expressed by eukaryotic cells and thus requires its own exogenous expression.

a tRNA ribozyme-based cleavage, or via the site-specific bacterial ribonuclease Csy4 (**Fig 3.4**).

Pol III – TetO system

The first sgRNA regulatory system explored, involved the combination of a Pol III promoter with two TetO sequences positioned near the promoters (Sun et al., 2019). This method had been previously tested with commonly used Pol III promoters, the human H1 (H1) and human small nuclear U6 (U6) promoter, along with a canonical TetO sequence specifically designed for regulating synthetic short hairpin RNA (Henriksen et al., 2007). To validate this strategy, two plasmids were created: H1 TetO and U6 TetO and used to express the MS2 sgRNA sequence.

The H1 TetO promoter sequence (**Fig. 3.5**) was synthesized (ThermoFisher) and cloned upstream of the SAM sgRNA vector (Addgene, #61424). The plasmid also featured two *BbsI* restriction sites directly following the promoter for the insertion of the sgRNA (**Fig. 3.5A**). The previously validated *CD2* sgRNA, serving as a test guide, was integrated into the plasmid using Golden Gate cloning via these *BbsI* restriction sites. The completed plasmid was validated by diagnostic digestion and Sanger sequencing (**Fig. 3.5B**).

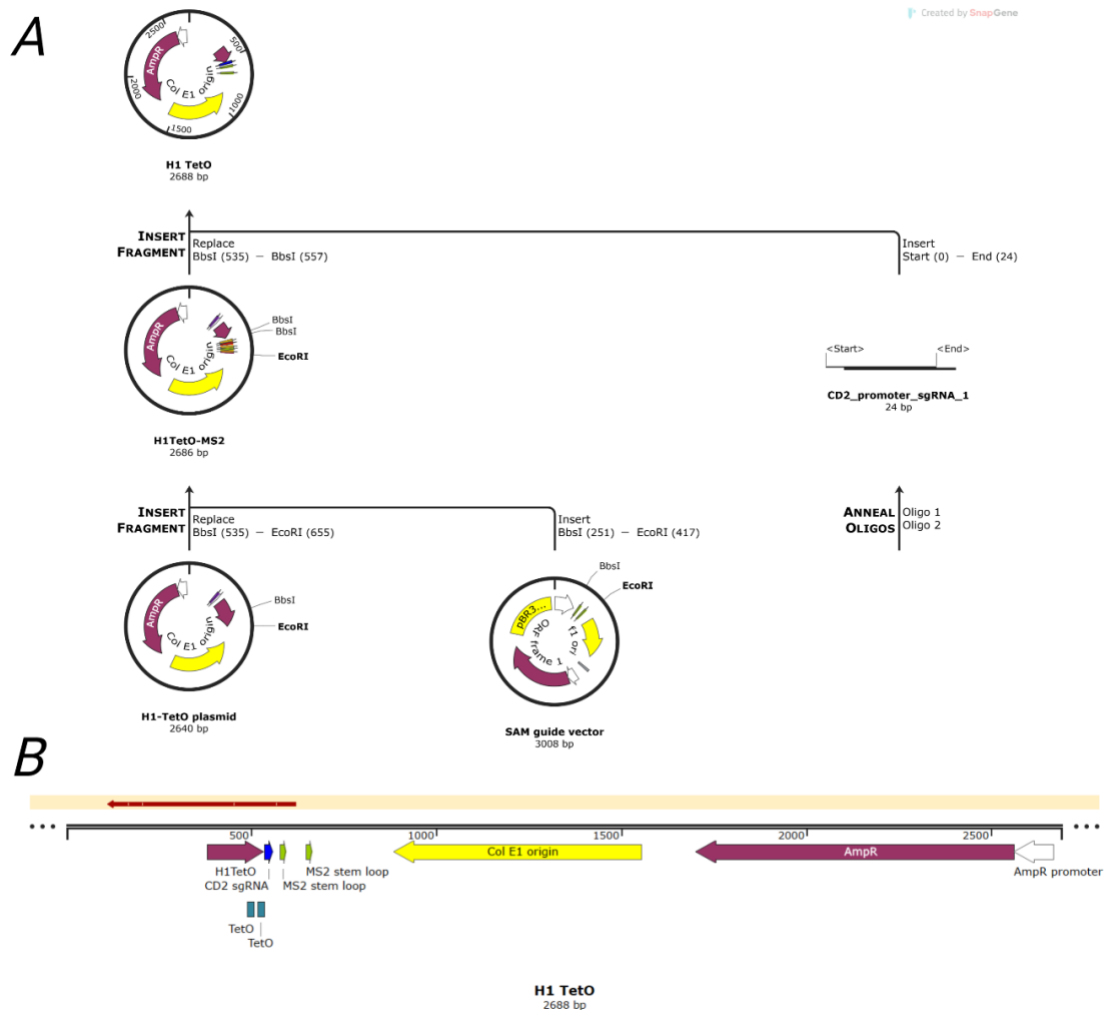


Figure 3.5 Cloning of H1 TetO plasmid. A) Cloning steps required for the generation of a H1 promoter fused to TetO sequences driving the expression of the CRISPRa-SAM sgRNA, followed by the integration of the CD2 sgRNA spacer sequence. B) Linearised H1 TetO vector driving the expression of the CD2 sgRNA. The red arrow shows the sequenced portions of the plasmid confirming the integration of the H1 TetO promoter and portion of the CD2 sgRNA.

A different approach was followed for the cloning of U6 TetO. Given that the original SAM sgRNA vector (Addgene, #61424) already contained the U6 promoter, a synthetic fragment comprising of two TetO sequences (ThermoFisher) was introduced downstream of the promoter and before the *BbsI* restriction sites (**Fig. 3.6A**) to recapitulate the promoter fusion reported by Sun et al. (2019). Subsequently, the CD2 sgRNA sequence was inserted by Golden Gate cloning to finalise the plasmid. The completed U6 TetO plasmid was confirmed by diagnostic digestion and Sanger sequencing (**Fig. 3.6B**).

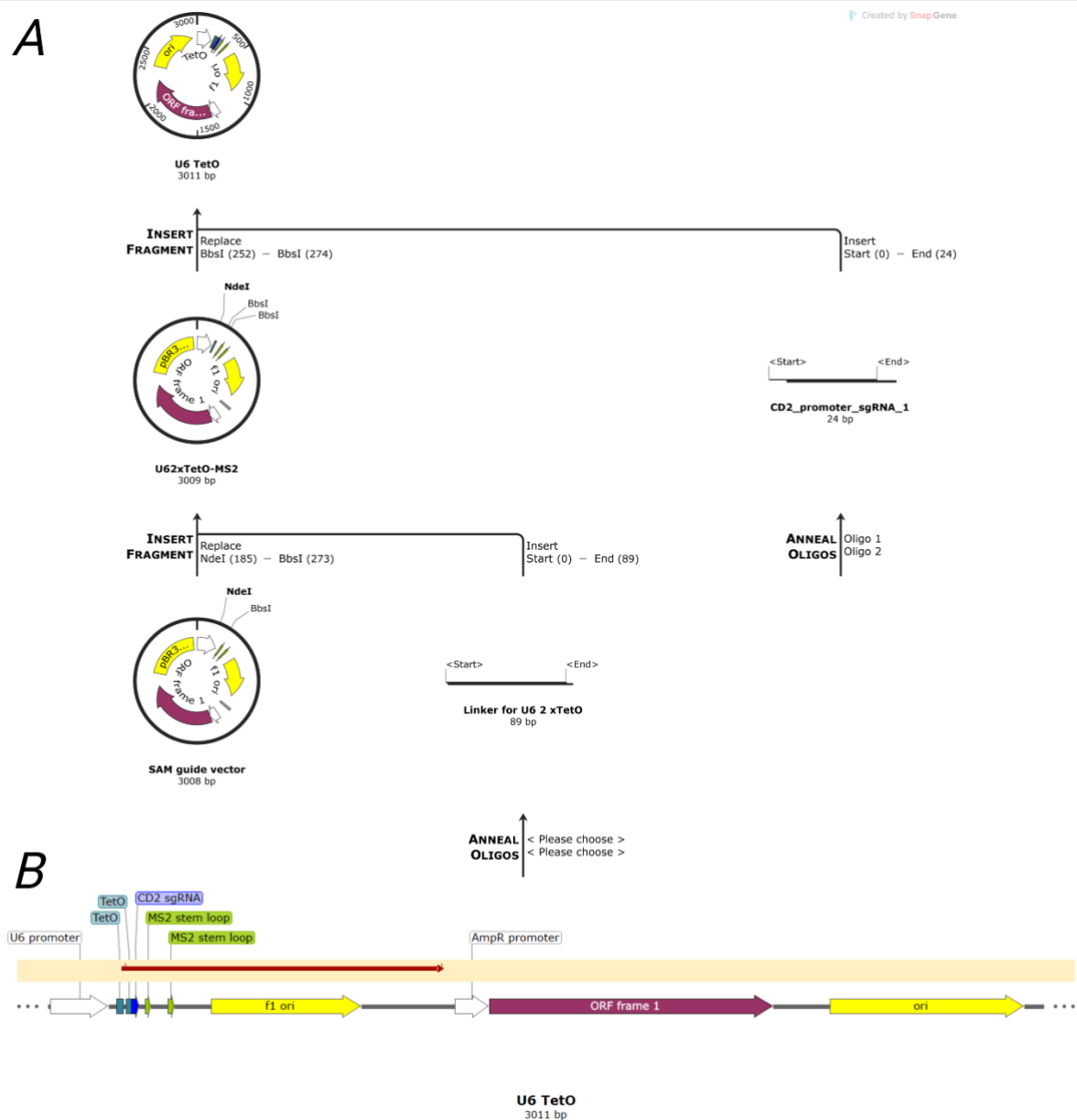


Figure 3.6 Cloning of U6 TetO plasmid A) cloning steps required for the generation of a U6 promoter fused the TetO sequences driving the expression of the CRISPRa-SAM sgRNA. Followed by the integration of the CD2 spacer

sequence. B) Linearised U6 TetO vector after cloning. The red arrow shows the sequenced portions of the plasmid confirming the integration of the TetO sequences and the CD2 sgRNA.

Both H1 TetO and U6 TetO plasmids were tested to determine whether the interaction with tTA was able to induce the expression of the CD2 sgRNA.

HEK293T cells were co-transfected with the BATV2 plasmid (see *Chapter 4*), providing the CRISPRa-SAM machinery, in conjunction with the completed H1 TetO or U6 TetO plasmids. The tTA plasmid was delivered to simulate the tTA-dependent activation, and compared with an empty vector. In the absence of tTA, it was anticipated that H1 TetO and U6 TetO would not express the CD2 sgRNA. The presence of tTA was expected to interact with the TetO sequences embedded in the promoters, inducing the expression of CD2 sgRNA and, consequently, the upregulation of the gene.

As shown in **Fig. 3.7A** the experiment yielded the opposite results. Cells transfected with H1 TetO and U6 TetO, in the absence of tTA, showed a robust upregulation of the endogenous CD2 gene, suggesting abundant sgRNA expression under these conditions. Conversely, the addition of tTA led to no upregulation of CD2, suggesting that the addition of tTA seemed to suppress the expression of the sgRNA.

To further investigate this phenomenon, the same experiment was repeated. However, 24 hours post-transfection, half of the samples were exposed to 1 µg/mL doxycycline, while the rest were exposed to PBS. As discussed before, doxycycline exposure causes a conformational change in tTA, preventing its binding to the TetO sequences. As illustrated in **Fig. 3.7B**, the combination of tTA and the H1 TetO or U6 TetO plasmids led to no upregulation, as before. However, the addition of doxycycline rescued the CD2 upregulation, suggesting sgRNA expression was occurring when the tTA binding to the TetO was inhibited. These results confirmed that the system behaved entirely opposite to the initial assumption, and tTA is acting as a transcriptional repressor when bound to the Pol III promoter fusions, as discussed in detail at the end of chapter. The results

indicated that an alternative approach for achieving tTA-activated sgRNA expression needed to be adopted.

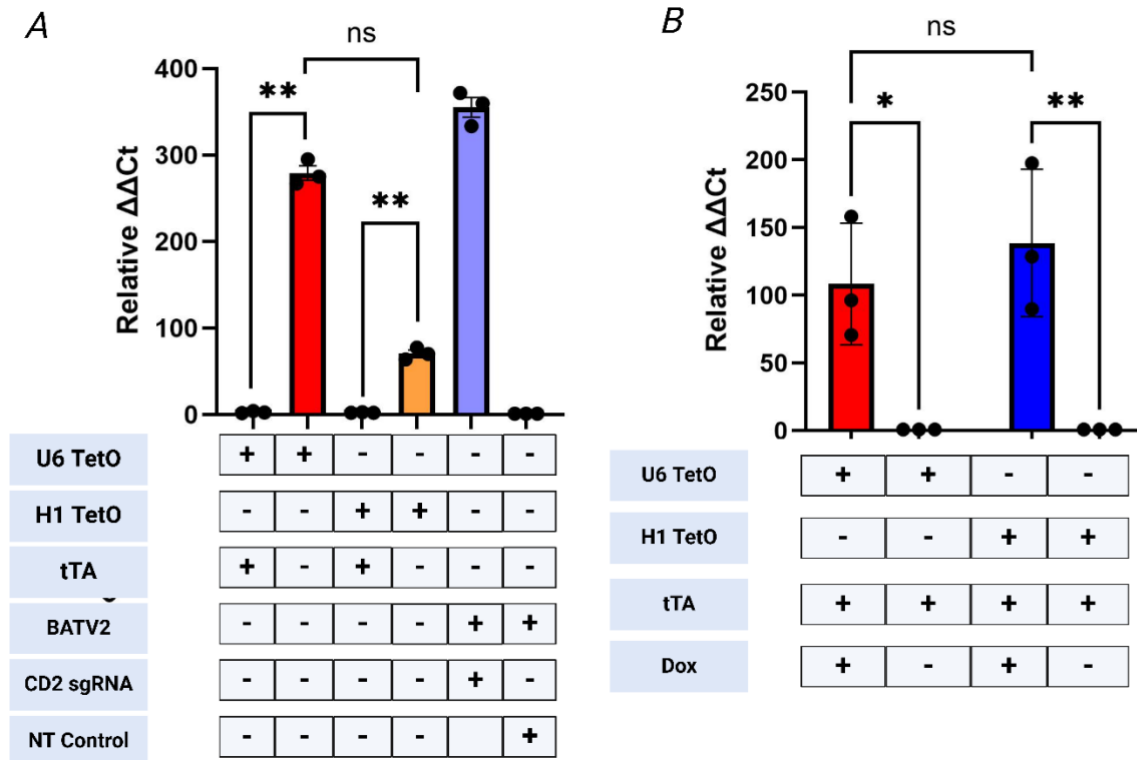


Figure 3.7 Investigation of H1 and U6 TetO plasmids **A)** Relative expression of CD2 compared to the negative control (NTC) in HEK293T cells transfected with the CRISPRa-SAM machinery combined with U6_TetO or H1_TetO promoters, with or without the addition of the tTA plasmid. A two-way ANOVA revealed that both promoters and the presence of tTA significantly influenced gene upregulation [Promoter type: $F(3,13) = 40.19$, $p < 0.001$; tTA presence: $F(1,13) = 34.25$, $p < 0.001$]. Tukey's multiple comparisons test indicated that gene expression driven by both U6_2xTetO and H1_TetO promoters was significantly increased without tTA ($p = 0.0011$). However, no significant difference was observed between the promoters themselves ($p = 0.054$). Error bars represent the standard deviation. **B)** Relative expression of CD2 compared to NT control in HEK293T cells transfected CRISPRa-SAM machinery and the tTA plasmid combined with: U6 TetO, H1 TetO and with or without the exposure of 1 μ g/mL of doxycycline. A two-way ANOVA revealed that the presence of doxycycline significantly influenced the activation of the promoters [$F(1,8) = 36.12$, $p = 0.0003$]. A Tukey's multiple comparisons test indicated that addition of doxycycline resulted in a significant upregulation of CD2 driven by both U6 TetO and H1 TetO promoters, compared to when doxycycline is not present ($p = 0.0242$ and $p = 0.0061$ respectively). Error bars represent standard deviation and sample size = 3.

tRNA-TetO system

As previously discussed, tetracycline-regulated gene expression is well characterised for Pol II transcripts, using a TetO system (Gossen and Bujard, 1992). In this context, I explored methods wherein the sgRNA could be cleaved from the Pol II-expressed tetracycline-regulated transcript. The first method exploited the natural tRNA processing mechanism, allowing the sgRNA to be embedded in a pre-tRNA sequence which is excised following transcription (Xie et al., 2015). A number of studies demonstrated the effectiveness of this system when combined with a Pol II promoter (Port and Bullock, 2016, Xu et al., 2016). Knapp et al. (2019) systematically optimized the human tRNA backbone sequence for improved expression in mammalian cells. This study determined the minimal components for the optimal tRNA:sgRNA scaffold as well as specific

mutations enhancing its efficiency, such as the $\Delta C55G$ mutation. This system was demonstrated to be effective in processing the flanked sgRNA expressed through a Pol II promoter, presenting itself as a potentially suitable tool for inducible sgRNA expression via the tTA system.

A new plasmid was synthesized (ThermoFisher), incorporating the $\Delta C55G$ minimal tRNA backbone flanking the *CD2* sgRNA, along with the entire MS2 scaffold required for CRISPRa-SAM upregulation, followed by a polyadenylation sequence. This plasmid was then combined with the TRE-TIGHT sequence (Takara™), featuring an optimized version of the TRE (TIGHT promoter) derived from the CB9-CB-rtTA-Ins-TRE-SMN plasmid previously generated in the Davies lab. Finally, a GFP reporter was introduced from the pCX-eGFP plasmid (Addgene, #15007) (**Fig. 3.8A**). The successful cloning of the plasmid (referred to as tRNA:sgRNA) was validated through restriction digestion and Sanger sequencing (**Fig. 3.8B**).

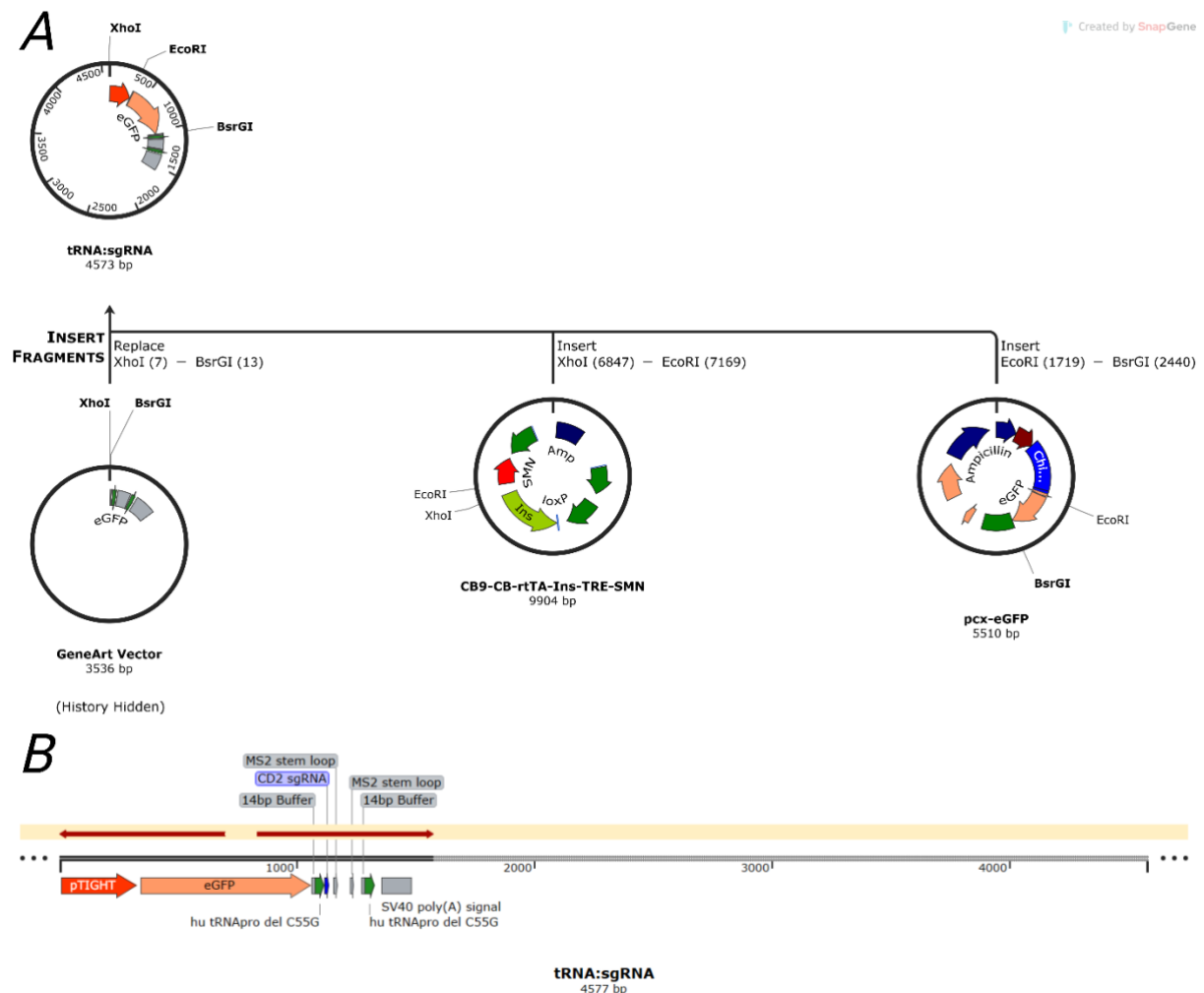


Figure 3.8 Cloning of tRNA:sgRNA plasmid. **A)** Cloning steps required for the generation of a Pol II TIGHT promoter driving the expression of a *CD2* sgRNA embedded in a self-cleaving tRNA ribozyme sequence. **B)** Linearised tRNA:sgRNA vector driving the expression of *CD2* sgRNA after cloning. The red arrow shows the sequenced portions of the plasmid confirming the integration of the TIGHT promoter followed by a GFP reporter gene and the *CD2* sgRNA embedded between the two tRNA sequences carrying the $\Delta C55G$ mutation.

Next, I investigated the efficiency of tTA to induce the expression of tRNA:sgRNA. HEK293T were co-transfected with BATV2, providing the CRISPRa-SAM machinery in combination with the tRNA:sgRNA plasmid and either the tTA or an empty plasmid. *CD2* expression was investigated by qPCR using reverse-transcribed RNA and harvested 48 hours following transfection. As illustrated in **Fig. 3.9**, *CD2* expression was not detected in the tRNA:sgRNA transfected cells, with or without the addition of tTA. Contrary to the successful expression of anagRNA using the tRNA scaffold reporter by Knapp et al. (2019), I was unable to show any upregulation of the *CD2* gene, suggesting that the mature sgRNA was not efficiently generated using my inducible construct.

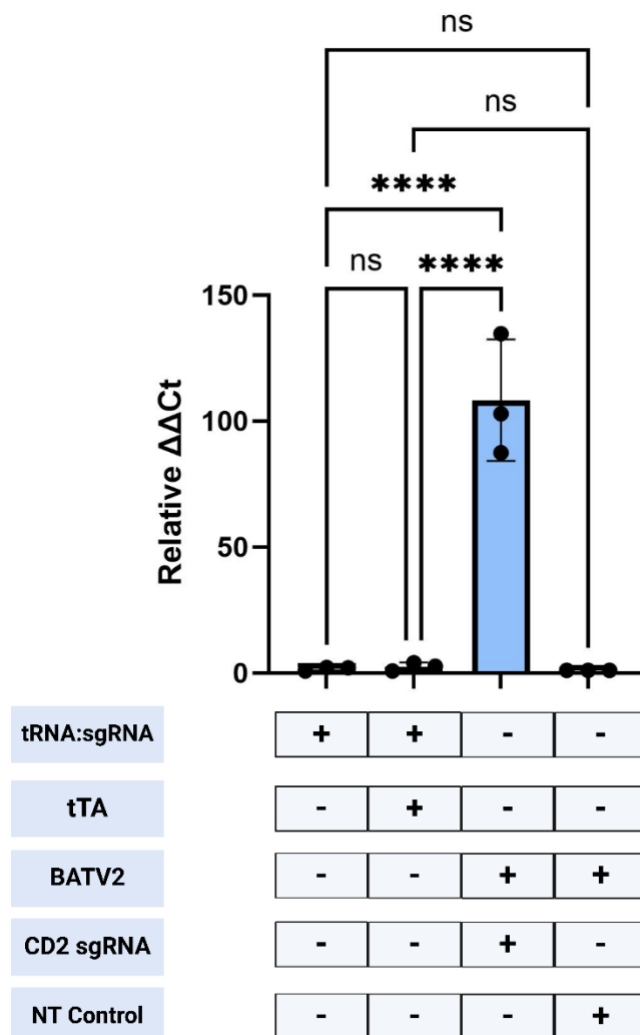


Figure 3.9 Testing of tRNA:sgRNA plasmid Relative expression *CD2* compared to NT control in HEK293T cells transfected with the BATV2 plasmid combined with: sgRNA:tRNA with and without tTA. A one-way ANOVA indicated a significant variation in *CD2* expression [F (3,8) = 58.38, $p < 0.0001$], due to positive control. A Tukey's multiple comparisons test showed no significant difference in *CD2* upregulation in the tRNA:sgRNA transfected cells, with or without tTA ($p = 0.9998$) as well as between tRNA:sgRNA w/ tTA or w/o tTA and the Non-Targeting control ($p = 0.9999$ and $p = 0.9986$, respectively). Errors bars represent standard deviation and sample size = 3.

Csy4 – TetO system

Knapp et al. (2019) also compared the effectiveness of different systems in releasing sgRNA for CRISPRa-SAM and demonstrated that the Csy4-induced system resulted in high fold change in upregulation while maintaining low background noise. This suggests that the Csy4 system might be a suitable alternative for the gene circuit.

To investigate Csy4-induced cleavage of the test *CD2* sgRNA, a new sequence was synthesized (ThermoFisher). The sequence featured the *CD2* sgRNA, with MS2 aptamer sequences, flanked by two nano-Csy4 cutting sites, demonstrated to be more effectively processed by Cys4 for the release of sgRNAs (Ferry et al., 2017). The sequence was incorporated into the tRNA:sgRNA plasmid which already contained the TIGHT promoter driving the expression of GFP. Finally, to complement the Csy4-mediated loss of the polyadenylation tail, the 3' end sequence of the mouse MALAT1 locus gene was added between the GFP and the first Csy4 site (**Fig. 3.10A**). The MALAT sequence was shown to stabilize transcripts lacking polyadenylation tails by generating a highly conserved 3' triple helical structure (also known as triplex) (Wilusz et al., 2012, Nissim et al., 2014). The plasmid – referred to as Csy4:sgRNA – was validated through double digest and Sanger sequencing (**Fig. 3.10B**). Finally, as mammalian cells do not express the Csy4 ribonuclease, the TR-Csy4 plasmid (Addgene, #80601) was used to supply its exogenous expression.

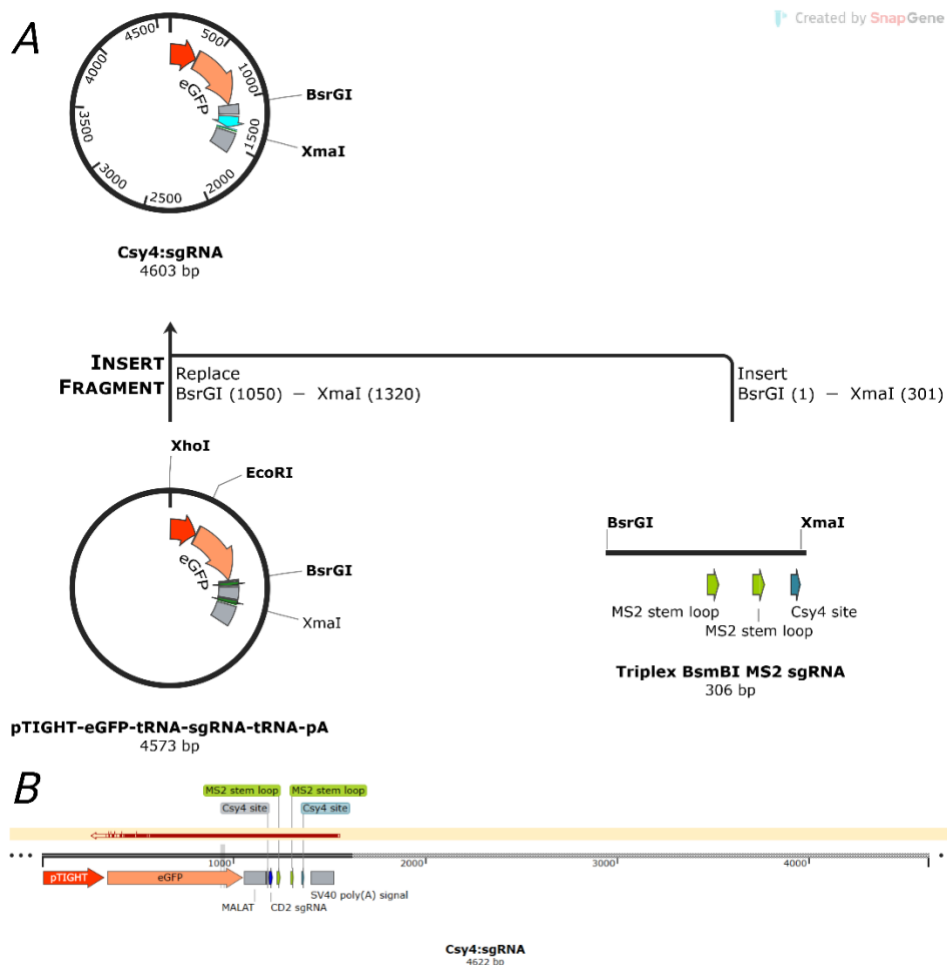


Figure 3.10 Cloning of Csy4:sgRNA plasmid **A)** Cloning steps required for the generation of a TIGHT promoter driving the expression of the CD2 sgRNA flanked by two Csy4 cleavage sites. **B)** Linearised Csy4:sgRNA vector driving the expression of CD2 sgRNA after cloning. The red arrow shows the sequenced portions of the plasmid confirming the integration of the MALAT sequence and the sgRNA flanked by Csy4 sites.

To assess the efficiency of the Csy4:sgRNA inducible mechanism, HEK293T cells were transfected with the BATV2 plasmid, and combinations of Csy4:sgRNA without tTA or Csy4, Csy4:sgRNA with tTA, Csy4:sgRNA with Csy4, and Csy4:sgRNA with both tTA and Csy4, along with positive and negative controls (**Fig. 3.11**). As depicted in **Fig. 3.11**, it was hypothesized that the

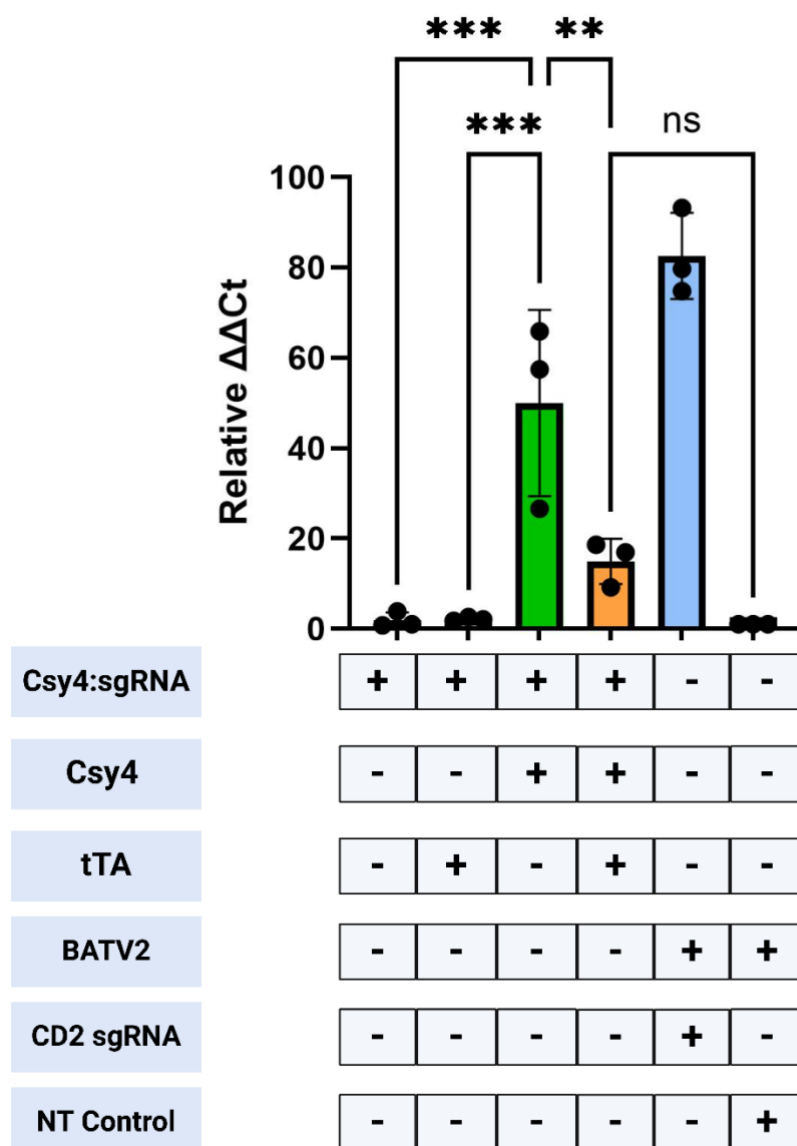


Figure 3.11 Testing of Csy4:sgRNA plasmid. Relative expression of CD2 compared to NT control in HEK293T cells transfected with the CRISPRa-SAM machinery combined with: Csy4:sgRNA with/without tTA and with/without Csy4 ribonuclease. A one-way ANOVA revealed that the Csy4:sgRNA plasmid induced different level of CD2 upregulation based on the combination of additional plasmids [$F(5,12) = 37.37, p < 0.0001$]. A Tukey's multiple comparison test suggests that no significant difference in upregulation can be noticed between the Csy4:sgRNA plasmid only and in combination with tTA ($p = >0.9999$), suggesting that the system was strong enough to avoid baseline activation. The test also demonstrated that the combination of Csy4:sgRNA and Csy4 plasmids induced strongest CD2 upregulation, relative to: Csy4:sgRNA only ($p = 0.0005$), Csy4:sgRNA + TetOFF ($p = 0.0006$) and interestingly Csy4:sgRNA + Csy4 + TetOFF ($p = 0.0073$) which was supposed to induce the strongest upregulation. The Csy4:sgRNA + Csy4 + TetOFF resulted non-significant compared to the negative control ($p = 5066$). Error bars represent standard deviation and sample size for each group = 3.

Csy4:sgRNA system would function as an AND gate, necessitating the presence of both tTA to induce the expression of the Csy4:sgRNA sequence and Csy4 to liberate the sgRNA.

48 hours after transfection, total RNA was extracted from the cells, and the cDNA was synthesized for the analysis of CD2 gene expression by qPCR. As shown in **Fig. 3.11**, cells transfected with the Csy4:sgRNA plasmid alone or with tTA did not exhibit CD2 upregulation. This

indicates the effective prevention of interactions between the embedded sgRNA sequence and CRISPRa-SAM within the primary transcript, in the absence of the Csy4 ribonuclease.

In contrast, cells co-transfected with Csy4:sgRNA and Csy4 showed a significant upregulation in *CD2* expression. Furthermore, *CD2* expression in cells co-transfected with Csy4:sgRNA, tTA, and Csy4 also resulted in substantial upregulation compared to both the negative control and cells transfected with Csy4:sgRNA only, which was however non-significant, due to the difference in variance between the Csy4:sgRNA, tTA and Csy4 samples and the negative controls. Although non-significant the combined results from Csy4:sgRNA + Csy4 and Csy4:sgRNA + Csy4 + tTA, especially compared to the minimal baseline levels of the Csy4:sgRNA only and the Csy4:sgRNA + tTA suggest that the Csy4-TetO system is an effective mechanism to induce sgRNA expression, with minimal to no background noise and ~20-fold change upregulation when activated.

Nevertheless, the considerable upregulation of *CD2* following Csy4:sgRNA and Csy4 ribonuclease co-transfection in the absence of tTA, implies that the Tet system is affected by background expression. This explanation suggests that the primary transcript is expressed by a leaky TIGHT promoter but remains inactive until the Csy4 nuclease releases the sgRNA, resulting in *CD2* upregulation. The higher upregulation seen in the Csy4:sgRNA + Csy4 cells compared to the Csy4:sgRNA + tTA + Csy4 cells was also unexpected, given the anticipated behaviour that tTA would induce a higher level of gene expression from the TIGHT promoter driving the primary transcript.

Despite the potential leakage of the TIGHT promoter, the Csy4 system emerged as the most effective mechanism for inducible expression in RNA. The system demonstrated minimal to no background noise when in its “OFF” state and significant gene upregulation when activated. Based on these results, it was concluded that Csy4:sgRNA would be the most suitable system to be incorporated into the final gene circuit.

Closing Remarks

In this chapter, I validated the components that would serve as the foundation for the final gene circuit. CMKLR1-Tango was demonstrated to effectively interact with A β and induce the downstream activation of reporter genes. Furthermore, the CMKLR1-Tango's activation through chemerin and A β exposure was found to be specific, with minimal affinity compared to a control FPRL1-Tango receptor. This validation established the CMKLR1-Tango system as a viable sensing component of the final circuit, capable of detecting extracellular A β .

I validated the sgRNA sequences targeting three genes—*CD2*, *HHEX* and *MSX2*. *CD2*, lacking expression in HEK293T and pluripotent stem cells, was chosen as an output reporter for testing individual components and more complex gene circuit combinations in subsequent phases. *HHEX* sgRNA was also confirmed to induce robust gene expression, making it suitable for testing the CRISPRa-SAM system during iPSC differentiation into neurons. Finally, although *MSX2* sgRNA induced a significant upregulation of the gene, it was modest compared to the other two, potentially due to its higher basal expression. Nevertheless, *MSX2* sgRNA proved effective, enabling the investigation of its upregulation in relation to its putative role as a neuroprotective gene (see Chapter 6).

A component regulated by tTA was needed to complete the circuit, capable of preventing unwanted gene upregulation in its OFF state. Studies have shown that the fusion of a Cas9 molecule to the Tango system could be used as an alternative mechanism to control CRISPR activation. However, the same studies reported significant levels of leakiness, attributed to the basal signal activity of the receptor and the basal cleavage of the TEV site, making the release of Cas9 less optimal (Kipniss et al., 2017). Leakiness was improved by Baeumler et al. (2017) by splitting the Cas9 into two components each tethered to a different receptor, a strategy that would have posed serious constraints to the design of my circuit due to its size.

I explored an alternative approach—tTA-inducible sgRNA expression. Three strategies were investigated: fusion of a Pol III promoter (H1 or U6) with TetO sequences or embedding sgRNA into a larger mRNA sequence using a tetracycline-regulated Pol II promoter followed by either tRNA post-transcriptional or Csy4 cleavage.

Firstly, I investigated the fusion of H1 and U6 Pol III promoters to TetO sequences (de Solis et al., 2016, Kumar et al., 2018, Sun et al., 2019). However, I found that the activation of these Pol III – TetO promoters was not feasible with tTA. The canonical TRE Pol II promoter relies on a combination of the minimal CMV promoter and several TetOs. The CMV is unable to induce gene expression without the support of the tTA, where the tetracycline portion of the molecule binds to the TetO sequence and the fused VP16 domain recruits the polymerase system. H1 or U6 promoters are instead fully active promoters that support sgRNA expression in the absence of tTA. Therefore, the presence of tTA prevents the polymerase system from interacting with the promoter, effectively functioning as a transcriptional block, reverting the Pol III tTA system to the original tetracycline repressor complex in its native *E.coli*.

Given the established success of inducible Tet-Pol II systems, I decided to concentrate on utilizing a Pol II promoter for the gene circuit, as it offers a well-established platform. To achieve

sgRNA expression from a Pol II promoter, the sgRNA needed to be incorporated into a larger mRNA sequence. Initially, I explored the use of the tRNA post-transcriptional modification mechanism, leveraging ribozymes (P and Z), which naturally excise specific RNA sequences. This mechanism is endogenous to mammalian cells, streamlining the gene circuit design by requiring only the addition of a tRNA:sgRNA sequence.

Moreover, the tRNA scaffold had been investigated in depth by Knapp et al. (2019) and optimized to uncouple the tRNA expression from its native Pol III promoter allowing expression from other Pol II promoters such as CMV. However, when combined with a TIGHT promoter, the tRNA:sgRNA system did not exhibit any upregulation of *CD2*, regardless of the presence of tTA. The lack of sgRNA expression in this system raises the need for further investigation to pinpoint the exact cause. The study by Knapp et al. (2019) relied on the CMV promoter as a core element upon which the optimization of the tRNA relied. It is possible that the lack of *CD2* upregulation might be the result of the unoptimized combination of the Δ C55G tRNA excision mechanism with the tTA inducible system which could have led to an inefficient cleavage. For example, potentially, the expression level is critically important for effective tRNA processing, and thus, the use of a different promoter in my work driving expression at a different level from the CMV promoter used by Knapp et al., 2019, may provide some explanation. Another explanation may be provided by the differences in sgRNA structure, with the MS2 aptamer sequences present in the precursor mRNA presenting difficulties for efficient tRNA processing.

In the final exploration of inducible systems, the Csy4 cleavage mechanism emerged as a promising strategy. Csy4, a ribonuclease, naturally functions in cleaving CRISPR RNA by targeting unique 28nt sequences, making it an optimal tool for sgRNA cleavage. The comparative study, by Knapp et al. (2019) revealed that the Csy4 system demonstrated an effective induction-to-low-background-noise ratio when compared to the Δ C55G tRNA system. Despite the initial hesitation due to the need for integrating the Csy4 ribonuclease into the circuit, as it is not naturally expressed in eukaryotic cells, and concerns about potential cellular toxicity at high concentrations (Nissim et al., 2014), the Csy4:sgRNA inducible system, when combined with the TIGHT promoter, yielded superior results. This system exhibited no sgRNA expression in the absence of the Csy4 RNase and was able to achieve significant, albeit modest, *CD2* target gene upregulation.

The combination of the Csy4 and TetO sequences was thought to establish a double-lock mechanism, offering what I believed would have been a more sensitive control over the final gene circuit. However, an interesting observation was made when comparing the outcomes of tTA and

Csy4 co-expression with Csy4 alone. Surprisingly, the *CD2* expression fold change was higher in the presence of Csy4 only, compared to the combination of tTA and Csy4. A number of factors could account for this result. Firstly, while the total DNA mass was kept constant by the addition of an empty vector, plasmid size was not considered, potentially resulting in a better transfection rate of the Csy4:sgRNA and Csy4 plasmids when transfected with a smaller empty vector than with the three Csy4:sgRNA /tTA/Csy4 plasmids. Additionally, the expression of exogenous genes can impact the cell's metabolism by diverting energy expenditure (Bolognesi and Lehner, 2018). It is possible that the combination of both Csy4 and tTA had a more pronounced effect on the cell's metabolism and protein translation than Csy4 alone. Finally, a structural explanation is worth considering, where the recruitment of the polymerase complex by tTA may have somehow impaired the interaction of Csy4 with the mRNA. Nevertheless, the upregulation of *CD2* in the presence of Csy4 only confirms that the TIGHT promoter still exhibits residual leakiness. This result suggested the need for further optimization to enhance the effectiveness of regulation and eliminate any remaining background expression in the system.

Nonetheless, having validated the three primary components of the gene circuit, the subsequent step involved the cloning of constructs to assemble the final circuit in the following Chapter.

Chapter 4 Design and validation of the final circuit

Introduction

The previously validated components were arranged in a sequential manner allowing the translation of the input into the expected output (**Fig. 4.1**). In its resting state, the circuit would be inactive. The sgRNA expression, being inhibited by the lack of tTA, still fused to the CMKLR1-Tango receptor (**1**). The ligand-binding interaction between A β and CMKLR1 would trigger the recruitment of β A-TEV protease. The proximity of β A-TEV to the TEV cleavage site, would result in the release of tTA (**2**). Interaction of the tTA with the TetO sequences would then trigger the Polymerase system, inducing the expression of the sgRNA. The sgRNA would then complex with the CRISPRa-SAM machinery (**3**). CRISPRa-SAM would then upregulate expression from the endogenous target gene (**4**).

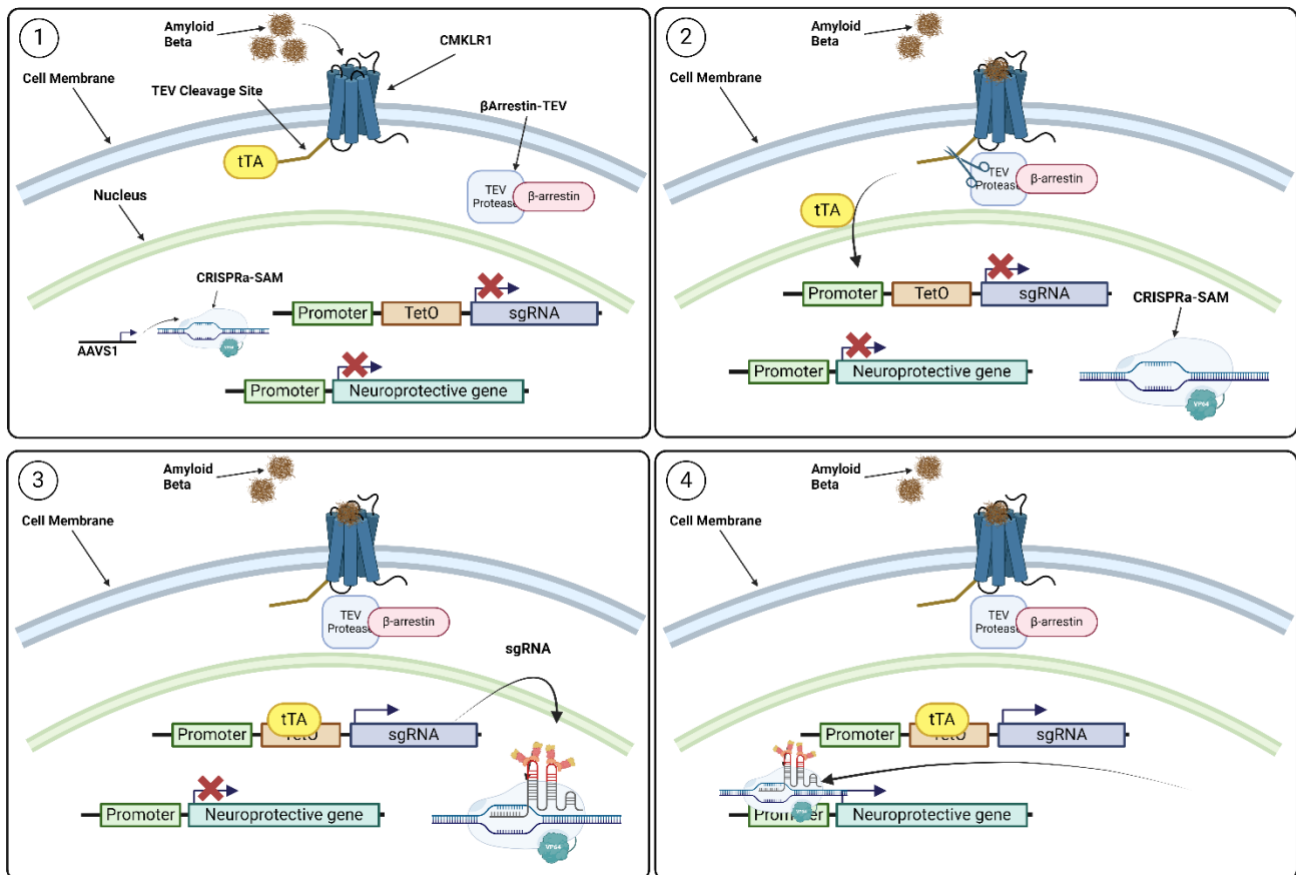


Figure 4.1 Diagram of stepwise circuit activation. It illustrates the sequential steps required to activate the circuit by interaction with A β and subsequent upregulation of neuroprotective gene as output signal. **1)** sgRNA expression is inhibited by the lack of tTA required to activate the promoter. tTA is fused to the Tango receptor via a TEV cleavage site. Lack of sgRNA results in an inactive CRISPRa-SAM. **2)** Ligand binding interaction between the receptor and A β results in the recruitment of β -arrestin fused to a TEV protease (β A-TEV) which releases the tTA. **3)** Interaction of the tTA with the promoter induces the expression of sgRNA which can interact with the rest of the CRISPRa-SAM machinery. **4)** The CRISPRa-SAM machinery travels to the relevant site directed by the sgRNA and induces expression of the gene of interest.

While the core elements have been validated individually, integrating these discrete components into a unified and linear system can be challenging. Specifically, stochastic dynamics between the circuit and endogenous genetic landscape as well as between each individual component, can significantly impact the functionality of the circuit (Zhang et al., 2016a). Individual factors, such as protein degradation rates and promoter strength, can vary between components. In fact, one of the most common points of failure in gene circuit engineering occurs when the output of an upstream component does not reach the dynamic threshold required to stimulate the subsequent one, resulting in a signal decay (Brophy and Voigt, 2014).

Based on this information, it is crucial to investigate the behaviour of the complete gene circuit before its integration into the genome of iPSC. Finally, due to the size of the final circuit, I decided to split the components into two separate plasmids. The plasmids were designed to be integrated into the *AAVS1* and *CLYBL* safe harbour loci (Smith et al., 2007, Cerbini et al., 2015a)(**Fig. 4.2**). In this section, I will outline the construction and testing of targeting vectors suitable for integrating the circuit within the human genome.

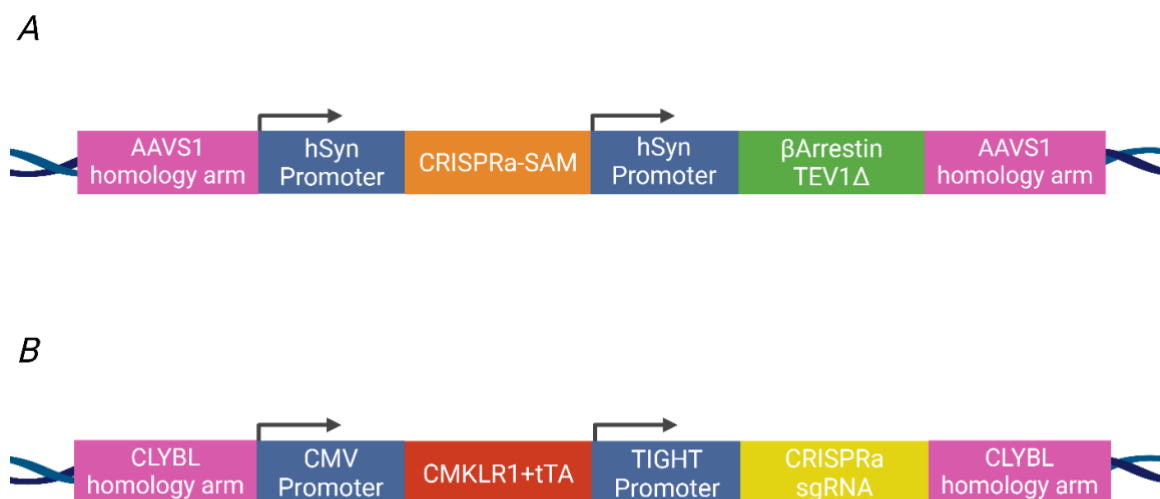


Figure 4.2 Diagram illustrating the two genetic sequences composing the complete circuit. A) Outlines the structure of the first sequence composed of the CRISPRa-SAM machinery and the β Arrestin-TEV sequence, both under the control of the hSyn promoter. This construct was designed for integration into the AAVS1 safe harbour locus. B) Outlines the structure of the second sequence composed of the CMKLR1-Tango system, driven by a CMV promoter and the Csy4:sgRNA sequence regulated by a tTA inducible TIGHT promoter. This construct was designed for integration into the CLYBL safe harbour locus.

Design of the BATV2 plasmid

The first targeting vector (named BATV2) was designed to achieve the integration of the CRISPRa-SAM machinery and the β A-TEV expression cassette, into the *AAVS1* safe harbour locus. The hSyn

promoter was used to drive the expression of both components with the intention of inducing expression in iPSC-derived neurons (see *Chapter 5*). Moreover, a puromycin selection cassette was incorporated to enable the selection of recombinant cells (**Fig. 4.3A**) and the entire sequence was designed to be flanked by homology arms to target the integration into the first intron of *AAVS1* (**Fig. 4.3B-C**).

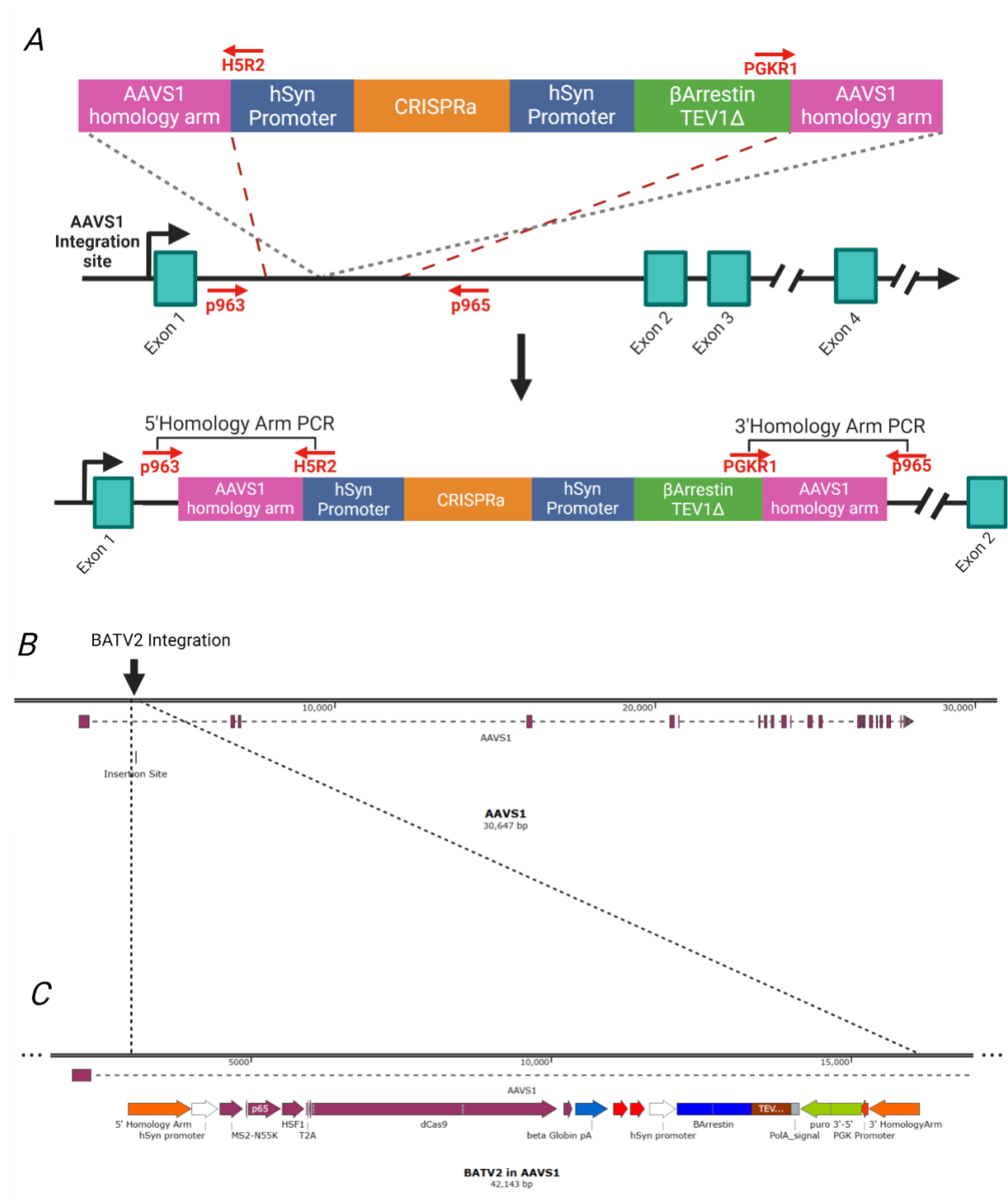


Figure 4.3 Diagram illustrated the inserted exogenous sequence into the first intronic region of *AAVS1*. **A)** The sequence contains two transgenes: CRISPRa-SAM machinery and the β A-TEV protease driven by the human synapsin promoter (hSyn). The TEV protease was optimized by integrating a point mutation, making the final molecule β A-TEV1 Δ . The primers employed to investigate correct integration of the sequence at the 5' and 3' ends, are delineated by red arrows. One primer in the set is located in the WT section and the other in the transgene. **B)** Complete map of the *AAVS1* safe

harbour locus. The map shows intronic regions defined by the dotted lines interspersed by the exonic regions defined by the red boxes (bottom). **C)** Integration of the BATV2 sequence into the first intron of AAVS1

The construction of the BATV2 targeting vector involved several steps distributed across three parallel assemblies resulting in the final cloned plasmid.

Assembly 1 led to the generation of a plasmid containing the β A-TEV1 Δ sequence driven by the hSyn promoter. The pcDNA3.1(+)-CMV-bArrestin2-TEV plasmid (Addgene, #107245), was sourced from the Roth lab. It already contained the β A-TEV sequence tested in the Tango system (Kroeze et al., 2015) and is driven by the CMV promoter. As shown in **Fig. 4.4A**, the CMV promoter was substituted with a synthetic hSyn promoter sequence. Additionally, Sanchez and Ting (2020) demonstrated enhanced catalytic activity of a truncated S153N-mutated TEV protease named TEV1 Δ . Consequently, the plasmid was further refined by replacing the original TEV sequence with the TEV1 Δ protease, also generated synthetically. Confirmation of the new plasmid, hSyn- β A-TEV1 Δ , was achieved through diagnostic digestion and Sanger sequencing (**Fig. 4.4B**).

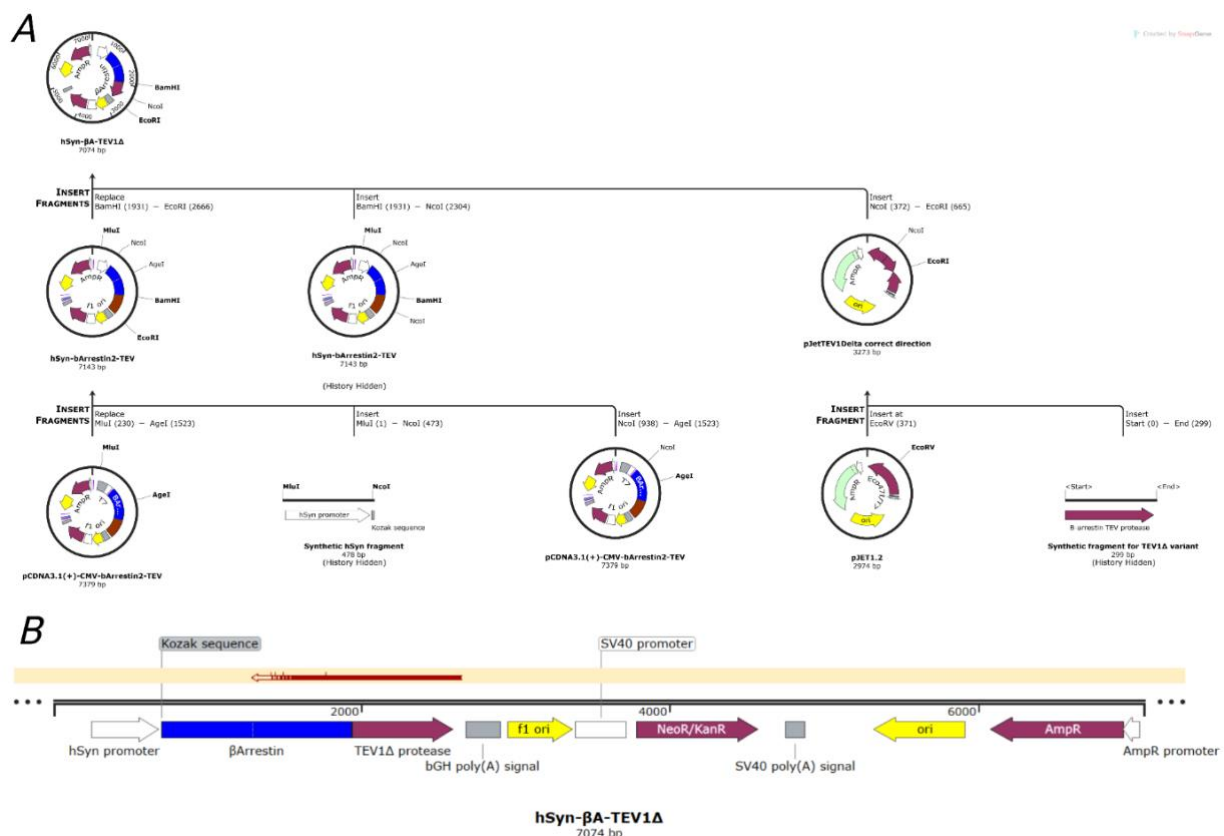


Figure 4.4 Assembly 1 of BATV2. **A)** Cloning steps required to replace the CMV promoter for the synthetic hSyn promoter driving the β A-TEV expression as well as the replacement of the sequence encoding the wild type of TEV protease for an optimized version carrying the S153N mutation (TEV1 Δ). **B)** Linearised hSyn- β A-TEV1 Δ vector after cloning. The red arrow shows the sequenced portion of the plasmid confirming the integration of the β Arrestin sequence, the TEV1 Δ protease and the polyadenylation sequence.

Simultaneously, **Assembly 2** focused on the integration of insulating sequences between the CRISPRa-SAM machinery and the β A-TEV1 Δ . For this, two chicken β -globin DNase-I hypersensitive site 4 regulatory elements (HS4) were integrated upstream of the β A-TEV1 Δ sequence, due to their strong insulating qualities (Chung et al., 1997) (**Fig. 4.5A**). The inclusion of

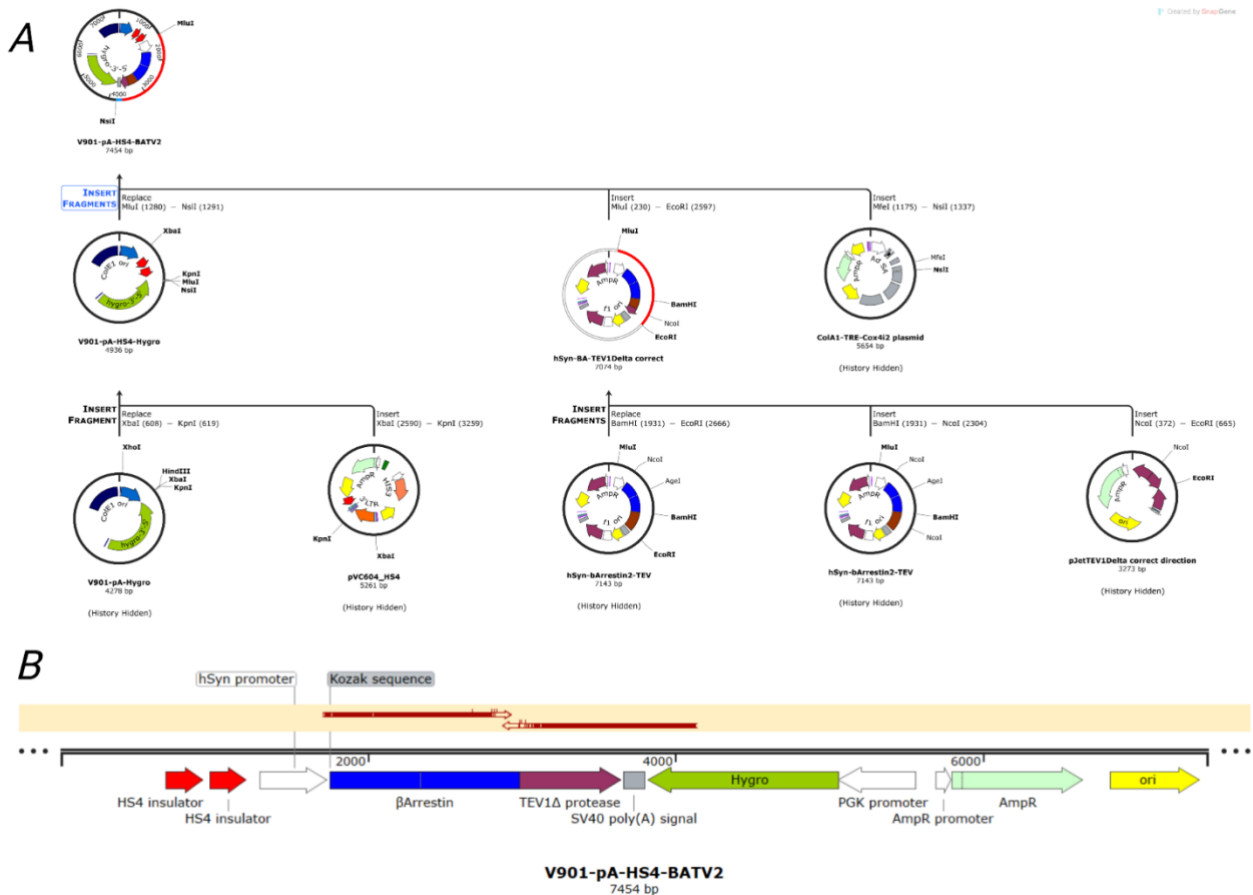


Figure 4.5 Assembly 2 of BATV2. **A)** Cloning steps required for the integration of HS4 insulating sequences and a polyadenylation sequence upstream and downstream of hSyn- β A-TEV1 Δ respectively. The Hygromycin resistance cassette driven by a PGK promoter was also integrated during this cloning **B)** Linearised V901-pA-HS4-BATV2 vector after cloning. The red arrows show the sequenced portions of the plasmid confirming the integration of the β A sequence, the TEV1 Δ protease and the polyadenylation sequence.

HS4 insulators was to limit any transcriptional interference occurring between the two components of this plasmid. Furthermore, a rabbit beta globin polyadenylation sequence was integrated downstream of β A-TEV1 Δ and upstream of the HS4 insulators.

Two HS4 elements were inserted in the V901-pA-Hygro plasmid, a generic plasmid backbone containing a Hygromycin resistance cassette and the rabbit beta globin polyadenylation sequence, to create the V901-pA-HS4-Hygro plasmid. Subsequently, the hSyn- β A-TEV1 Δ construct was integrated into V901-pA-HS4-Hygro, and a bovine Growth Hormone polyadenylation sequence derived from the ColA1-TRE-Cox4i2 plasmid was integrated between

the TEV1 Δ protease and the Hygromycin cassette. The resultant plasmid, V901-pA-HS4-BATV2, was confirmed via diagnostic digestion and Sanger sequencing. The complete cloning strategy is shown in **Fig. 4.5B**.

Concurrently, **Assembly 3** used as starting material the original AAVS1-CAG-SAMv2 plasmid from the Davies lab, which had already been used successfully to integrate a construct encoding the CAG-driven CRISPRa-SAM machinery at the AAVS1 locus of KOLFSAM, A synthetic linker

containing specific *AscI* and *NcoI* sites was incorporated into the AAVS1-CAG-SAMv2 plasmid through the *XhoI* site, to facilitate the subsequent replacement of the CAG promoter with the hSyn promoter, resulting in the AAVS1-hSyn-SAM-Delta-Neo (**Fig. 4.6A**). Secondly, the β A-TEV1 Δ and the Hygromycin cassette from the AAVS1-hSyn-SAM-Delta-Neo plasmid was combined within the V901-pA-HS4-Hygro-BATV2 plasmid, producing a complete plasmid, BATV2-Hygro, containing both the CRISPRa-SAM machinery and β A-TEV1 Δ sequence, driven by the hSyn promoter, together with a Hygromycin resistance selection cassette. Cloning validation was performed on the BATV2-Hygro plasmid through diagnostic restriction digestion and Sanger sequencing (**Fig. 4.6B**).

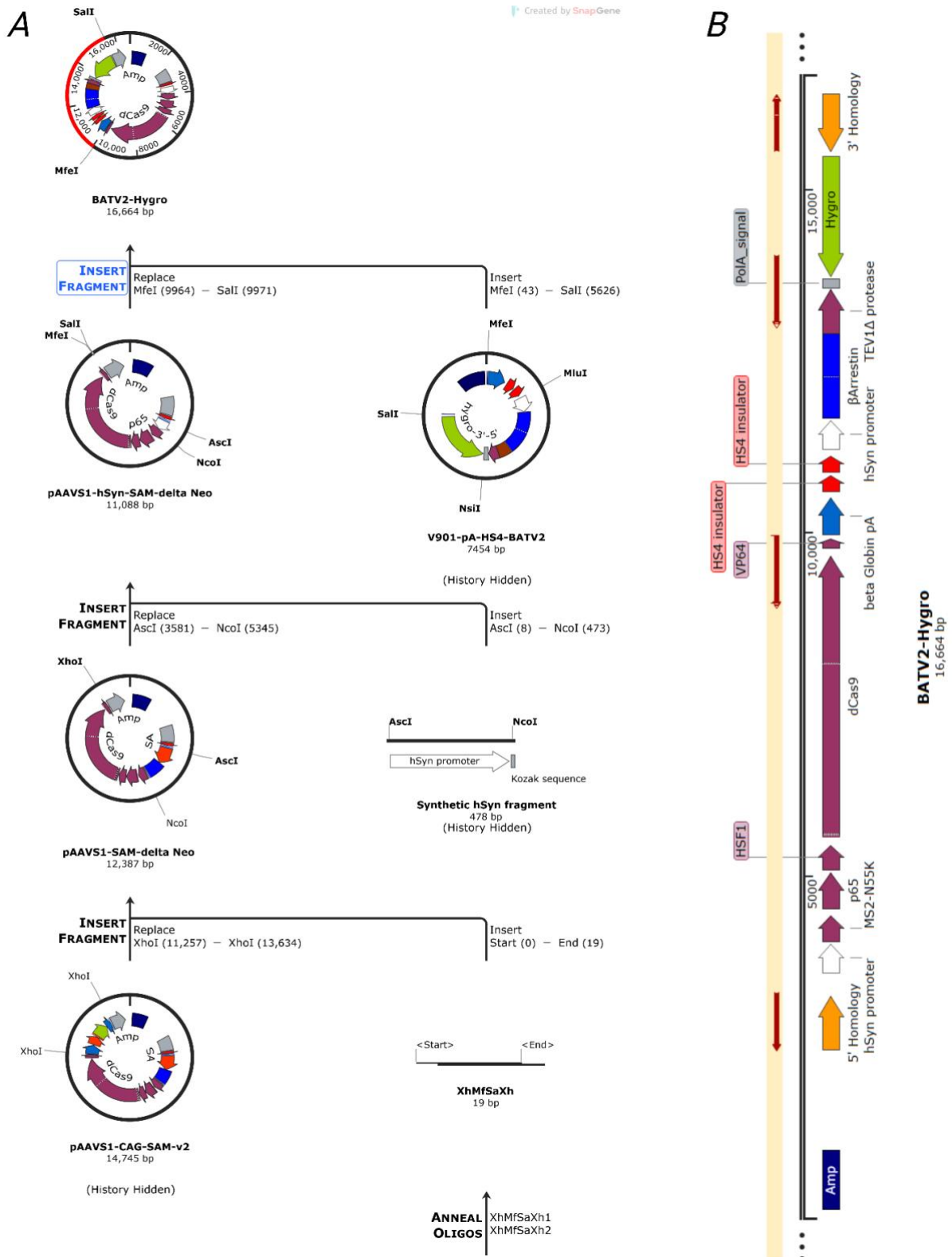


Figure 4.6 Assembly 3 of BATV2. **A)** Description of the cloning steps required for the integration of the hSyn promoter upstream of the CRISPRa machinery and final integration of the hSyn-CRISPRa sequence in the plasmid containing hSyn-βA-TEV1Δ and Hygromycin resistance cassette. **B)** Linearised BATV2 Hygro vector after cloning. The red arrows show the sequenced portions of the plasmid confirming the integration of the two Homology Arms, the CRISPRa sequence followed by the poly-A sequence as well as the βA-TEV1Δ and Hygromycin cassettes.

A

BATV2
15,961 bp

INSERT FRAGMENT
Replace
NsiI (13,719) — SalI (15,536)

Flip and Insert
XhoI (3) — NsiI (1136)

SalI

pAAVS1_hSyn_SAM_BATV2
16,653 bp
(History Hidden)

B

BATV2
15,961 bp

C

BATV2
15,961 bp

~ 91 ~

Validation of BATV2 in HEK293T cells

Following the cloning of BATV2, I validated the hSyn-driven CRISPRa-SAM machinery and β A-TEV1 Δ , integrated into this vector through functional testing in HEK293T cells.

HEK293T cells were lipofected with the BATV2 plasmid and the CD2 sgRNA plasmid or the non-targeting control. qPCR analysis was carried out 48 hours post-transfection and demonstrated that the combination of BATV2 and the CD2 sgRNA significantly upregulated CD2 expression, by an average of ~300 fold relative to the control. This outcome confirmed the functionality of the CRISPRa-SAM machinery within the BATV2 targeting vector (**Fig. 4.8A**).

Given that hSyn is a neuronal promoter, an investigation of promoter strength and its impact on CRISPRa-SAM levels was performed. The above experiment was thus repeated, with the inclusion of a CAG-driven CRISPRa-SAM construct (pAAVS1-CAG-SAMv2) in combination with the CD2 sgRNA to compare the strength of the two promoters. As illustrated in **Fig. 4.8B**, the CAG promoter induced upregulation of CD2 expression several orders of magnitude higher than that achieved by the hSyn promoter. These results align with the expectation that the hSyn promoter is naturally less active in non-neuronal cells, but nonetheless the experiments confirmed the functionality of the cloned hSyn promoter-driven CRISPRa-SAM machinery.

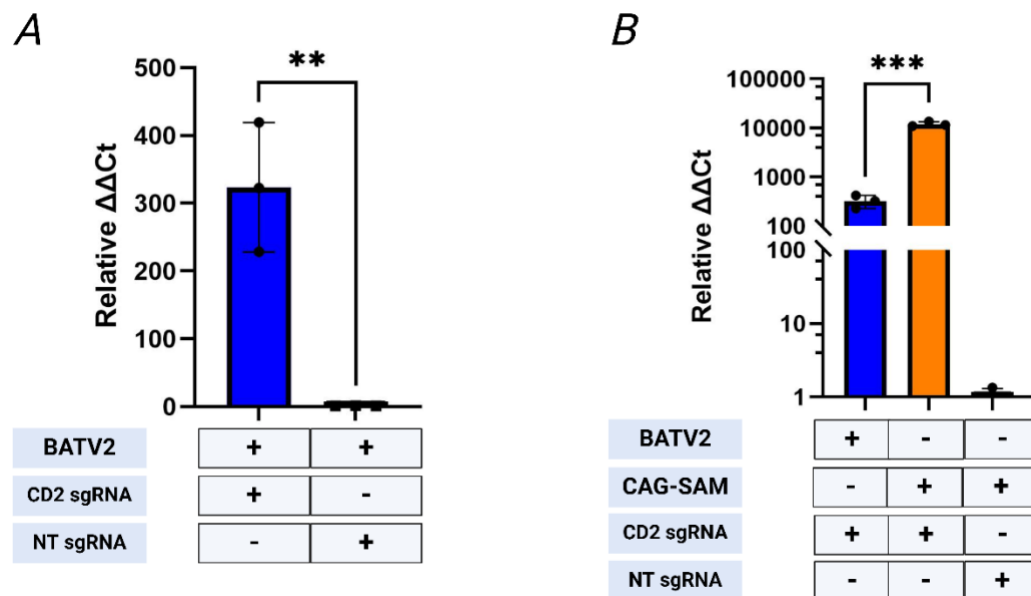


Figure 4.8 Testing of CRISPRa-SAM in BATV2. **A)** Relative expression of CD2 compared to NT control in HEK293T cells transfected with the BATV2 plasmid and the CD2 sgRNA or a non-targeting control. A t-test demonstrated that the co-transfection of the BATV2 plasmid and a sgRNA targeting CD2 resulted in a significant upregulation of CD2, compared to the co-transfection of BATV2 and a non-targeting control ($p = 0.0043$). Error bars represent standard deviation and sample size for each group = 3. **B)** Relative expression of CD2 = comparison in promoter strength between the hSyn promoter – driving CRISPRa-SAM in BATV2 – and the CAG promoter – driving CRISPRa-SAM in CAG-SAM. The one-way ANOVA a significant variation in CD2 upregulation [$F(2,6) = 237.1, p < 0.0001$]. A Tukey's multiple comparisons test demonstrated that demonstrated that the CAG-driven CRISPRa induced significantly higher CD2 expression than the hSyn-driven CRISPRa ($p < 0.0001$). Error bars represent standard deviation and sample size for each group = 3.

The functionality of the hSyn-driven β A-TEV1 Δ component within the BATV2 targeting vector was also assessed. HEK293T cells were transiently transfected with TIGHT-GFP, a plasmid generated in the lab to express a GFP reporter gene regulated by the TIGHT promoter (Takara Bio™) as I hypothesised that the GFP signal in this testing system would have been functionally equivalent to the luciferase signal from the HTLA cells. For GFP to be expressed following CMKLR1-chemerin interaction, β A-TEV1 Δ must be expressed by the hSyn promoter within the BATV2, enabling the

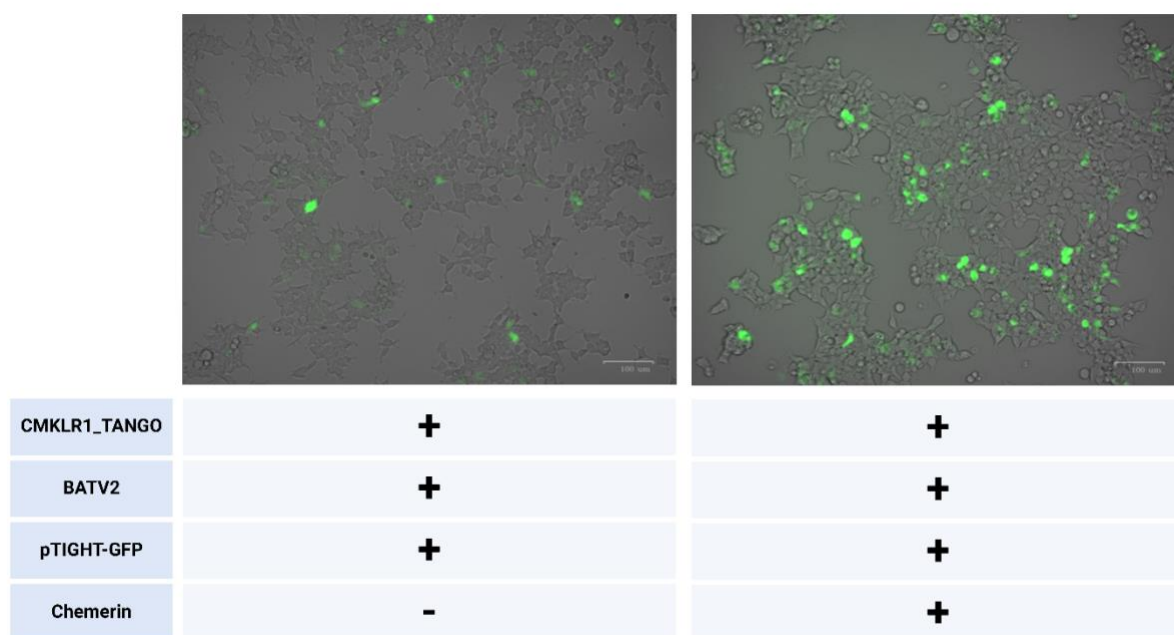


Figure 4.9 Testing of β A-TEV1 Δ in BATV2. Images of GFP expression of HEK293T cells co-transfected with CMKLR1-Tango receptor and the BATV2 plasmid containing β A-TEV1 Δ driven by the hSyn promoter. The cells were also transfected with a TIGHT-GFP plasmid to act as a reporter gene. Cells were either exposed to 1mM of chemerin or in PBS. The cells exposed to chemerin, showed an increased GFP expression

cleaving of the tTA fused to the CMKLR1 receptor. HEK293T cells were transfected with a combination of CMKLR1-Tango (Addgene, #66253), BATV2, and TIGHT-GFP, and exposed to either 1mM of chemerin or PBS. The increased GFP expression following chemerin exposure (**Fig. 4.9**) confirmed the functionality of this component, and the hSyn promoter can thus express sufficient β A-TEV1 Δ to complete the circuit post-ligand binding interaction between CMKLR1 and chemerin.

Design of the CLICK4 plasmid

With the validation of the BATV2 plasmid and the previous results suggesting that the Csy4 system is the most effective method to induce sgRNA expression through tTA activation, the second and final plasmid for the complete gene circuit was cloned. This vector (referred to as CLICK4) was designed to equip the iPSC with constitutive expression of the CMKLR1-Tango receptor, the Csy4 ribonuclease, and the TIGHT-regulated Csy4:sgRNA. To facilitate the selection

of gene targeted cells, a Neomycin resistance cassette was inserted (**Fig. 4.10A**). Ultimately, the entire sequence was flanked by homology arms to allow gene targeting into the second intronic region of the *CLYBL* safe harbour locus (**Fig. 4.10B-C**).

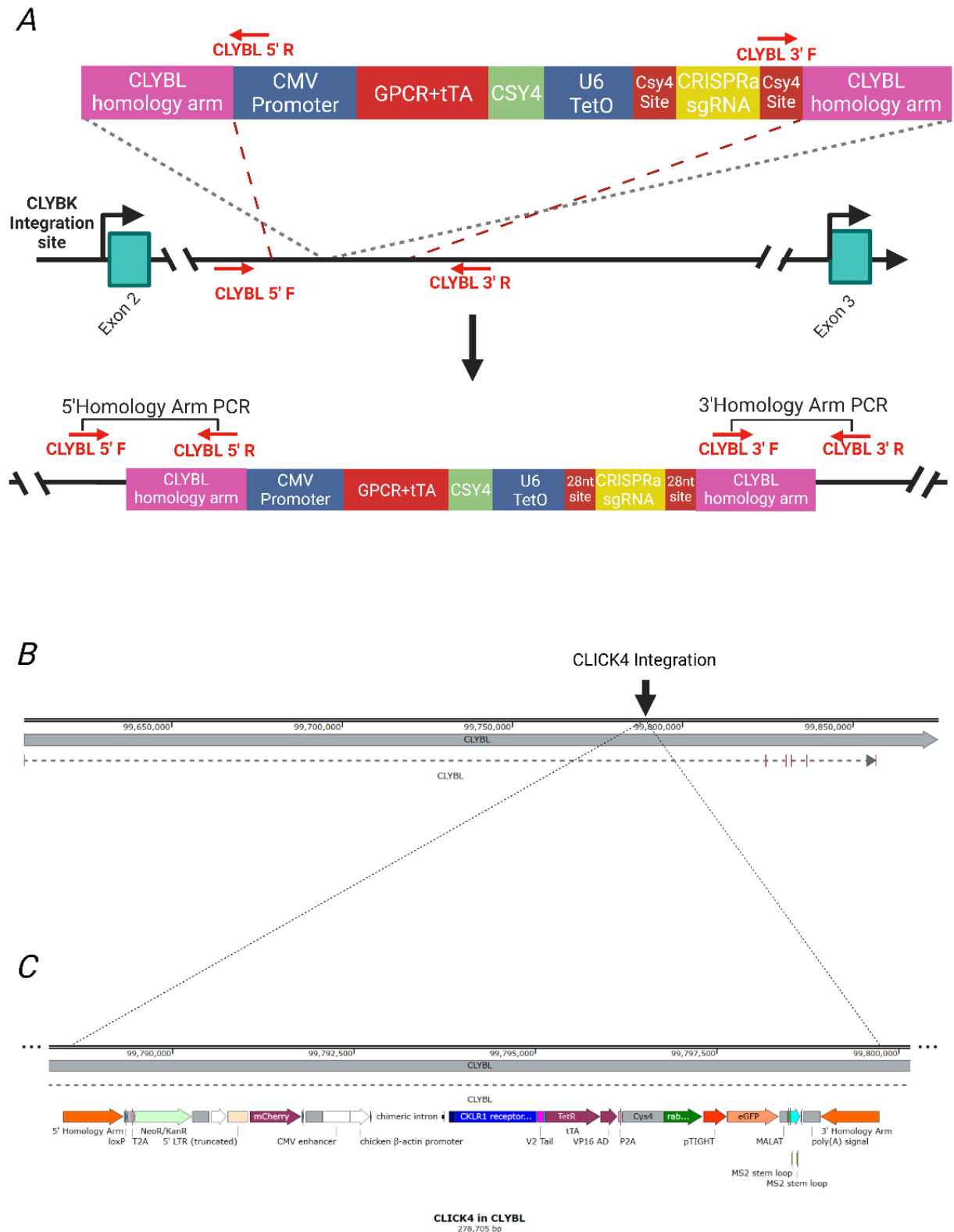


Figure 4.10 Diagram of the inserted exogenous *CLICK4* sequence into the second intron of *CLYBL*. **A)** The sequence contains two transgenes: *CMKLR1-Tango*, fused to the *Csy4* ribonuclease interspaced by a self-cleaving *T2A* sequence and a *TIGHT* promoter driving the expression of a *CRISPRa-SAM* sgRNA flanked by two *Csy4*-nano cleaving sites. **B)** Map of the *CLYBL* sequence showing the integration site into the second intronic region of the locus. **C)** A focus on the integrated *CLICK4* sequence within the *CLICK4* plasmid.

Due to the complex nature of the cloning, the generation of the plasmid involved numerous cloning steps, which were executed in several independent assemblies sequentially.

Assembly 1 involved cloning the CLYBL homology arms and the MS2 sgRNA structure. The previously developed U6 TetO plasmid containing the majority of these elements, was combined with pUCM-CLYBL-hNIL vector (Addgene, #105841), which contained the homology arms required for targeted integration into the *CLYBL* safe harbour locus (Cerbini et al., 2015a). The U6 TetO sequence was initially incorporated into a cloning vector v901-2 v2, and equipped with a *BsmBI* linker before being cloned into pUCM-CLYBL-hNIL (**Fig. 4.11A**). The resulting plasmid, CLYBL TETO MS2, underwent validation through restriction digestion and Sanger Sequencing (**Fig. 4.11B**).

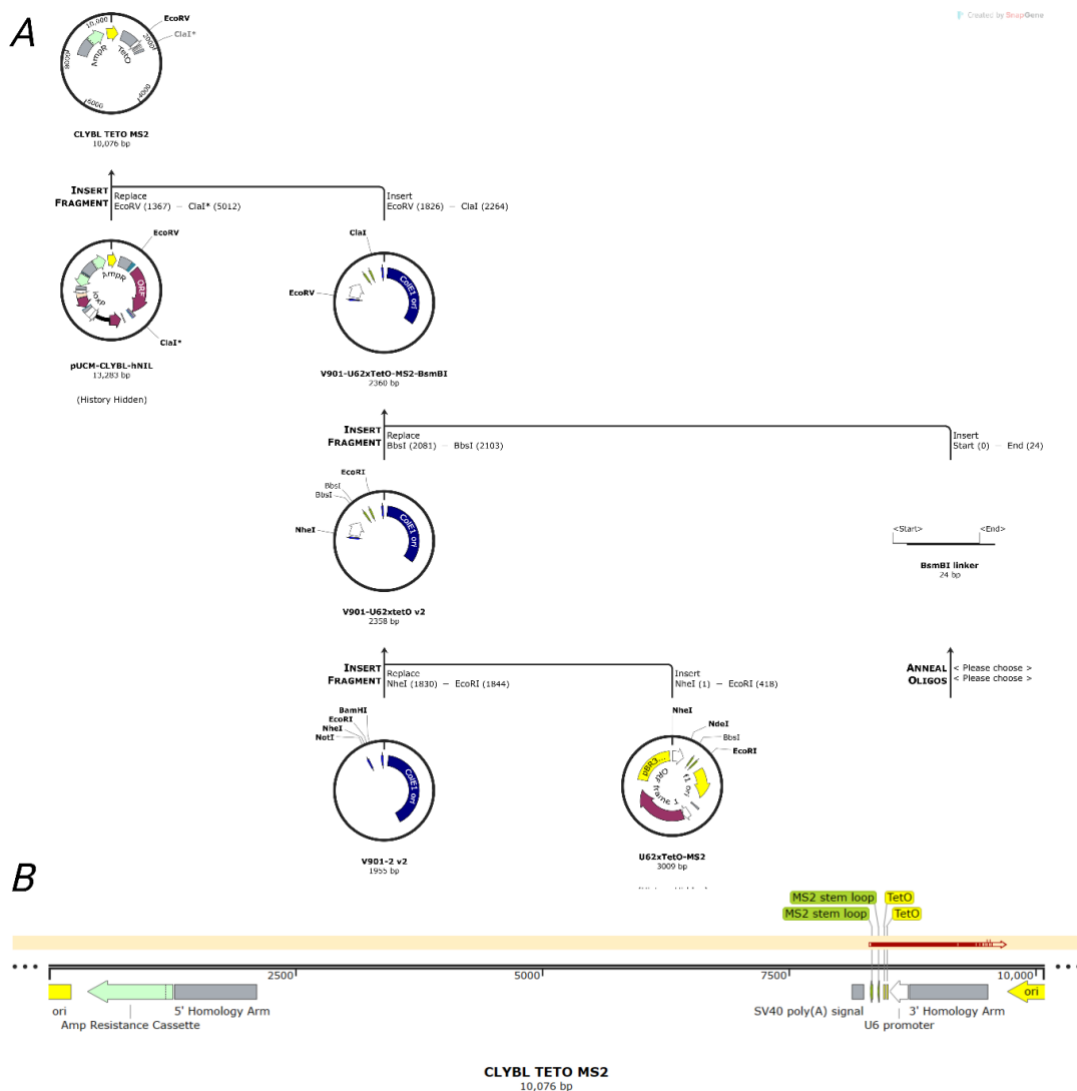


Figure 4.11 Assembly 1 of CLICK4. **A)** Cloning steps required for the integration of the CLYBL homology arms followed by the U6 TETO CRISPRa sgRNA sequence. **B)** Linearised CLYBL TETO MS2 vector after cloning. The red arrow shows the sequenced portion of the plasmid confirming the integration of the 3' Homology Arm, followed by the U6 promoter driving the expression of the CRISPRa sgRNA sequence.

Next, **Assembly 2** involved the integration of the CMKLR1-Tango receptor expression cassette. Initially, a generic cloning vector, V924, was modified to include a synthetic sequence containing an *XbaI* and *NotI* sites. This adjustment facilitated the subcloning of the CMKLR1-Tango sequence from the original plasmid (Addgene, #66253), into V924. The final step involved inserting this sequence into the CLYBL TETO MS2 plasmid using *Sall* in the vector site and *XhoI* in the insert site, taking advantage of their isoschizomeric nature (**Fig. 4.12A**). The resulting plasmid, CLYBL TETO MS2 CMKLR1, was validated through restriction digest and Sanger sequencing (**Fig. 4.12B**).

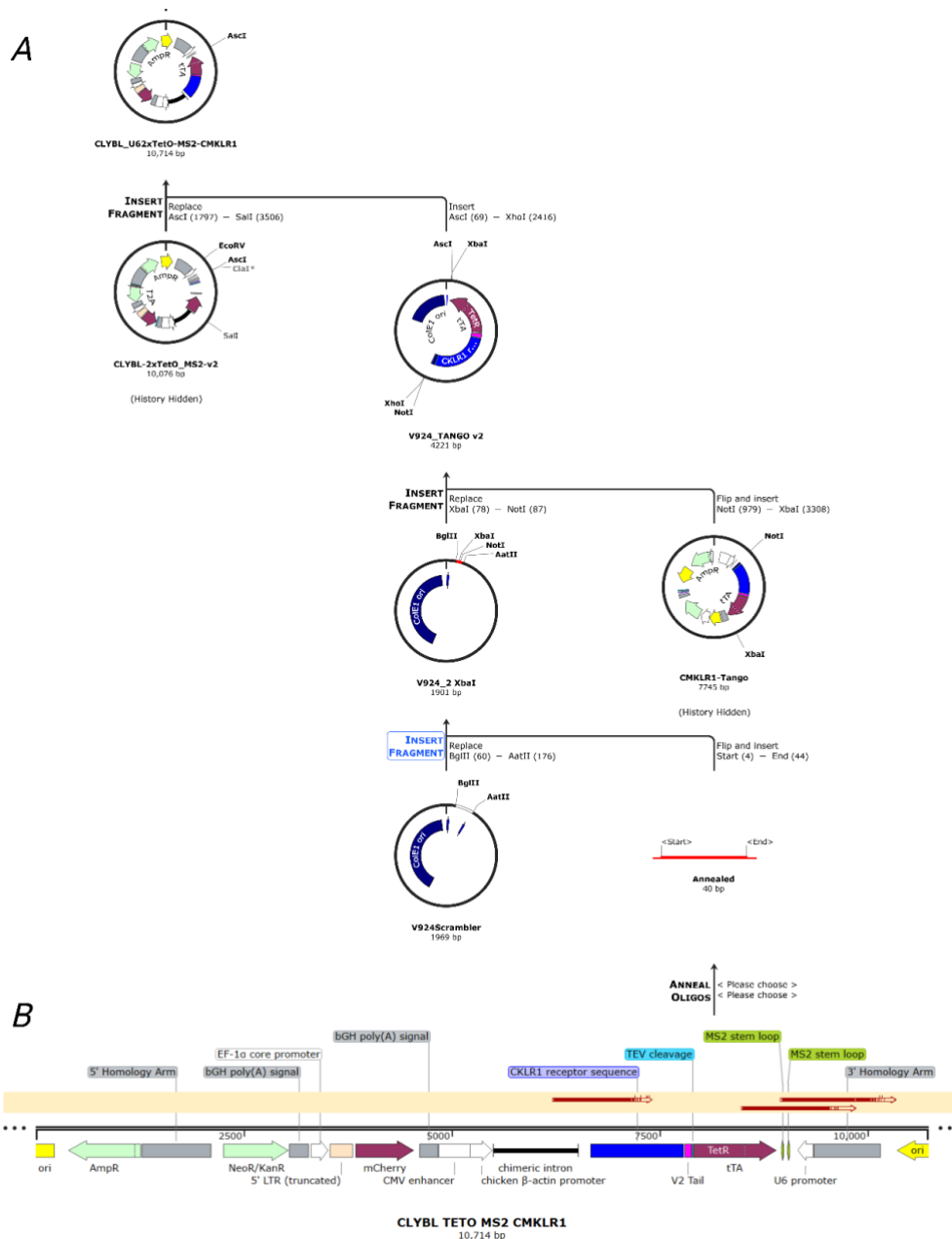


Figure 4.12 Assembly 2 of CLICK4. **A)** Cloning steps required for the integration of the CMKLR1-Tango sequence into the CLYBL plasmid. **B)** Linearised CLYBL TETO MS2 CMKLR1 vector after cloning. The red arrows show the sequenced portions of the plasmid confirming the presence of the priorly integrated U6-sgRNA sequence, the integration of the CMKLR1-Tango receptor and the 3' homology arm.

Created by SnapGene



~ 98 ~

arrows show the sequenced portion of the plasmid confirming the integration the *pol-A* sequence, the *TIGHT* promoter driving the expression of the GFP reporter sequence as well as the sgRNA scaffold.

In the final **Assembly 4**, the new Csy4 regulatory system with the Csy4 ribonuclease was integrated into the targeting vector. Firstly, the Csy4 ribonuclease was cleaved from the TR-Csy4 plasmid (Addgene, #80601) and inserted downstream of the CMKLR1-Tango sequence, interspaced by a *Thosea asigna* virus 2A (T2A) “self-cleaving” peptide sequence. This sequence has been demonstrated to allow bicistronic expression after transcription, with a tTA maintaining its efficiency while still being fused to the T2A remnant (Moriyama et al., 2013). This arrangement allows the Csy4 to be expressed together with the CMKLR1-Tango receptor. Secondly, a synthetic fragment containing a Csy4-nano sequence was integrated. Finally, the *CD2* sgRNA was

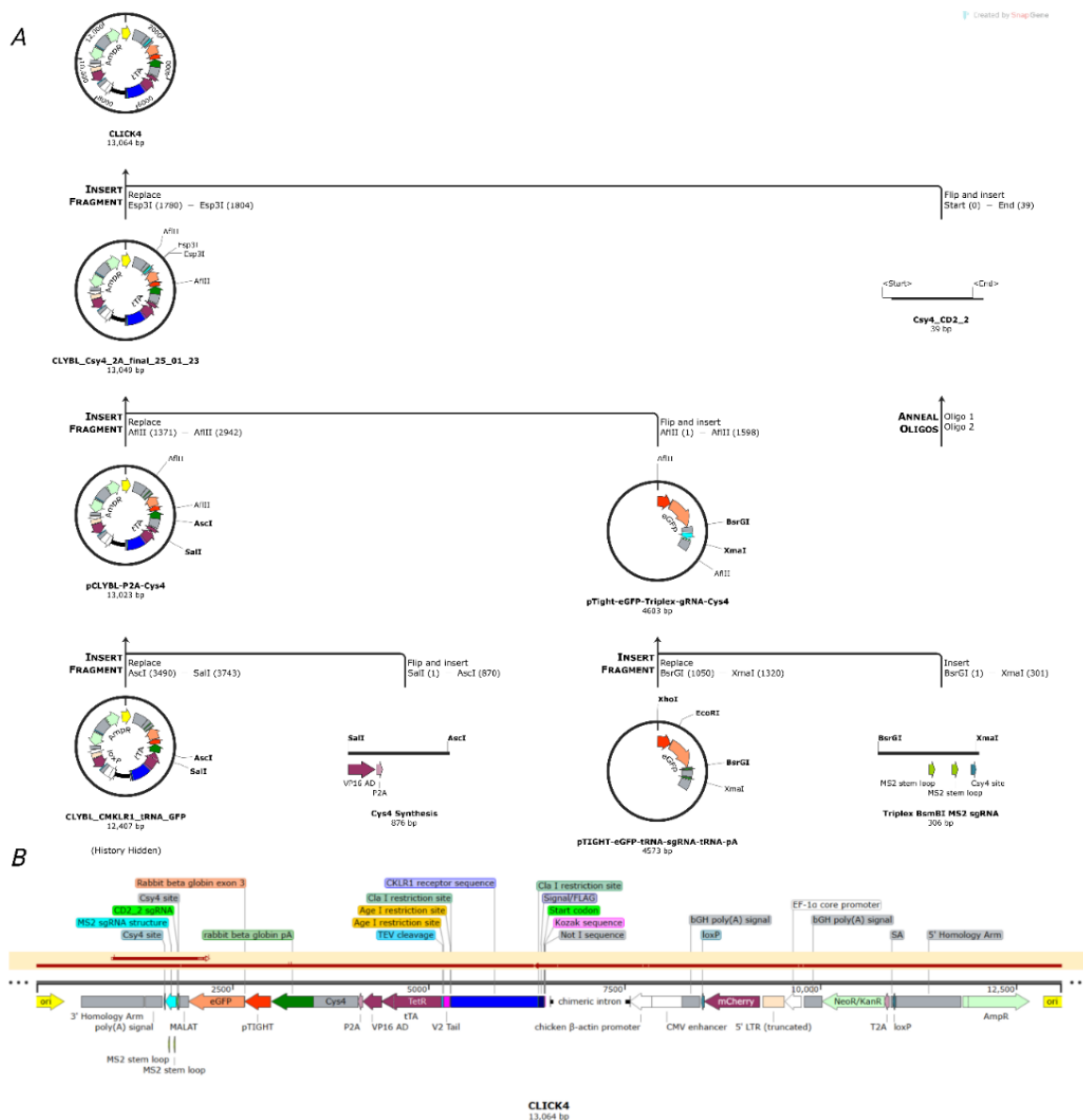


Figure 4.14 Assembly 4 of CLICK4. **A)** Cloning steps required for the integration of the Csy4 ribonuclease and the Csy4:sgRNA sequence regulated by the *TIGHT* promoter. **B)** Whole-plasmid sequencing – shown by the red arrow – demonstrating that the entire plasmid was cloned as intended.

integrated with the second Csy4 nano site, completing the CLICK4 plasmid (**Fig. 4.14**). This plasmid underwent validation through double digest as well as a comprehensive full plasmid sequencing analysis via Oxford Nanopore Technologies Long-Read Sequencing (**Fig. 4.14/15**).

Created by SnapGene



Figure 4.15 Final CLICK4 map. The map shows the integration of the CMKLR1-Tango system driven followed by the Csy4 ribonuclease interspersed by a P2A self-cleaving mechanism and driven by a CMV promoter. The GFP and Csy4:sgRNA sequences are driven by a TIGHT promoter and the entire transgene is flanked by CLYBL homology arms.

Validation of the CLICK4 plasmid

Following the successful cloning of CLICK4, I validated its two main components, both individually and in the context of a complete circuit.

Firstly, the efficiency of the CMKLR1-Tango system, which was previously tested using individual expression plasmids, had to be re-evaluated after its integration into the CLICK4 targeting vector.

HTLA cells, expressing β A-TEV and a TRE-regulated luciferase reporter, were transfected with either the CLICK4 plasmid or the original CMKLR1-Tango plasmid as a positive control. Cells were also transfected with FPRL1 Tango or an empty vector, acting as negative controls. 24 hours after transfection, the cells were exposed to 1mM chemerin or PBS, and the luciferase assay was conducted 24 hours after exposure. As shown in **Fig. 4.16A**, the relative luciferase expression of CLICK4 cells exposed to chemerin was significantly higher than that of FPRL1 and the negative control, while no significant difference could be noted between the CLICK4 and CMKLR1-Tango. These results suggested that the CMKLR1 sequence in the context of the CLICK4 targeting vector retained its function.

Having shown the functionality of the CLICK4 and BATV2 plasmids' individual components, I combined the machinery into the full circuit, testing initially in HEK293T cells. Since the primary Csy4:sgRNA transcript was embedded downstream of a GFP, this fluorescence cassette provides a simple readout of the induction of the transcript into which the sgRNAs are embedded. HEK293T cells were transfected with the BATV2 and the CLICK4 plasmids and exposed to either 1mM chemerin or PBS, 24 hours after transfection. I hypothesised that the combination of BATV2 and CLICK4 should have provided all the components to complete a circuit capable of transducing the presence of chemerin to a GFP reporter signal. To further control the experiment HEK293T cells were also transfected with the individual components as single expression plasmids (CMKLR1-Tango (Addgene, #66253), the β A-TEV plasmid (pcDNA3.1(+)-CMV-bArrestin2-TEV, Addgene #107245), and the TIGHT-GFP plasmid).

As shown in **Fig. 4.16B**, there was no significant difference in GFP expression between the cells exposed to chemerin and those exposed to the PBS in the BATV2+CLICK4 transfected cells nor in CMKLR1-Tango + β A-TEV + TIGHT-GFP transfected cells. The main observation from this experiment was an overall high level of GFP expression from the transfected cells irrespective of the induced state, that is to say, that the GFP expression from either the TIGHT-GFP plasmid or the CLICK4 plasmid was leaky. The leakiness masked any potential effect of chemerin and activation of the CMKLR1-Tango receptor, preventing the functional validation of the full circuit.

Despite the impact of the leakiness of the TIGHT promoter, when transiently transfected, on a functional evaluation of the circuit, I hypothesized that the capability of Cys4 to serve as a secondary control system might still provide some level of regulation over the Csy4:sgRNA system.

To explore this in more detail, HEK293T cells were transfected with the BATV2 and CLICK4 plasmids. Additionally, the cells were co-transfected with a combination of: tTA only, Csy4 only, or both tTA and Csy4. The results (**Fig. 4.16C**) revealed that the empty circuit (BATV2 and CLICK4 only) exhibited minor *CD2* upregulation, consistent with the TIGHT leakage observed earlier. Interestingly, the addition of tTA and Csy4, either independently or combined, resulted in a similar yet weaker response compared to when each component was investigated individually. Specifically, the addition of Csy4 led to nearly double the *CD2* upregulation observed with the empty circuit, and this effect was slightly enhanced with the addition of tTA. Conversely, tTA alone produced the lowest level of upregulation among the three conditions.

These findings confirm that Csy4 functions as a relatively effective regulator, as its absence prevents the processing of the TIGHT-driven transcript, and ensures that no functional sgRNA is available to interact with the CRISPRa machinery. However, the overall weak signal, combined with the leakiness of the TIGHT promoter, indicated that the CLICK4 plasmid was not yet suitable for genomic integration.

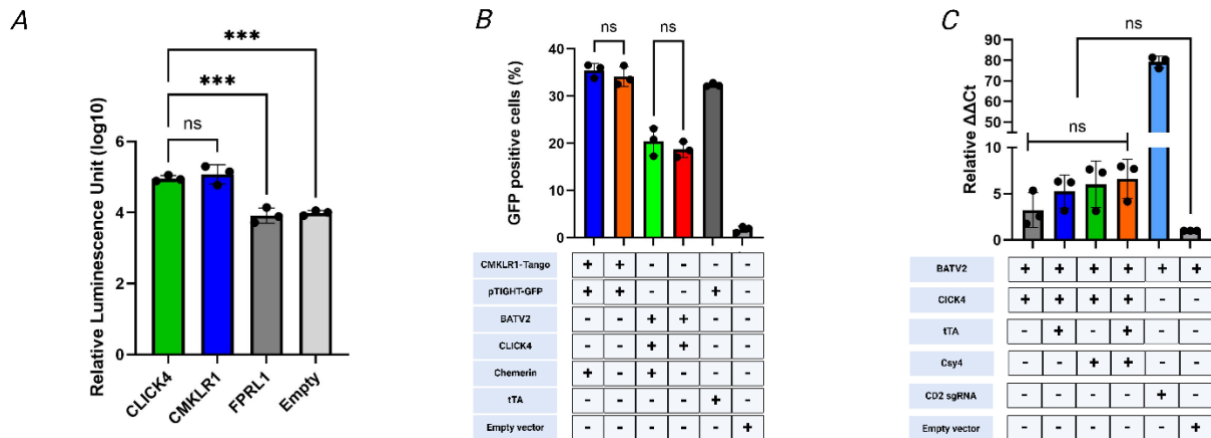


Figure 4.16 Investigation of CMKLR1 Tango and Csy4:sgRNA in CLICK4. **A)** Relative luciferase expression of HTLA cells transfected with CLICK4, CMKLR1, FPRL1 plasmid or an empty vector and exposed to Chemerin. Following normalisation by log10, one-way ANOVA demonstrated that chemerin exposure significantly influenced luciferase expression [$F(3,8) = 35.09$, $p < 0.0001$]. A Tukey's multiple comparisons test demonstrated that CLICK4-transfected cells produced significantly more relative luminescent units (RLU) than the FPLR1 transfected cells ($p = 0.0005$) and the control transfected with an empty vector ($p = 0.0008$). The RLU of CMKLR1-transfected cells was not significantly different from the positive control – where the Tango system is bypassed via addition of tTA ($p = 0.8585$). Error bars represent standard deviation and sample number = 3. **B)** Percentage of GFP expressing HEK293T cells following transfection of the original TANGO components (CMKLR1-CMV-BATEV and the TIGHT-GFP) or the full circuit (BATV2 + CLICK4) after exposure of Chemerin or PBS. The samples were also compared to a positive control, TIGHT-GFP combined to tTA and a untransfected HEK293T cells. A two-way ANOVA revealed no significant the exposure to Chemerin did not result in a significant variation in the number of GFP expressing cells [$F(1,13) = 2.414$, $p = 0.1443$]. The different plasmids on the onther hand, resulted in a significant variation in GFP expressing cells [$F(3,13) = 286.1$, $p < 0.0001$], due to the positive control. Error bars represent standard deviation and sample number = 3. **C)** Relative CD2 expression of HEL293T cells transfected with BATV2 + CLICK4 combined with: tTA only, Csy4 only, tTA+Csy4. A one-way ANOVA demonstrated that a significant variation of CD2 upregulation between groups [$F(5,12) = 672.0$, $p < 0.0001$]. A Tukey's multiple comparisons test indicated that this was however, fully driven by the positive control. No significant difference in CD2 expression could be seen between the CLICK4 transfected cells and the cells transfected with CLICK4 and: tTA ($p = 0.8264$), Csy4 ($p = 0.5713$) and tTA+Csy4 ($p = 0.3852$). Error bars represent standard deviation and sample number for each group = 3.

Closing Remarks

In this chapter, I detailed the cloning process aimed at generating the BATV2 and CLICK4 plasmids, housing the entire circuit in the context of targeting vectors for integration at the *AAVS1* and *CLYBL* loci, respectively. The BATV2 plasmid comprised the CRISPRa-SAM machinery and the β A-TEV1 Δ , driven by the hSyn promoter. The CLICK4 plasmid, comprised the CMV-driven CMKLR1-Tango and TIGHT-driven Csy4:sgRNA transcript.

Transient transfection of BATV2 alongside the *CD2* sgRNA plasmid illustrated that the hSyn promoter effectively drove CRISPRa-SAM expression, leading to significant upregulation of *CD2*. However, when compared to a CRISPRa-SAM machinery driven by a CAG promoter, the hSyn-CRISPRa-SAM machinery was notably less potent. This discrepancy aligns with the anticipated promoter strength in a non-neuronal cell line. The weakness of the hSyn promoter in this non-neuronal cell type might contribute to a weaker overall output signal when combined in the complete circuit.

Similarly, the expression of β A-TEV1 Δ via the hSyn promoter was confirmed through transient transfection of BATV2, CMKLR1-Tango, and TIGHT-GFP plasmids, successfully reconstituting the original Tango system using BATV2 to express β A-TEV1 Δ . However, while more GFP expression was observed following chemerin exposure, the subsequent results suggested that a leakiness of the TIGHT promoter could mask any potential inductions resulting from the receptor activation. Potentially a more rigorous assessment of whether the hSyn promoter effectively drives sufficient β A-TEV1 Δ expression is warranted.

The CLICK4 plasmid, containing the CMKLR1- tTango and TIGHT promoter-driven Csy4:sgRNA system, was also cloned successfully. While the CMKLR1-Tango system integrated into CLICK4 behaved similarly to the original CMKLR1 plasmid, the TIGHT promoter was again found to be considerably leaky. However, the addition of Csy4 increased *CD2* upregulation, suggesting that Cys4 provides a degree of regulation by releasing the mature sgRNA following cleavage of the Csy4:sgRNA precursor transcript. Nevertheless, *CD2* upregulation following Csy4 addition, or the combination of Csy4 and tTA, was found to be weaker than anticipated, making the complete circuit less effective.

As mentioned earlier, signal decay can pose a significant challenge in gene circuit development (Brophy and Voigt, 2014). Variations in individual component factors such as promoter strength and protein decay can impede the signal if the output strength of each component fails to activate the subsequent one. The testing of the circuit in a non-neuronal cell line, yet using neuronal specific promoters, might provide an explanation here.

The CRISPRa-SAM system was confirmed to be effectively expressed by the hSyn promoter but was considerably less able to induce gene expression from exogenous loci, when compared with the CAG promoter driven equivalent construct. Thus, when integrated within the complete circuit, it is possible that this low expression level could result in a signal too weak to be detected. Furthermore, it is possible that the hSyn promoter driven β A-TEV1 Δ expression was below a threshold required for robust Tango receptor functionality. When coupled with the leaky expression of Csy4:sgRNA, these concerns on the expression level may explain the very weak *CD2* upregulation observed from the complete circuit. The weakness of the hSyn promoter in non-neuronal cell lines, makes it challenging to test individual components as well as the complete circuits in cell lines such as HEK293T.

Two possible solutions would be to transfect the circuit into the iPSC-derived neurons or to use immortalised neuronal cell lines such as NB69, a neuroblastoma cell line. Transfecting BATV2 into these cell lines would increase the strength of the hSyn promoter and thus the signal produced by either CRISPRa-SAM or β A-TEV1 Δ , possibly overcoming the signal decay noticed in these results.

However, the introduction of the hSyn promoter was specifically intended for expression within the context of the *AAVS1* safe harbour locus in iPSC-derived neurons. Consequently, the behaviour of the CRISPRa-SAM system driven by a neuronal promoter in iPSC-derived neurons needs to be explored by integrating the construct into the *AAVS1* as reported in the following Chapter.

Chapter 5 Validation of hSYN-driven CRISPRa-SAM in iPSC-derived neurons

Introduction

The aim of the gene circuit was to allow maintained and robust expression of the system in iPSC-derived neurons. Following the generation and validation of the circuit components within BATV2, I integrated this sequence into the genome of KOLF2-C1 iPS cells via CRISPR/Cas9 mediated homologous recombination.

Insertion of exogenous DNA into the host genome requires integration into locations that avoid significant disruptions to the genomic landscape, while ensuring consistent and predictable gene expression of the transgene. The *AAVS1* locus, situated on chromosome 19, encodes for protein phosphatase 1 regulatory subunit 12C (*PPP1R12C*) (Kotin et al., 1992). Given that this locus is considered largely redundant, the integration of exogenous genes into its intronic region has not only proven non-disruptive to other genes but has also resulted in robust gene expression in defined cell states (Smith et al., 2007). Over the past decade, *AAVS1* has emerged as a central integration site in human cell lines (Kotin et al., 1992, Pavani and Amendola, 2021).

Inducible transgenes have been successfully integrated into the *AAVS1* locus of human iPS cells. The integration was realized through various methods, such as Zinc-finger nucleases, TALENs, and CRISPR. However, in most cases, the integration and expression of the exogenous construct was not tracked across differentiation, usually focusing on only a single cell state (González et al., 2014, Hockemeyer and Jaenisch, 2016, Hayashi et al., 2020a, Nandy et al., 2023).

When investigating expression of an integrated construct across cell differentiation, significant variation in the expression of *AAVS1*-integrated transgenes has been observed (Ordovás et al., 2015, Bhagwan et al., 2020, Karbassi et al., 2024). Notably, considering the reported silencing effects of the *AAVS1* locus, an investigation was undertaken to explore the use of myeloid-specific promoters to drive the expression of transgenes in iPSC-derived myeloid cells (Klatt et al., 2020). The study revealed that myeloid promoters were heavily impacted by methylation when the cells were in their pluripotent state, leading to a complete shutdown of gene expression. However, expression recovered in the myeloid state, where promoter methylation decreased. This study demonstrated that the *AAVS1* locus is susceptible to methylation effects, which, in turn, affect gene expression during cell differentiation and highlight the potential of tissue-specific promoters for overcoming gene silencing effects.

To date, no studies have explored the integration of neuron-specific promoters to overcome silencing *AAVS1*-integrated transgenes in iPSC-derived neurons.

In the previous chapters, the hSyn promoter was observed to drive the expression of CRISPRa-SAM in HEK293T cells, albeit with weaker efficiency compared to the synthetic CAG promoter. However, it would be expected that hSyn would drive robust expression during neuronal differentiation and circumvent potential silencing phenomena of *AAVS1*-integrated CRISPRa machinery, as similarly demonstrated in the aforementioned study (Klatt et al., 2020).

Targeting of SAM-BATV2 into *AAVS1* Safe Harbor Locus

BATV2 was designed to be integrated via homology-directed recombination into the first intronic region of the *AAVS1* safe harbour locus of KOLF2-C1 iPS cells (**Fig. 5.1A**). First, sgRNAs targeting the integration sites of both *AAVS1* and *CLYBL* were tested for efficiency of target site cleavage. The *AAVS1* targeting sgRNA had been previously validated by Castaño et al. (2017), while the *CLYBL* targeting sgRNA was designed using the CRISPOR algorithm as detailed in the Methods section. Pluripotent KOLF2-C1 iPS cells were transfected with a combination of Cas9 ribonucleoprotein and *AAVS1* or *CLYBL* sgRNAs. 48 hours after transfection, DNA was extracted, and the region of interest was amplified by PCR with appropriate primers across the *AAVS1* or *CLYBL* target sites. PCR products were sequenced and the percentage of indels generated at the target site was quantified using the Synthego-ICE analysis tool. Results were compared to the *hFLNC*-targeting sgRNA commonly used in the Davies lab as a positive control. As depicted in **Fig. 5.1B**, the *CLYBL* sgRNA induced ~30% indels, similar to *hFLNC*, suggesting it to be a reliable guide. On the other hand, the *AAVS1* sgRNA yielded only ~8% indels, indicating a lower efficiency. However, this guide had previously succeeded in integrating the CAG-driven CRISPRa-SAM into the *AAVS1* locus of KOLF2-C1, giving rise to the KOLFSAM cell line. To minimize variability between KOLFSAM and the cell line generated in this chapter, the *AAVS1* sgRNA was chosen for integration of the BATV2 sequence.

KOLF2-C1 cells were transfected with the ribonucleoprotein (Cas9 complexed with the *AAVS1* sgRNA), with the addition of the BATV2 plasmid as an HDR template (**Fig. 5.1A**). Cells were then exposed to puromycin for selection. Homologously recombined clones that survived selection were identified by PCR amplification over the 5' and 3' homology arms of the BATV2 targeting vector. Each primer pair comprised one primer outside the BATV2 homology sequence (p963 or p965) in the endogenous *AAVS1* gene locus, while the other primer was positioned at the 5' or 3' extremities of the sequence to be inserted (H5R2 for the 5' PCR or PGKR1 for the 3' PCR) (**Fig. 5.1A**). Consequently, amplification of either fragment relied on the presence of the BATV2

sequence at the expected AAVS1 location. Wild-type KOLF2-C1 cells were used as a negative control as no amplification was expected to occur without the presence of BATV2. The KOLFSAM cell line, served as a positive control for the 5' PCR, given its 5' sequence matched that of BATV2. Preliminary genotyping confirmed the integration of the BATV2 sequence into the cell pool (Fig. 5.1C).

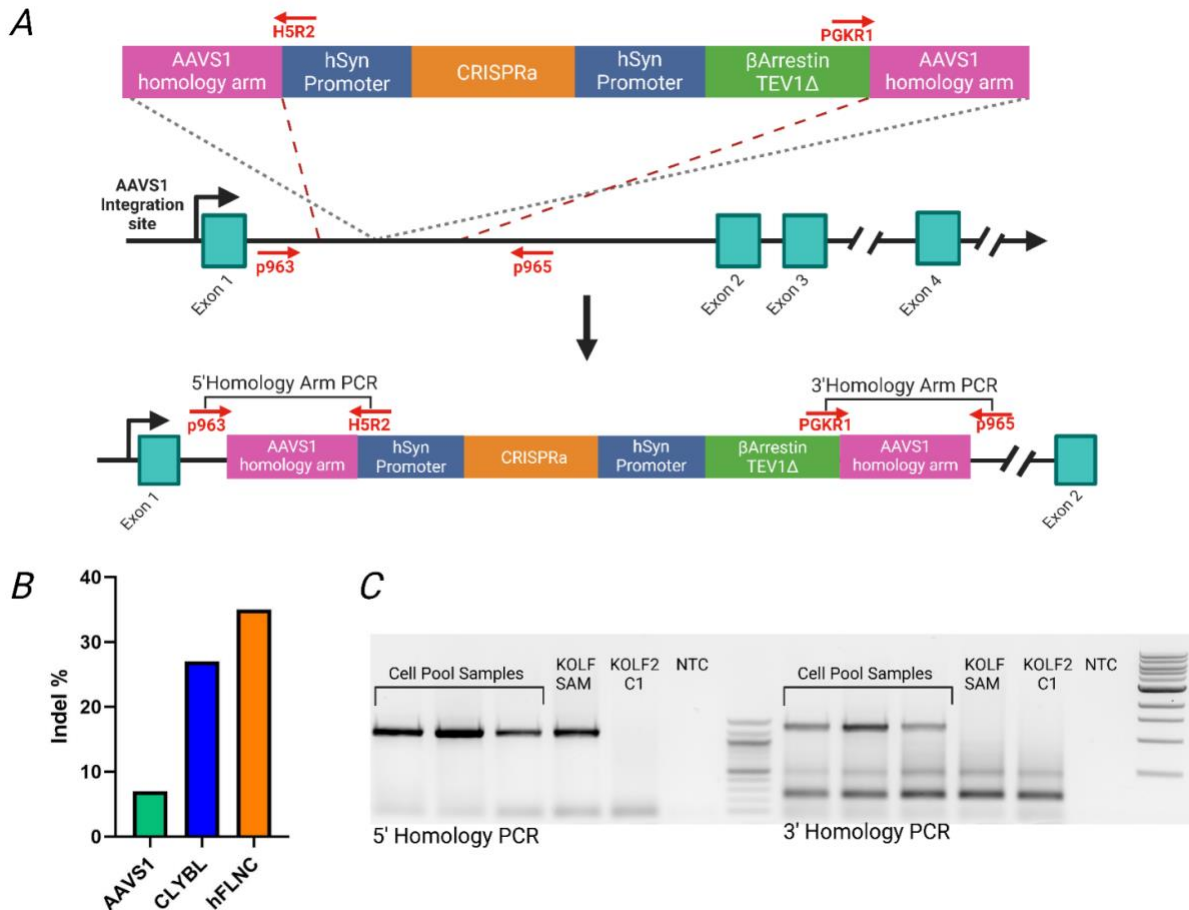


Figure 5.1 Integration of BATV2 in KOLF2-C1 iPSC. **A)** Diagram of the BATV2 insertion into the first intron of AAVS1. The PCR sequences for the 5' and 3' Homology Arms are also shown. For each primer set, one primer (p963-p965) is located in the endogenous AAVS1 gene sequence, while the other primer (H5R2-PGKR1) is located in the BATV2 sequence. **B)** Cutting efficiency of the sgRNA used to target AAVS1, CLYBL and human FLNC. KO score describes the percentage of INDELS generated by Cas9 at the target site **C)** Gel electrophoresis of PCR amplification of 5' and 3' homology arm PCRs for the cell pools containing the BATV2 sequence. Given the location of the primers – on the endogenous sequence and the BATV2 sequence – amplification is only possible if BATV2 has been correctly integrated. KOLFSAM, characterised by a similar sequence at the 5' was used as positive control.

Following positive pool identification, cells were seeded at low density and cultured. Colonies from a 96-well plate were isolated and clonally expanded for further genotyping. 21 colonies displayed the expected amplification of the 5' homology arm product and were chosen for further genotyping (Fig. 5.2A). Successful 3' homology arm amplification was observed in 19 of these 21 colonies, confirming the successful integration of both sides of the BATV2 sequence into the AAVS1 safe harbour locus (Fig. 5.2B). Subsequently, a PCR aimed at determining the zygosity of

the cell lines was conducted to select a heterozygous and a homozygous cell line. This strategy involved selecting primers p963 and p965, targeting the *AAVS1* site. In a wild-type cell line, the proximity of these primers allows for fragment amplification. However, if BATV2 is correctly integrated, the increased distance between p963 and p965 renders amplification unviable. Therefore, successful amplification of the *AAVS1* PCR fragment in a cell line confirmed to possess the BATV2 sequence suggests that one *AAVS1* copy is intact, indicating heterozygosity. Conversely, a lack of amplification in a positive cell line suggests that both *AAVS1* copies were affected by Cas9 editing, resulting in the integration of the BATV2 sequence in both copies and indicating homozygosity. It is worth noting that Cas9-double strand breaks without integration of BATV2 can also result in a loss of one or both primer sites, resulting in a lack of amplification without the BATV2 integration. 2 homozygous and 2 heterozygous cell lines were identified using this approach (**Fig. 5.2C**). One of each cell line was taken forward for further analysis in this study, 3F (KOLFBAT Hom) and 9G (KOLBAT Het).

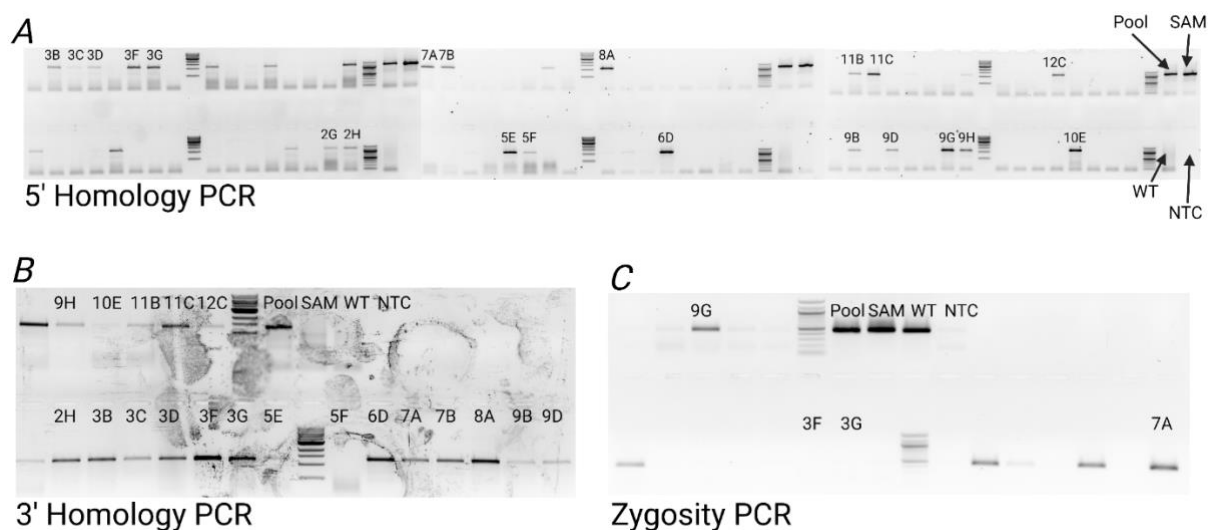


Figure 5.2 Gel electrophoresis of KOLFBAT. **A)** Gel electrophoresis of 5' homology arm amplification in the 96 colonies. Presence of an amplicon indicates integration of the 5' end of the BATV2 construct. **B)** Gel electrophoresis of 3' homology arm amplification of the colonies which had already amplified the 5' homology arm. **C)** Gel electrophoresis of the *AAVS1* of the colonies that already showed 5' and 3' amplification. Presence of an amplification product suggests heterozygosity, lack of amplification suggests possible homozygosity.

Quality Control of KOLFBAT Hom and KOLFBAT Het

Genetic manipulation of pluripotent stem cells can exacerbate the genetic instability inherent in human iPS cells, potentially leading to the generation of unviable cells that fail to differentiate correctly or at all (Lund et al., 2012). Therefore, quality control tests were conducted to assess the genomic stability and pluripotency of the KOLFBAT Hom and KOLFBAT Het cell lines.

Pluripotency was assessed through the expression of key markers *POUF51* (*OCT4*), *NANOG*, and *SOX2*, analysed via staining and FACS analysis (**Fig. 5.3A**). The blue cluster, representing stained cells, exhibited high levels of pluripotency factors with no overlap with the controls stained with non-specific antibodies (in red), indicating that both cell lines retained their pluripotent characteristics. Furthermore, the expression of pluripotency markers in the KOLFSAM cell line, as previously conducted by Dr. Rath, also confirmed the pluripotency of this line, used in later studies (**Fig. 5.3B**).

Genomic stability was evaluated through chromosome counting, performed by Daniela Moralli at the Chromosome Dynamics Core (WCHG), using DAPI-stained metaphases. Chromosome counting (**Fig. 5.3C**) enables the detection of possible chromosomal abnormalities such as aneuploidy, as well as insertions or deletions at the scale of a few megabases (Lund et al., 2012). Both KOLFBAT Hom and Het cell lines were deemed karyotypically normal, with a median chromosome count of 46 per cell, and no metaphases containing more than 46 chromosomes was observed (**Fig. 5.3D**).

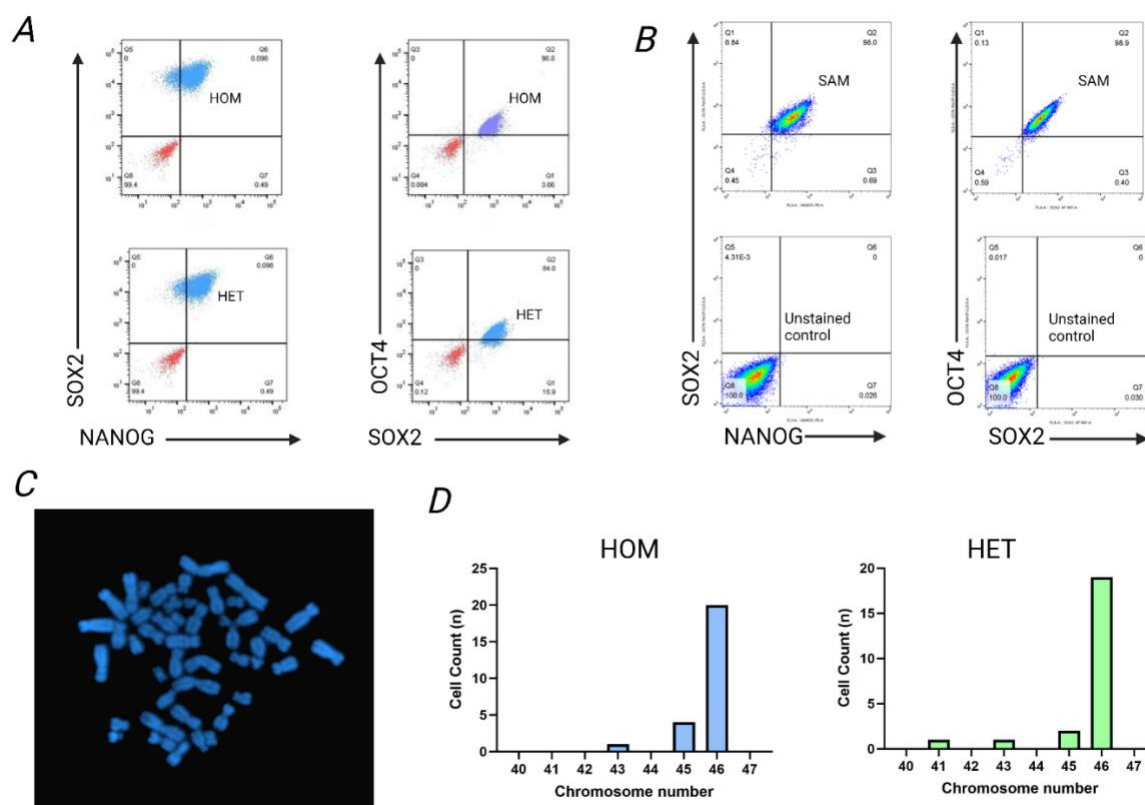


Figure 5.3 Quality control of KOLFBAT. **A)** FACS analysis of KOLFBAT Hom, KOLFBAT Het following staining with NANOG, SOX2 and OCT4 pluripotency markers. Blue cluster represent the stained population. Red cluster the cell group stained with isotype control antibodies. **B)** FACS analysis of KOLFSAM following staining with NANOG, SOX2 and OCT4 pluripotency markers provided by Dr. Rath. Top box shows the stained population, bottom box the population stained with the isotype control antibodies. **C)** DAPI-based chromosome staining of KOLFBAT Hom for karyotyping, provided by Dr. Moralli. **D)** Chromosome count in KOLFBAT Hom (blue) and KOLFBAT Het (green) population.

Validation of the Integrated CRISPRa-SAM in Pluripotent Stem Cells

Following the validation of BATV2 integration into the KOLFBAT cell lines, additional tests were conducted to establish whether the neuronally active hSyn promoter is able to drive the expression of the CRISPRa-SAM machinery in the pluripotent state.

To assess the correct expression of CRISPRa-SAM and β A-TEV1 Δ , total RNA was extracted from KOLFBAT (Hom and Het) and KOLFSAM cells. Subsequently, the RNA was reverse transcribed into cDNA, and the expression of dCas9 and β A was investigated through PCR amplification (**Fig. 5.4A/B**). The amplification of dCas9 and β A in the cDNA of both KOLFBAT Hom and Het was observed, with the amplification of dCas9 from KOLFSAM cDNA serving as the positive control. Importantly, since both Cas9 and β A sequences were without intronic regions, genomic contamination could have caused a false positive. To address this concern, non-reverse transcribed RNA was also used as a template for dCas9 and β A amplification, revealing a lack of bands for KOLFBAT Hom, Het, and KOLFSAM RNA, indicating the purity of the RNA from genomic

DNA contamination. Therefore, the observed amplification in the reverse transcribed templates suggests expression from the hSyn promoter.

Following that, a qPCR for dCas9 was conducted to explore the variation in expression strength between the CAG and hSyn promoters. Both KOLFBAT Hom and Het exhibited significantly lower expression levels, approximately 30% and 15% relative to the dCas9 expression seen in KOLFSAM (Fig. 5.4C).

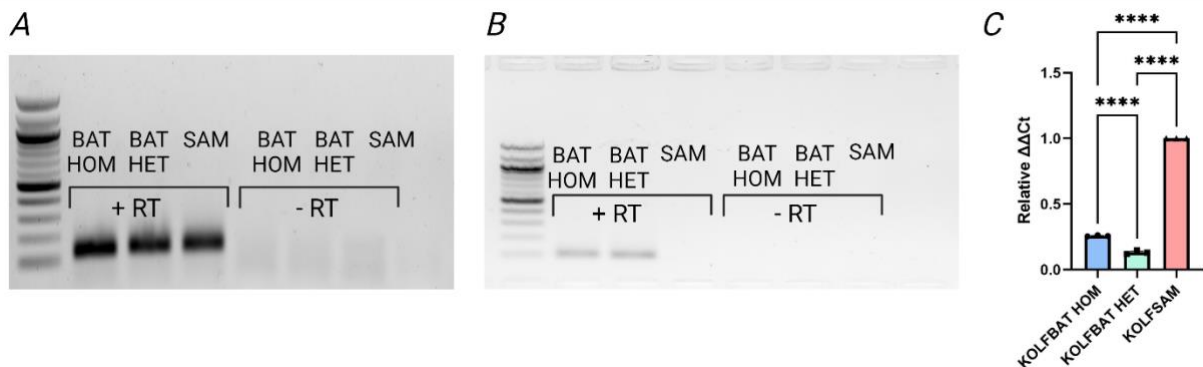


Figure 5.4 Investigation of Cas9 and βA expression in KOLFBAT. **A)** Gel electrophoresis of Cas9 PCR using the cDNA of KOLFBAT Hom, Het and KOLFSAM as template. **B)** Gel electrophoresis of PCR amplification of β Arrestin using the cDNA of KOLFBAT Hom, Het and KOLFSAM as template. **C)** Expression of Cas9 of KOLFBAT Hom (Blue) and KOLFSAM Het (Green) relative to KOLFSAM (Red). A one-way ANOVA demonstrated that the relative expression of the Cas9 in the three cell lines was significantly different [$F(2,6) = 8143$, $p < 0.0001$]. A Tukey's multicomparisons test indicated that each cell line expressed dCas9 significantly differently than the others ($p < 0.0001$). KOLFSAM expressing significantly more dCas9 due to the CAG promoter, followed by KOLFBAT Hom expressing Cas9 through the hSyn promoter located in both AAVS1 alleles and finally KOLFSAM Het expressing dCas9 through a hSyn promoter integrated into a single allele. Error bars represent standard deviation and sample number for each group = 3.

To further investigate the implications of the reduced hSyn driven CRISPRa-SAM expression in the pluripotent state, KOLFBAT (Hom) and KOLFSAM were transfected with plasmids containing the sgRNAs targeting *HHEX* and *MSX2*. As illustrated in Fig. 5.5A, when in the pluripotent state, the CAG-driven CRISPRa-SAM machinery in KOLFSAM induced an average upregulation of approximately 130-fold for the human *HHEX* gene (Fig. 5.5A) and an average of 30-fold for *MSX2* (Fig. 5.5B). In contrast, there was no discernible upregulation of the same target genes in KOLFBAT.

Based on this it can be inferred that in the pluripotent state, the observed lower expression level of hSyn promoter-driven CRISPRa machinery in the KOLFBAT cell lines was below a critical threshold for functionality, whereas the strong activity of the CAG promoter driving the machinery in the KOLFSAM cell line was sufficient to achieve robust CRISPR activation from endogenous promoters. These results underscore the differences in promoter strength between the CAG and hSyn in the pluripotent state. This data is consistent with the nature of the two promoters, where the CAG promoter is a synthetic promoter designed to induce high levels of ubiquitous expression

(Farokhimanesh et al., 2010). On the other hand, the human synapsin promoter is an endogenous promoter that is predominantly active in neuronal cells (Kügler et al., 2003).

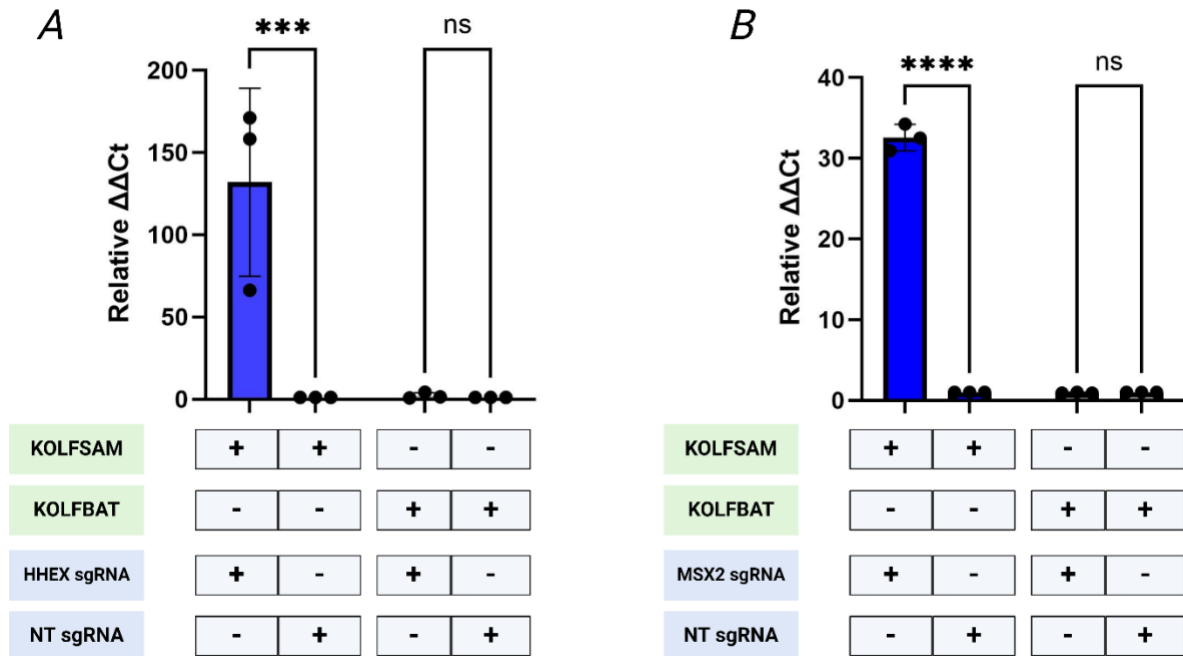


Figure 5.5 Investigation of CRISPRa-SAM expression in pluripotent KOLFSAM and KOLFBAT. **A)** Relative expression of HHEX pluripotent KOLFSAM (Blue) or KOLFBAT (Red) transfected with either the HHEX sgRNA or NT sgRNA. Green legend represents the cell line used and Blue legend represents the sgRNA transfected. A two-way ANOVA suggested a significant variation in hHEX upregulation based on cell line [$F(1,8) = 15.42$, $p = 0.0044$] and sgRNA [$F(1,8) = 15.97$, $p = 0.0040$]. A Fisher's least significant difference test indicated transduction of hHEX sgRNA in pluripotent KOLFBAT iPSC resulted in a significantly higher hHEX upregulation than transduction with the non-targeting control ($p = 0.0005$). On the other hand, no significant upregulation of hHEX in the pluripotent KOLFBAT iPSC transduced with either HHEX or NTC sgRNA could be noticed ($p = 0.9618$). Error bars represent standard deviation and sample size for each group = 3. **B)** Relative expression of MSX2 in pluripotent KOLFSAM (Blue) or KOLFBAT (Red) transfected with either the MSX2 sgRNA or NT sgRNA. Green legend represents the cell line used and Blue legend represents the sgRNA transfected. A two-way ANOVA suggested a significant variation in MSX2 upregulation based on cell line [$F(1,8) = 1096$, $p < 0.0001$] and sgRNA [$F(1,8) = 1090$, $p < 0.0001$]. A Fisher's least significant difference test indicated transduction of MSX2 sgRNA in pluripotent KOLFBAT iPSC resulted in a significantly higher MSX2 upregulation than transduction with the non-targeting control ($p < 0.0001$). On the other hand, no significant upregulation of MSX2 in the pluripotent KOLFBAT iPSC transduced with either MSX2 or NTC sgRNA could be noticed ($p = 0.9999$). Error bars represent standard deviation and sample size for each group = 3.

Differentiation of KOLFBAT/SAM into Cortical Neurons

To explore the potential activation of hSyn and subsequent CRISPRa functionality in the neuronal state, both cell lines underwent neuronal differentiation. Following the protocol outlined by Shi et al. (2012), KOLFBAT (Hom and Het), and KOLFSAM were exposed to Neuronal Maintenance medium (NMM) supplemented with SB431542 and Dorsomorphin dihydrochloride for 12 days after reaching 100% confluency. SB431542 and Dorsomorphin inhibit SMAD-dependent transforming growth factor morphogen (TGF- β) signalling, guiding differentiation toward the anterior neuroectodermal lineage (Chambers et al., 2009, Shi et al., 2012). Diverging

from the original protocol, the cells were singularized with Accutase after 12 days and further differentiated until they reached the required maturation (**Fig. 5.6**).

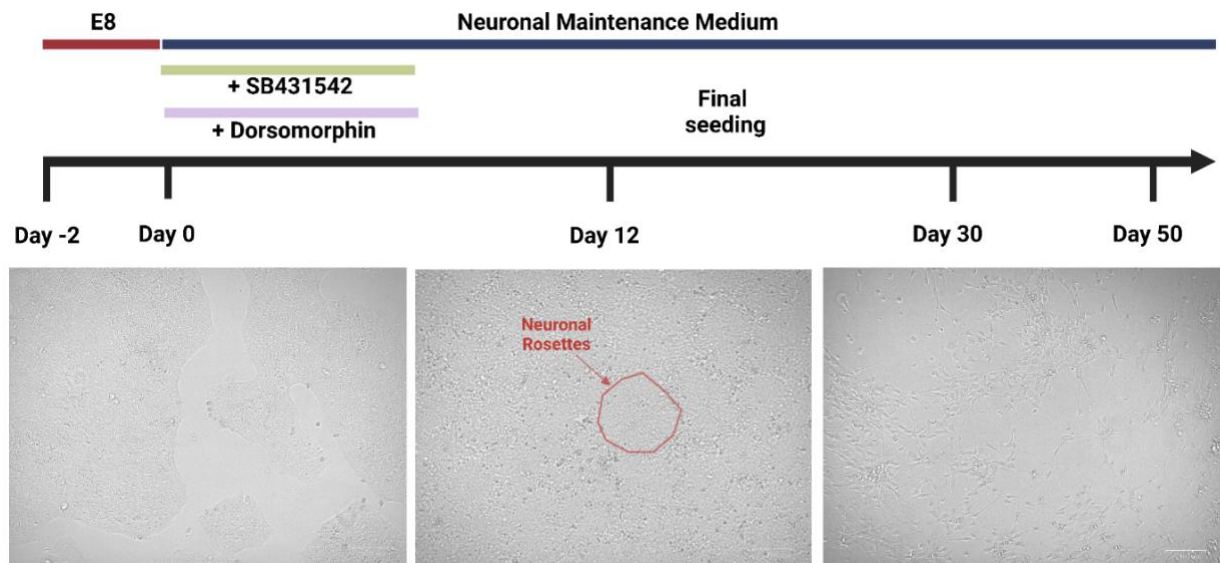


Figure 5.6 Neuronal differentiation of iPSC. Timeline of iPSC neuronal differentiation. The top portion shows the media used at specific timepoints. The media is changed from E8 to Neuronal Maintenance Media (NMM) combined with Dorsomorphin and SB431542. Cells are exposed to NMM only after the initial 10 days exposure. The bottom three pictures show the three major phenotypes seen through neuronal differentiation. The first picture shows classic pluripotent morphology characterised by large nuclei and clusters of cells separated by empty gaps between clusters. The middle picture shows the formation of neuronal rosettes. Cells are densely grouped in outward facing “rosettes”. The final picture shows the morphology of the cells as they reach the neuronal progenitor stage. Cells are well separated and are starting to display axonal growth.

Alongside the differentiation, RNA was extracted from the cell lines at specific time points (day 0, 24, 48, and 72), and qPCR analysis of pluripotency markers, as well as early, mid, and late neuronal markers, was performed to assess the effective differentiation into neurons.

In **Fig. 5.7A**, *OCT4* and *NANOG*, serving as markers of the pluripotent state, significantly declined with the differentiation, remaining consistently under 1-fold relative to day 0. Neuroectodermal markers *NESTIN* and *PAX6* were expected to increase in expression around the 10th day of differentiation and be sustained as neuronal rosettes formed (Callaerts et al., 1997, Shi et al., 2012). *PAX6* exhibited a peak upregulation at the day 24 time point, aligning with the anticipated differentiation, while *NESTIN* showed a robust peak in KOLFSAM but not in the two KOLFBAT lines, suggesting potential differences in the cell lines.

Late-stage biomarkers *SYNAPSIN 1*, and *vGLUT*, critical for cortical neuron maturation, all demonstrated upregulation at the day 72 time point, confirming the expected maturation of the cells.

KOLFBAT Hom underwent further characterization through immuno-staining of selected markers. At the Day 24 time point, *TBR2*, a neuronal progenitor marker expressed in the nucleus

(Hevner, 2019), and NESTIN, expressed in developing axons, were used for staining. The resulting images, depicted in **Fig. 5.7B**, reveal the presence of round, purple cellular compositions representing the developing neuronal rosettes at this stage. By day 55, when KOLFBAT was stained with the same markers, a significant transformation in cellular composition was observed. The rosettes had disappeared, and the neuronal progenitor cells had adopted a morphology more akin to mature neurons. In the final stage, at day 72, immune-staining was performed using TUJ1, which expresses neuron-specific tubulin (Menezes and Luskin, 1994), and BRN2 (Hagino-Yamagishi et al., 1997).

dCas9 expression was also analysed throughout differentiation (**Fig. 5.7C**). In KOLFSAM, where dCas9 is driven by the CAG promoter, a steady decline was observed across the time points, dropping below 1-fold relative to day 0. Conversely, KOLFBAT Hom and Het exhibited the opposite trend, with a relatively steady increase in Cas9 expression, reaching ~3-fold expression for KOLFBAT Hom compared to day 0.

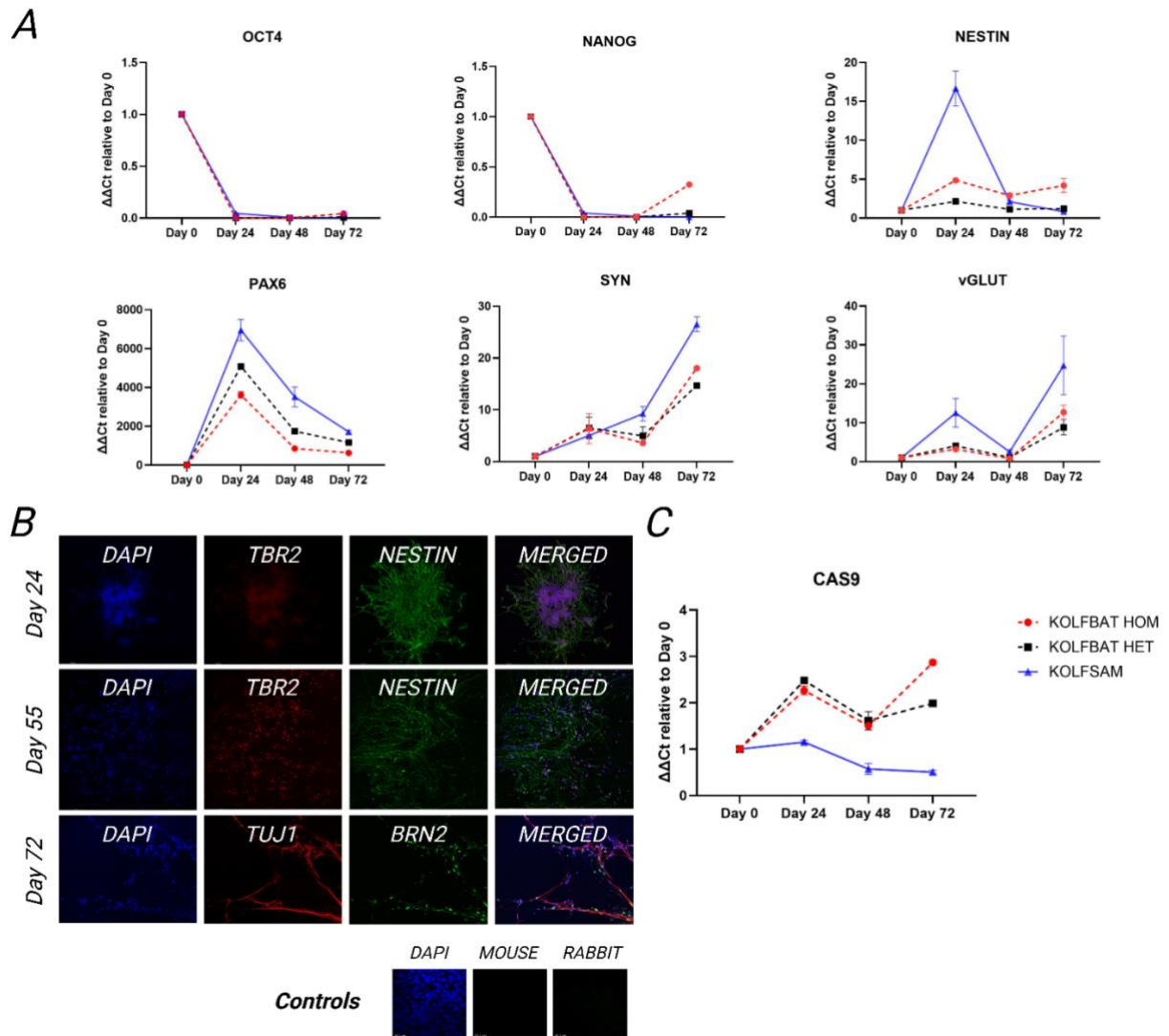


Figure 5.7 Investigation of neuronal markers across differentiation. **A)** qPCR analysis of neuronal biomarkers across specific differentiation timepoints, relative to day 0, for KOLFSAM (Blue), KOLFBAT HOM (Red) and KOLFBAT HET (Black). OCT4 and NANOG – two pluripotency markers – are shown to decrease after the first timepoint suggesting loss of pluripotency, NESTIN and PAX6 – two intermediate neuronal biomarkers – are shown to increase in expression around day 24. SYNAPSIN and vGLUT – two later stage neuronal markers – are shown to increase after day 48 and at the day 72 mark. **B)** Staining of KOLFBAT HOM across three differentiation timepoints (24/55/72). At day 24, KOLFBAT is stained with TBR2 a nucleic marker and NESTIN expressed in the developing axons. The same staining was carried out for day 55. At day 72, BRN2 and TUJ1 – two later stage markers – are successfully stained showing the developed axons. **C)** qPCR analysis of Cas9 across different time points relative to day 0, for KOLFSAM (Blue), KOLFBAT HOM (Red) and KOLFBAT HET (Black). Cas9 expression appears to fall below the 1-fold threshold after the day 12 time points in KOLFSAM and continuing to decrease in the later timepoints as well. Cas9 expression increases for KOLFBAT HOM/HET reaching about 3-fold by day 72.

These results confirmed the successful neuronal differentiation of the three cell lines into cortical neurons. Additionally, the results suggest that the ubiquitous synthetic promoter, CAG, is robust in the pluripotent cell state, but the expression is progressively silenced during differentiation. Conversely, the neuronally active hSyn promoter is weakly expressed in the pluripotent cell state but becomes more active as the cells differentiate toward a mature neuronal lineage. This variation in dCas9 expression is consistent with other studies, where the expression of the

transgene in the AAVS1 safe harbour locus is influenced by methylation during differentiation (Klatt et al., 2020, Karbassi et al., 2024).

To validate whether the increase in the hSyn-driven expression of the CRISPRa-SAM machinery was sufficient to induce expression of a target gene in differentiated neurons, mature cultures of neurons were transduced with guide RNA targeting specific genes of interest, utilizing lentiviruses to integrate the sgRNA due to their low transfection efficiency. The initial lentiviruses, containing sgRNA targeting *MSX2* and a non-targeting control, were obtained from VectorBuilder.

The efficiency of lentiviral sgRNA delivery was first validated on pluripotent KOLFSAM cells. Transduction of 50,000 cells with a titration of the *MSX2* lentiviruses determined that the multiplicity of infection (MOI) of 10 lentiviruses per cell resulted in the highest *MSX2* upregulation without significant cell death, as shown in **Fig. 5.8A** and **B**. To confirm the functionality of the CRISPRa-SAM systems in KOLFSAM and KOLFBAT using lentiviral sgRNA delivery in the pluripotent state, both cell lines were transduced with 10 MOI of *MSX2* lentiviruses and RNA was extracted 48 hours later. qPCR analysis of *MSX2* demonstrated that in KOLFSAM, *MSX2* upregulation was ~20 fold, but no upregulation of *MSX2* was observed in the pluripotent KOLFBAT cells, consistent with previous results (**Fig. 5.8C**).

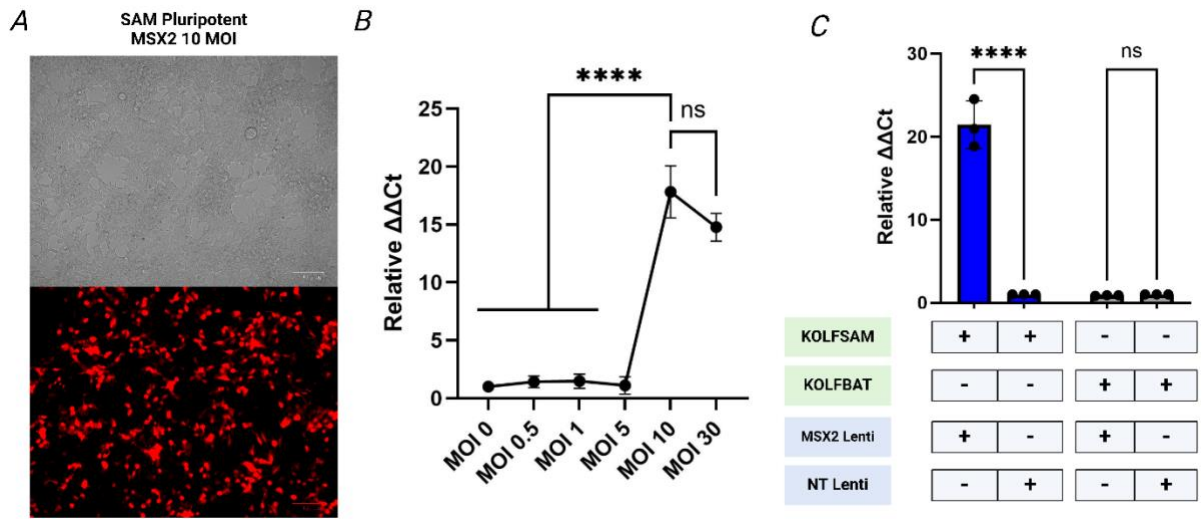


Figure 5.8 Lentiviral transduction in pluripotent iPSC. **A)** Fluorescence of mCherry expressed in the lentiviruses provided by VectorBuilder, demonstrating effective transduction in pluripotent iPSC. **B)** Curve of relative *MSX2* expression following transduction of sgRNA *MSX2* lentiviruses at different MOI for 50,000 cells. A one-way ANOVA analysis confirmed a significant difference between concentrations [$F(5, 12) = 142.8, p < 0.0001$]. A Tukey's multiple comparisons test demonstrated that MOI 10 induced significantly higher *MSX2* upregulation than the other MOIs ($p < 0.0001$) apart from MOI 30 ($p = 0.0565$) which however showed a decrease in *MSX2* upregulation. Error bars represent standard deviation and the sample size for each group = 3. **C)** Relative expression of *MSX2* in pluripotent KOLFSAM and KOLFBAT following transduction of sgRNA *MSX2* or non-targeting lentiviruses. A two-way ANOVA suggested a significant variation in *MSX2* upregulation based on cell line [$F(1,8) = 153.9, p < 0.0001$] and sgRNA [$F(1,8) = 150.7, p < 0.0001$]. A Tukey's multiple comparisons test indicated transduction of *MSX2* sgRNA in pluripotent KOLFBAT iPSC resulted in a significantly higher *MSX2* upregulation than transduction with the non-targeting control ($p < 0.0001$). On the other hand, no significant upregulation of *MSX2* in the pluripotent KOLFBAT iPSC transduced with either *MSX2* or NTC sgRNA could be noticed ($p > 0.9999$). Error bars represent standard deviation and sample size for each group = 3.

Following the validation of lentiviral sgRNA delivery in pluripotent stem cells, the *MSX2* sgRNA was tested in day 50 neurons. The selection of day 50 was based on large-scale RNA sequencing analysis by the Buckley lab, which quantified gene expression variation during the differentiation of iPS cells into neurons. The data revealed a significant upregulation of *SYNAPSIN* starting from day 50, consistent with literature results and the above data showing increased expression of *SYNAPSIN* between 48 and 72 days (Togo et al., 2021).

As performed with pluripotent stem cells, titrated *MSX2* lentiviruses were tested in KOLFBAT-derived neurons to determine the optimal MOI. As depicted in **Fig. 5.9A** and **B**, qPCR analysis of *MSX2* upregulation indicated that 60 MOI represented the optimal lentiviral concentration with the least percentage of cell death. As expected, the overall efficiency was considerably lower than observed in pluripotent stem cells.

Following the confirmation of the validated lentiviral concentration, KOLFSAM and KOLFBAT day-50 neurons were exposed to 60 MOI of *MSX2* or non-targeting control sgRNA lentiviruses. After 48 hours, RNA extraction for qPCR analysis revealed that *MSX2* upregulation in KOLFSAM-derived neurons was only ~5-fold compared to the ~25-fold expression seen in KOLFBAT-derived neurons (**Fig. 5.9C**) The reversal of CRISPRa-SAM efficiency observed in KOLFSAM and KOLFBAT between their pluripotent and neuronal states confirmed that the hSyn promoter is able to reverse the silencing of CRISPRa-SAM and instead allow CRISPR-activation in mature neurons. On the other hand, the CAG-driven CRISPRa-SAM is eventually silenced as the cell line moves away from its pluripotent state.

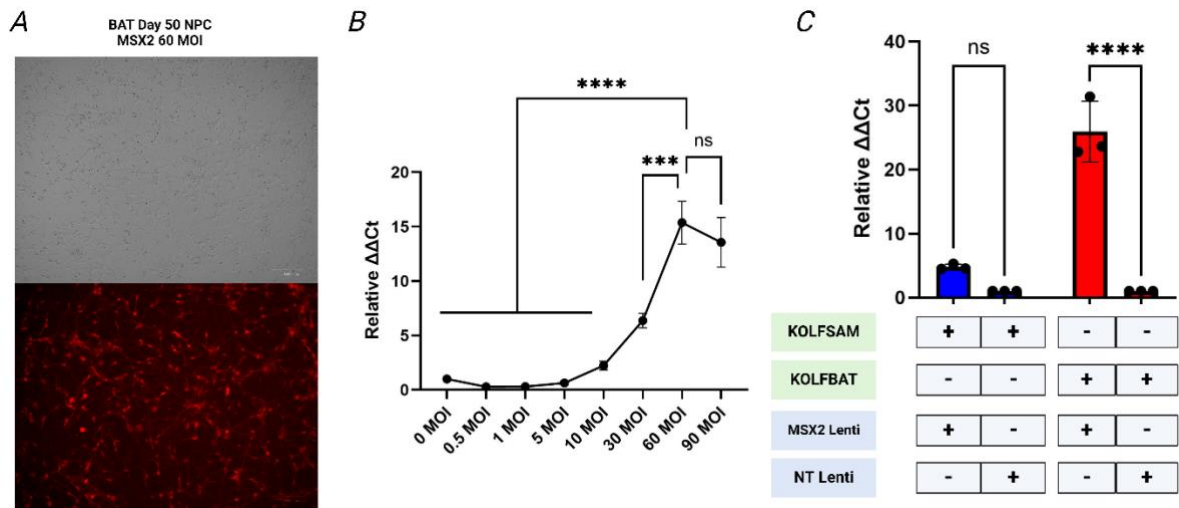


Figure 5.9 Lentiviral transduction in iPSC-derived neurons. **A)** Fluorescence of mCherry expressed in the lentiviruses provided by VectorBuilder, demonstrating effective transduction in Day 50 neuronal iPSC. **B)** Curve of relative *MSX2* expression following transduction of sgRNA *MSX2* lentiviruses at different MOI for 50,000 cells. A one-way ANOVA analysis confirmed a significant difference between concentrations [$F(7, 16) = 31.86, p < 0.0001$]. A Tukey's multiple comparisons test demonstrated that MOI 60 induced significantly higher *MSX2* upregulation than the other MOIs ($p < 0.0001$ and MOI 30 $p = 0.0006$) apart from MOI 90 ($p = 0.9305$) which however showed a decrease in *MSX2* upregulation

and increased cell death. Error bars represent standard deviation and the sample size for each group = 3. **C)** Relative expression of *MSX2* in pluripotent KOLFSAM and KOLFBAT following transduction of sgRNA *MSX2* or non-targeting lentiviruses. A two-way ANOVA suggested a significant variation in *MSX2* upregulation based on cell line [$F(1,8) = 59.01$, $p < 0.0001$] and sgRNA [$F(1,8) = 108.6$, $p < 0.0001$]. A Tukey's multiple comparisons test indicated transduction of *MSX2* sgRNA in pluripotent KOLFBAT NPC resulted in a significantly higher *MSX2* upregulation than transduction with the non-targeting control ($p < 0.0001$). On the other hand, no significant upregulation of *MSX2* in the KOLFSAM NPC transduced with either *MSX2* or NTC sgRNA could be noticed ($p = 0.2868$). Error bars represent standard deviation and sample size for each group = 3.

To further validate this result using an independent gene activation, a *HHEX* sgRNA and NT sgRNA were delivered to KOLFBAT and KOLFSAM-derived neurons by lentiviruses, generated in-house. Dr. Ping Zhang from the Knight lab produced and titrated the *HHEX* and NT lentiviruses. Day-50 KOLFBAT and KOLFSAM neurons were transduced with 60 MOI, and RNA extraction was performed after 48 hours. Similar to the CRISPR-activation of *MSX2*, hSyn-driven CRISPRa induced approximately a 25-fold upregulation of *HHEX* relative to the NT control, in contrast to the 4-fold upregulation achieved by the CAG-driven CRISPRa-SAM (**Fig. 5.10A**).

Finally, when comparing the relative upregulation in the pluripotent to the neuronal state following *MSX2* transduction, it is clear that the efficiency of CAG-driven CRISPRa-SAM decreases drastically as the cells differentiate into neurons whereas the hSyn-driven CRISPRa-SAM increases in efficiency (**Fig. 5.10B**). These findings confirm that the hSyn promoter can induce neuron-specific expression of CRISPRa-SAM when integrated into the *AAVS1* safe harbour locus, while stronger but more ubiquitous promoters such as the CAG promoter are silenced in the *AAVS1* locus as the cell differentiates into the neuronal lineage.

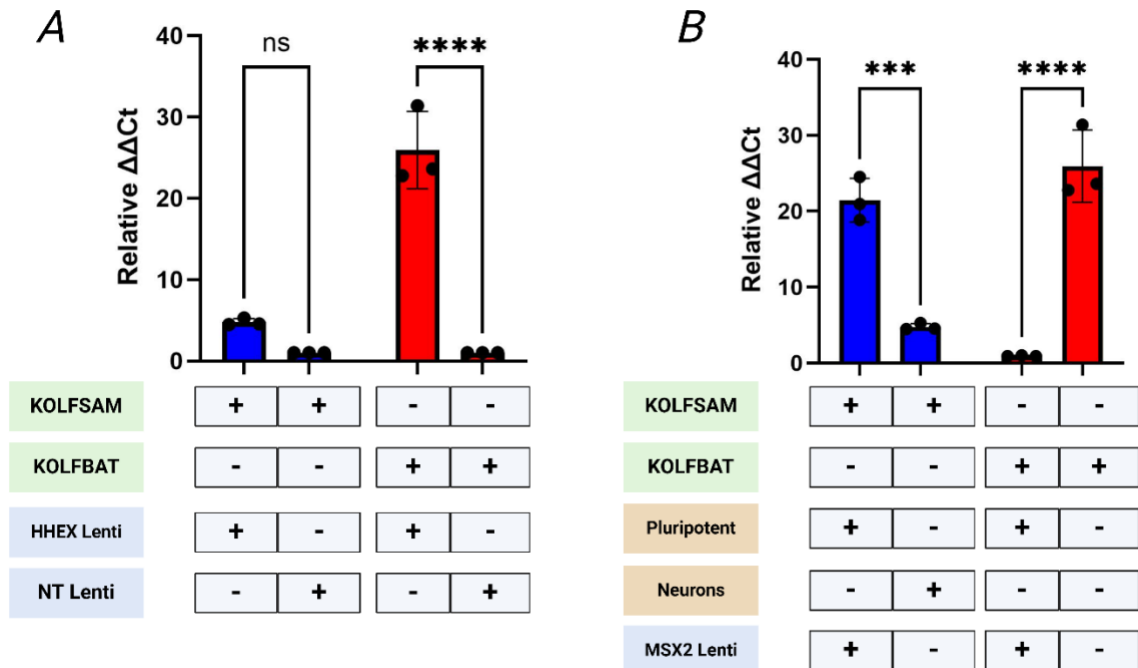


Figure 5.10 Comparison of CRISPRa-SAM driven gene expression between pluripotent and neuronal cell state. **A)** Relative expression of the *HHEX* gene in day 50 KOLFBAT (Blue) and KOLFSAM (Orange) derived neurons transduced

with lentiviruses containing either the HHEX sgRNA or non-targeting (NT) sgRNA. A two-way ANOVA suggested a significant variation in MSX2 upregulation based on cell type [$F(1,8) = 55.23$, $p < 0.0001$] and the sgRNA [$F(1,8) = 79.85$, $p < 0.0001$]. A Tukey's multiple comparisons test revealed no significant difference in MSX2 upregulation in the KOLFSAM NPC transduced with the lentiviral MSX2 sgRNA or the NTC sgRNA ($p = 0.7195$). On the other hand, a significant difference in the upregulation of MSX2 was found in the KOLFBAT NPC transduced with the lentiviral MSX2 sgRNA as opposed to the NTC sgRNA ($p < 0.0001$). Error bars represent standard deviation and sample size for each group = 3. **B)** Comparison of the relative expression of the MSX2 gene in either pluripotent or neuronal KOLFBAT and KOLFSAM cell lines. A two-way ANOVA suggested a significant variation in MSX2 upregulation based on differentiation [$F(1,8) = 6.673$, $p = 0.0325$] but not between cell types [$F(1,8) = 0.0442$, $p = 0.8387$]. A Tukey's multiple comparisons test suggested a significant drop in MSX2 upregulation between pluripotent and NPC KOLFSAM ($p = 0.0004$) while the opposite trend could be seen in KOLFBAT with a significant increase in MSX2 upregulation from pluripotent to NPC ($p < 0.0001$). Error bars represent standard deviation and sample size for each group = 3.

Closing Remarks

In this chapter, I detailed the process for the integration of the BATV2 sequence into the *AAVS1* safe harbour locus of KOLF2-C1 iPS cells by CRISPR-assisted gene targeting, resulting in the generation of KOLFBAT cell lines. For several years, the *AAVS1* locus has been considered a reliable safe harbour for mammalian integration. However, over the past decade a number of studies have reported the silencing of transgenes expressed in this locus during differentiation, especially CRISPR systems (Oceguera-Yanez et al., 2016, Ordovás et al., 2015, Karbassi et al., 2024). While the cause of the silencing is still unclear, I hypothesized that the use of an endogenous promoter, naturally expressed in the desired cellular state, might be able to counteract the natural silencing occurring at the *AAVS1* locus.

Having established the successful integration of BATV2 into *AAVS1*, I investigated the efficiency of *hSyn* to drive the expression of CRISPRa-SAM in both a pluripotent and neuronal state. As the Davies lab had already generated a similar cell line in which CRISPRa-SAM is driven by the CAG promoter and inserted into the same position within *AAVS1*, I used this opportunity to compare the role of the two promoters in maintaining robust CRISPRa-SAM expression as the cells differentiated into neurons. Throughout differentiation, the expression of CRISPRa-SAM was shown to decrease when driven by CAG and to increase when driven by *hSyn*, corroborating the initial hypothesis.

While establishing that the use of an endogenous cell-type specific promoter permits CRISPRa-SAM expression in the differentiated state, I also investigated if this expression was sufficient for functional CRISPR activation, that is to allow the efficient upregulation of a gene of interest. Neurons are very sensitive cell types, and lipofection can have adverse effects on morphology and survival, while also resulting in very low efficiency (Karra and Dahm, 2010). Therefore, I chose to deliver the sgRNA through lentiviral transduction, the generation of which was first outsourced and then generated in-house. Through lentiviral transduction of the *MSX2* and *HHEX*, I was able to induce target gene upregulation in the KOLFBAT-derived neurons.

With the generation of this cell line, I demonstrated that endogenous human promoters such as *hSyn* can be effective in counteracting the silencing of transgenes in the *AAVS1* locus of differentiating iPS cells and demonstrated that the cell line can be employed for the investigation of target gene upregulation in neurons. In the following chapter, I will employ KOLFBAT to investigate how *MSX2* upregulation in neurons affects the general transcriptome in relation to its possible role as a neuroprotective gene

Chapter 6 Effect of MSX2 upregulation on the transcriptome of iPSC-derived neurons

Introduction

The muscle segment homeobox 2 (*MSX2*) is a transcription factor associated with muscle, heart, and craniofacial development (Kwang et al., 2002, Chen et al., 2007, Han et al., 2007) playing a role in directing stem cell differentiation towards the mesendoderm lineage (Wu et al., 2015). *MSX2* is part of a family that also includes *MSX1* and both share a highly conserved homeodomain and protein structure, with the main distinction being the N-terminal region, which is 30 nucleotides longer in *MSX1* (Liu et al., 2021).

Initial studies suggested that *MSX2* primarily regulated gene expression through transcriptional repression (Newberry et al., 1997). However, recent research has highlighted its ability to induce the activation of other downstream regulators (Wu et al., 2015, Zhuang et al., 2009).

Among the pathways most closely associated with *MSX2* is the bone morphogenetic protein (BMP) pathway, a subset of the transforming growth factor- β family (TGF- β), comprising approximately 30 proteins involved in numerous pathways and cellular processes (Lowery and Rosen, 2018).

In the context of neuronal development, knock out of *MSX1* and 2 in mouse neural stem cells resulted in severe kinematic and athletic impairment, suggesting an involvement in neurogenesis (Yu et al., 2019). Moreover, *MSX2* activation has been associated with *ATOH1* expression, crucial for dorsal spinal neurogenesis (Duval et al., 2014) and cranial neural crest cell differentiation through *RUNX2* regulation (Han et al., 2007). Additionally, *MSX1* has been implicated in the accumulation of neurofibrillary tangles in AD patients, suggesting a potential role for the *MSX* family in Alzheimer's disease (Liang et al., 2018).

As mentioned in the introduction, *clusterin* (*CLU*) has been identified as the third most significant risk factor for late onset AD development (Lambert et al., 2013). The relationship between *CLU* and A β remains unclear, with some studies suggesting a role in A β clearance, while others imply a role of *CLU* in A β toxicity (Foster et al., 2019). Recent findings from the Buckley lab demonstrated that deletion of *CLU* in iPSC-derived neurons provided increased protection against A β -driven stress, suggesting that the role of *CLU* might be more detrimental than protective in the context of A β (Robbins et al., 2018). Furthermore, unpublished data from the same study revealed that *MSX2* is the most upregulated gene following A β -stress in the *CLU*-KO

iPSC-derived neurons. While the relationship between *CLU* and *MSX2* has not been well defined, a recent study unveiled that *CLU* promoted *MSX2* degradation to maintain *SOX2*-mediated pluripotency in oral cancer stem cells, indicating that *CLU* and *MSX2* could be working antagonistically (Praharaj et al., 2023).

Given *MSX2*'s involvement in neuronal and neurogenesis pathways, along with its emerging connection to *CLU*, it is hypothesized that *MSX2* may play a neuroprotective role in neuronal cells. By manipulating *MSX2* expression, we can potentially uncover the pathways regulated by this transcription factor and explore its neuroprotective effects. However, because *MSX2* is also critical in stem cell differentiation, its premature activation could disrupt neuronal development, highlighting the need for precise regulation.

In this chapter, I will explore the upregulation of *MSX2* in differentiated KOLFBAT neurons and assess its impact on the transcriptome through RNA sequencing analysis. Additionally, I will investigate potential downstream targets of *MSX2* in relation to neuroprotection. These experiments serve as proof-of-concept for the neuronal CRISPRa-SAM system developed in iPSC-derived neurons in the previous chapters and provide insight into this transcription factor as a potential therapeutic target.

MSX2 sgRNA transduction, RNA sequencing and quality control

To examine the impact of *MSX2* upregulation on the transcriptome, 50,000 day-50 KOLFBAT-derived neurons were transduced with either *MSX2* or NT sgRNA lentiviruses at 60 MOI in triplicates. Previous experiments indicated a clear upregulation of *MSX2* at the 48-hour mark. However, considering the potential need for a longer timeframe to allow the *MSX2* transcription factor to affect the transcriptome, RNA was extracted 48 hours and 10 days following transduction. Prior to sequencing, the expression level of *MSX2* was assessed via qPCR. Consistent with the previous results, qPCR analysis confirmed significant upregulation of *MSX2* after 48 hours (~20-fold) (**Fig. 6.1**). Notably, substantial upregulation of *MSX2* persisted 10 days post-transduction (~10-fold), albeit reduced compared to the 48 hours timepoint. Despite this reduction, significant *MSX2* upregulation was confirmed, prompting RNA purification and RNA sequencing, as outlined in the Methods section, on triplicates of *MSX2*-transduced and NT-transduced KOLFBAT-derived neurons at both 48 hours and 10 days timepoints.

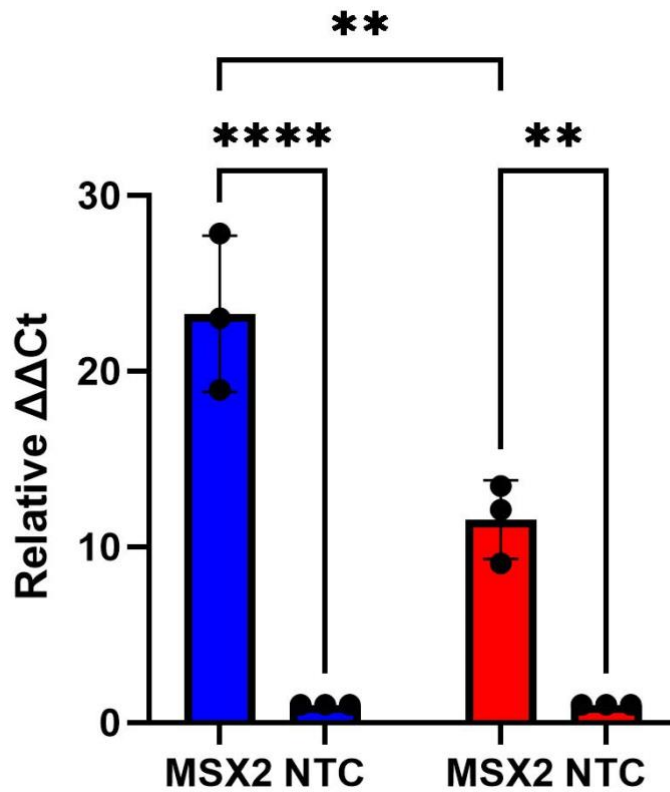


Figure 6.1 MSX2 upregulation at 48 hours and 10 days timepoints. Relative expression of MSX2 in 50-days KOLFBAT-derived neurons following transduction of the MSX2 or non-targeting sgRNA lentiviruses. The RNA was extracted at the 48 hours and 10 days timepoints, relative to the time of transduction. A two-way ANOVA suggested a significant variation in MSX2 upregulation based on timepoints [$F(1,8) = 16.53$, $p = 0.00036$] sgRNAs [$F(1,8) = 129.6$, $p < 0.0001$]. A Tukey's multiple comparisons test revealed a significant difference in MSX2 expression between the NPCs transduced with the MSX2 sgRNA and the NTC at both 48 hours ($p < 0.0001$) and 10 days ($p = 0.0037$). A significant drop in MSX2 upregulation in the NPCs transduced with MSX2 sgRNA from the 48 hours and the 10 days timepoints ($p = 0.0019$). Error bars represent standard deviation and sample size for each group = 3.

RNA sequencing quality control

Triplicate libraries for both MSX2 and NT control at 48 hours and 10 days were sequenced with NextSeq 500/550 High Output Kit v2.5 (75 Cycles) single-reads (Illumina®), yielding a mean read depth of $45,559,164 \pm 1.8$ million reads per sample (**Fig. 6.2A**). For the analysis, datasets were labelled based on the treatment (MSX2 and NT), the time point (48 hours and 10 days) and the triplicate as such: MSX2_48h_A/B/C, MSX2_10d_A/B/C, NT_48h_A/B/C, NT_10d_A/B/C. Euclidean distance between groups was calculated using the following formula $\sqrt{\sum_{i=1}^n (x_i - y_i)^2}$ to determine the square root of the sum of the square differences between the two corresponding groups. This value indicates the similarity between groups, 0 indicates no differences between the groups and an increase in the value is correlated to the distance between the groups. Euclidean distance analysis revealed no outliers within each triplicate group and demonstrated

a positive correlation between triplicates and treatment groups (MSX2 vs NT). This correlation correctly grouped individual reads within their respective sample group and treatment (**Fig. 6.2B**). Furthermore, principal component analysis confirmed the distinct separation of the four groups, with the treatment (sgRNA) explaining 70% of the variability, notably with the MSX2_10d group showing the greatest variation (**Fig. 6.2C**). Approximately 16% of the variability was attributed to temporal variation, as evidenced by the clustering of 48-hours samples at the bottom and 10-days samples at the top. These quality control analyses confirmed the high-quality sequencing of the transcriptomes and demonstrated clear variability between samples.

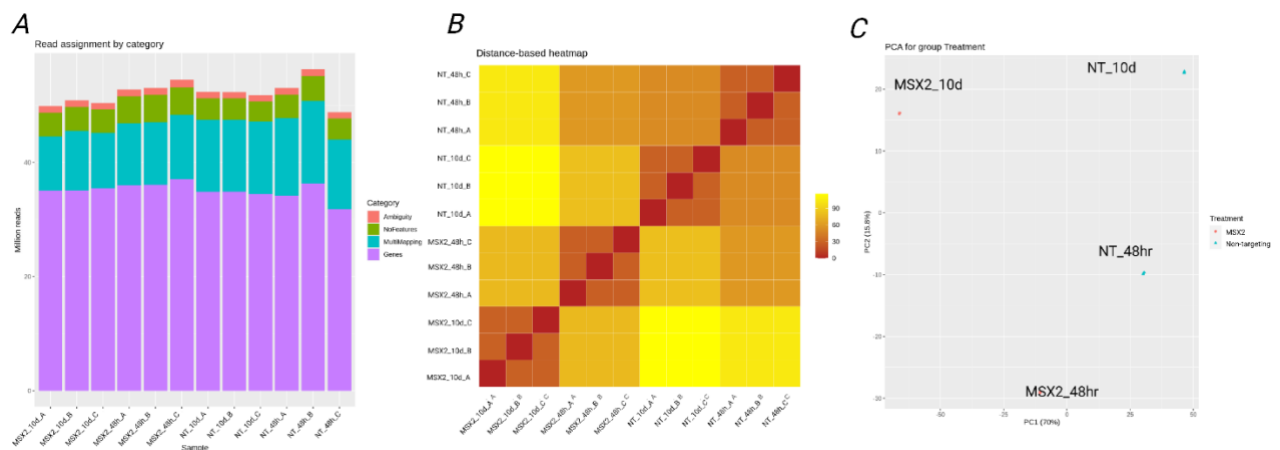


Figure 6.2 Quality Control analyses of RNA sequencing results. **A)** Read count for each sample. Read count showed a consistent average of ~45,000,000 reads per sample, except for NT_48h_C which produced ~41,000,000. The gene count (pink) varied around ~35,000,000 reads per sample, which was within the expected range. Similarly, ~9,000,000 reads were found to be multimapped, i.e. mapped to more than one genomic region (turquoise), which was within the expected threshold. Minimal ambiguous (orange) and “non-features” (green) reads were recorded. **B)** Distance based heatmap. Euclidean distance is calculated between each sample to determine similarity. Similarity is defined by a colour gradient: red (0 distance between samples) and yellow (> 90 distance between samples). All triplicates were found to be correctly grouped within each treatment group confirming the absence of outliers. The groups were also most closely correlated by treatment, with MSX2_10 and MSX2_48hr being shown to be more closely correlated than NT_10d and NT_48h. **C)** Principal component analysis plot. The plot shows 4 distinct clusters confirming clear variation between groups. 70% of the variation (x-axis) was explained by the treatment, with MSX2_10d being the most different group, followed by MSX2_48h before the controls on the right-side of the plot. ~16% of the variance (y-axis) was explained by time, with the 10 days samples closer to the top of the plot and the 48 hours samples at the bottom.

Differential analysis of 48 hours and 10 days samples

To examine the impact of the MSX2 transcription factor on the transcriptome, the R package DESeq2 (Love et al., 2014) was used for the differential analysis, comparing treatment datasets by timepoints *i.e.* MSX2_48h with NT_48h and MSX2_10d with NT_10d. An adjusted p-value (padj) for the differentially expressed genes was carried out using the Benjamini-Hochberg procedure to decrease the false discovery rate (filtered for a $\text{padj} \leq 0.05$). With filtering, a total of 10,460 differentially expressed genes were found at the 48-hours timepoint and 12,094 for the 10-days timepoint. A dispersion plot was used to determine how well the DESeq2 model fitted the data. As shown in **Fig. 6.3**, dispersion estimates for each gene were generated (black), based on maximum likelihood estimation of dispersion. Finally, each gene dispersion estimate was “shrunk” towards the fitted curve (blue). The shrunk dispersion estimates that follow the curve suggest that they fit the DESeq2 model correctly. Dispersion estimates (blue circles) that did not fit the curve suggested that their variability was too high to follow the modelling assumptions set by DESeq2 and so their dispersion values were not shrunk towards the curve. As anticipated, with the increase in the mean of normalized counts, gene dispersion decreased. Notably, the majority of genes in the datasets aligned well with the curve, indicating a good fit for the DESeq2 model (**Fig. 6.3A** for 48 hours and **6.3B** for 10 days). The use of technical replicates, along with the robustness of *MSX2* upregulation, likely accounts for the extensive number of differentially expressed genes and their low dispersion.

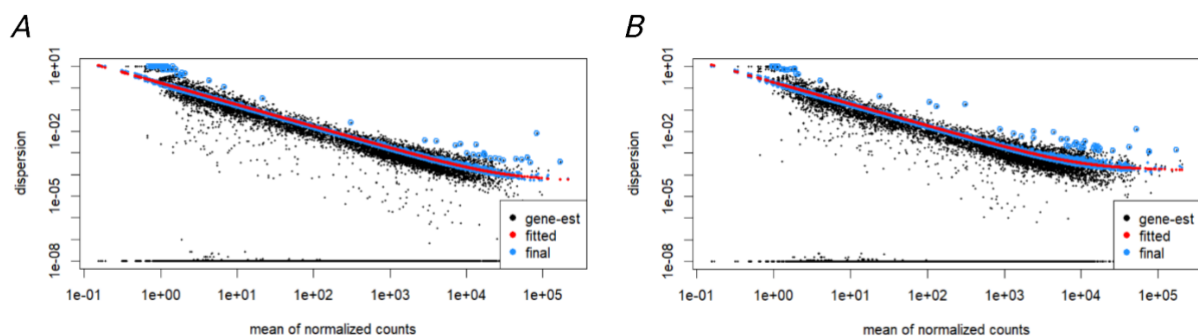


Figure 6.3 Dispersion plots for the 48 hours and 10 days datasets. The line (red) is fit to the data shrinking resulting in the final fitted dispersion estimates (blue). As normalised counts (x-axis) increase, dispersion (y-axis) decreases. Resulting in a curved shaped plot. **A)** Dispersion plot of the genes differentially expressed in the MSX2_48h RNA sequencing compared to the NT_48h sequencing. Each dot represents the dispersion. A few genes (labelled in blue) do not fit the modelling assumption and as such do not fit the fitted curve (blue). The well-fitted curve and the low number of outliers suggest that data was a good fit for the DESeq2 model. **B)** Dispersion plot of the genes differentially expressed in the MSX2_10d RNA sequencing compared to the NT_10d sequencing. Each dot represents the dispersion. A few genes (labelled in blue) do not fit the modelling assumption and as such do not fit the fitted curve (blue). The well-fitted curve and the low number of outliers suggest that data was a good fit for the Deseq2 model.

To visualize the extent of differential expression for each of the genes, volcano plots with stringent criteria ($\text{Log}_2\text{Fold change threshold} > 2$ and $\text{padj} < 10^{-50}$) were employed. The volcano plot of the differentially expressed genes at 48 hours revealed that, although most genes exhibited significantly low padj values ($< 10^{-10}$), their $\text{Log}_2\text{Fold change}$ predominantly hovered around 1, as evidenced by the blue cluster (**Fig. 6.4A**). Conversely, the volcano plot at 10 days showed a notable increase in both the quantity of differentially expressed genes and the magnitude of difference. Many genes exhibited both very low padj values ($< 10^{-50}$) and $\text{Log}_2\text{Fold changes} > 2$ (red clusters), particularly among downregulated genes (left-hand section of the plot) (**Fig. 6.4B**). The top 15 most differentially expressed genes were highlighted in both plots. A comparison between the two volcano plots clearly demonstrates an escalation in differentially expressed genes from 48 hours to 10 days, supporting the hypothesis that *MSX2* has a cumulative effect on the transcriptome over time.

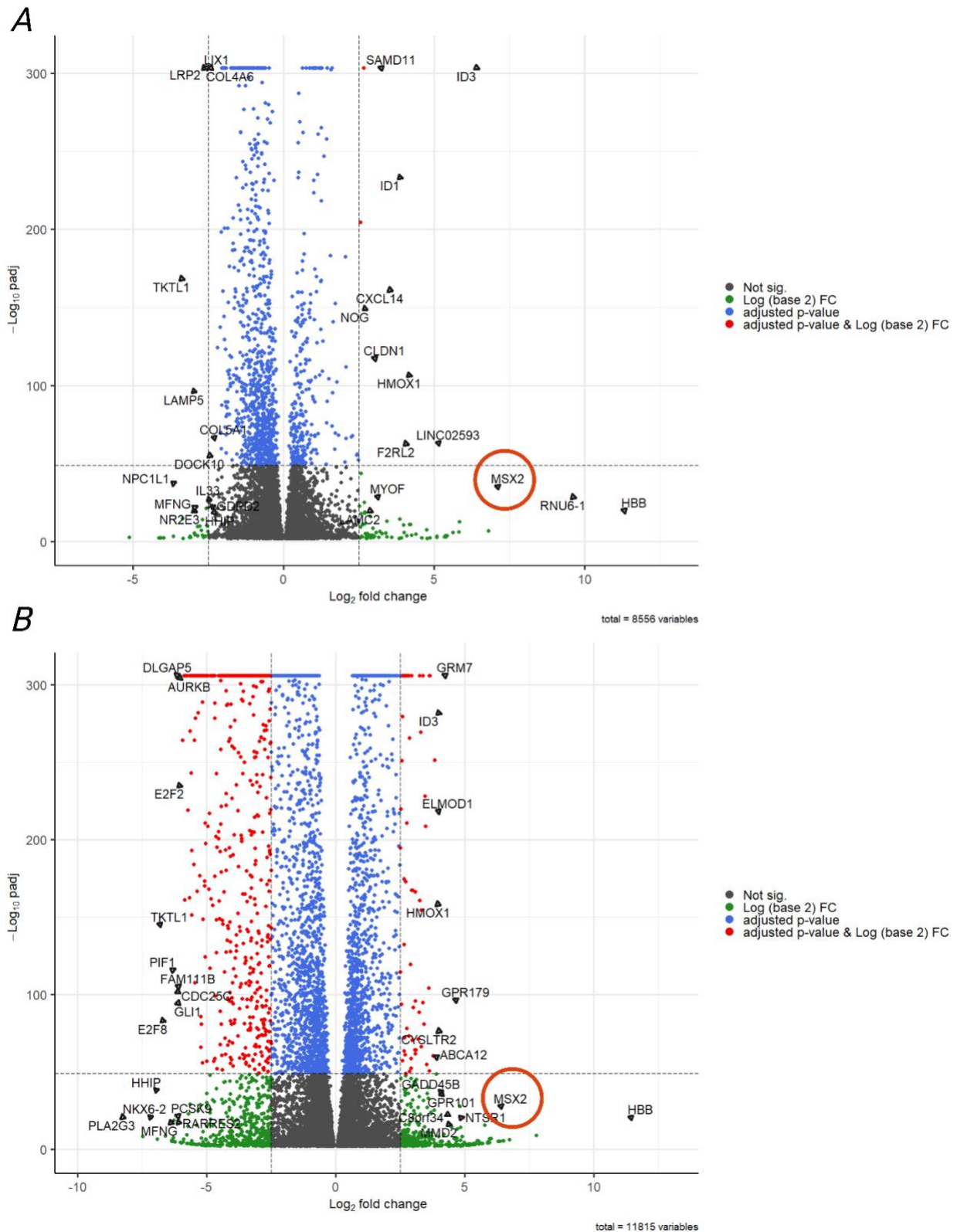


Figure 6.4. Volcano plots of differentially expressed genes 48 hours and 10 days after MSX2 transduction. A very stringent padj of 10^{-50} (dotted line on the y-axis) was selected together with a Log_2Fold change of 2 (two dotted lines on the x-axis) were set as threshold to visualise the extent of significant in the two datasets. The genes below this threshold are in grey. The genes with a $\text{Log}_2\text{Fold} > 2$ but a $\text{padj} > 10^{-50}$ are shown in green. The genes with a $\text{Log}_2\text{Fold} < 2$ and $\text{padj} < 10^{-50}$ are shown in blue. Finally, the genes with a $\text{Log}_2\text{Fold} > 2$ and a $\text{padj} < 10^{-50}$ are shown in red. The highlighted names are the top 15 genes based on Log_2Fold and with a $\text{padj} < 0.05$. **A)** The volcano plot showing the differentially expressed genes at the 48 hours timepoint, shows a large quantity of genes within the grey and blue range, suggesting that most differential expression was $< 2 \text{ Log}_2\text{Fold}$. **B)** The volcano plot showing the differentially expressed genes at the 10 days timepoint, shows a marked increase in genes with $\text{Log}_2\text{Fold} > 2$ and a $\text{padj} < 10^{-50}$ suggesting an increase in the number and extent of differentially expressed genes.

A heatmap was generated to visualize the top and bottom 25 differentially expressed genes at the 48-hours timepoint (**Fig. 6.5**) and the 10-days timepoint (**Fig. 6.6**). *MSX2* was notably highly expressed in both datasets, ranking as the third most upregulated gene in the 48-hours dataset (**Fig. 6.5**) with a 7.11 Log₂ fold change and the second most upregulated in the 10-days dataset (**Fig. 6.6**) with a 6.4 Log₂ fold change. This finding validates the overexpression of *MSX2* initially confirmed through qPCR and further illustrates the decrease in *MSX2* expression from 48 hours to 10 days, confirming the reliability of the sequencing results. Additionally, both datasets exhibited haemoglobin subunit beta (*HBB*) as the most consistently upregulated gene across the two timepoints, with an 11.31 Log₂fold change at 48 hours (**Fig. 6.5**) and an 11.43 Log₂ fold change at 10 days (**Fig. 6.6**).

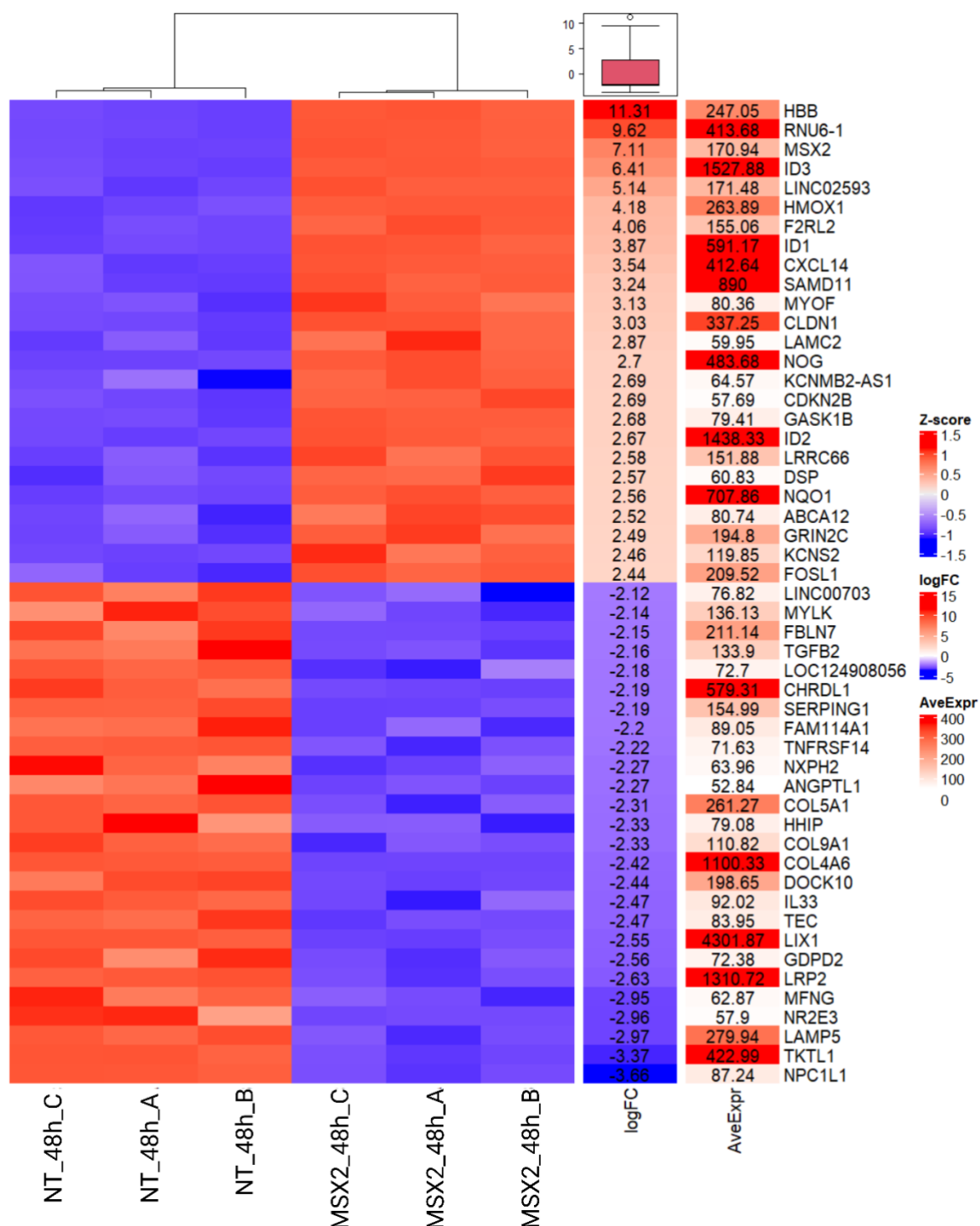


Figure 6.5 Heatmap of the top and bottom 25 differentially expressed genes in the 48 hours timepoint analysis. The heatmap shows three variables (left to right). The standard score (Z-score) shows the number of standard deviations (SD) by which each gene differs in expression compared to the average gene expression within each dataset. Red depicts a 1.5 SD increase from the mean and blue depicts -1.5 SD decrease from the mean. The top two quadrants show the most upregulated genes in the MSX2 dataset relative to the NT dataset, the bottom two show the opposite. The second column shows the Log₂ fold change of each gene from the MSX2 triplicates compared to the NT control. The last column shows the Average RNA expression of each gene with a red gradient going from white = 0 gene copies to red = > 400 gene copies.

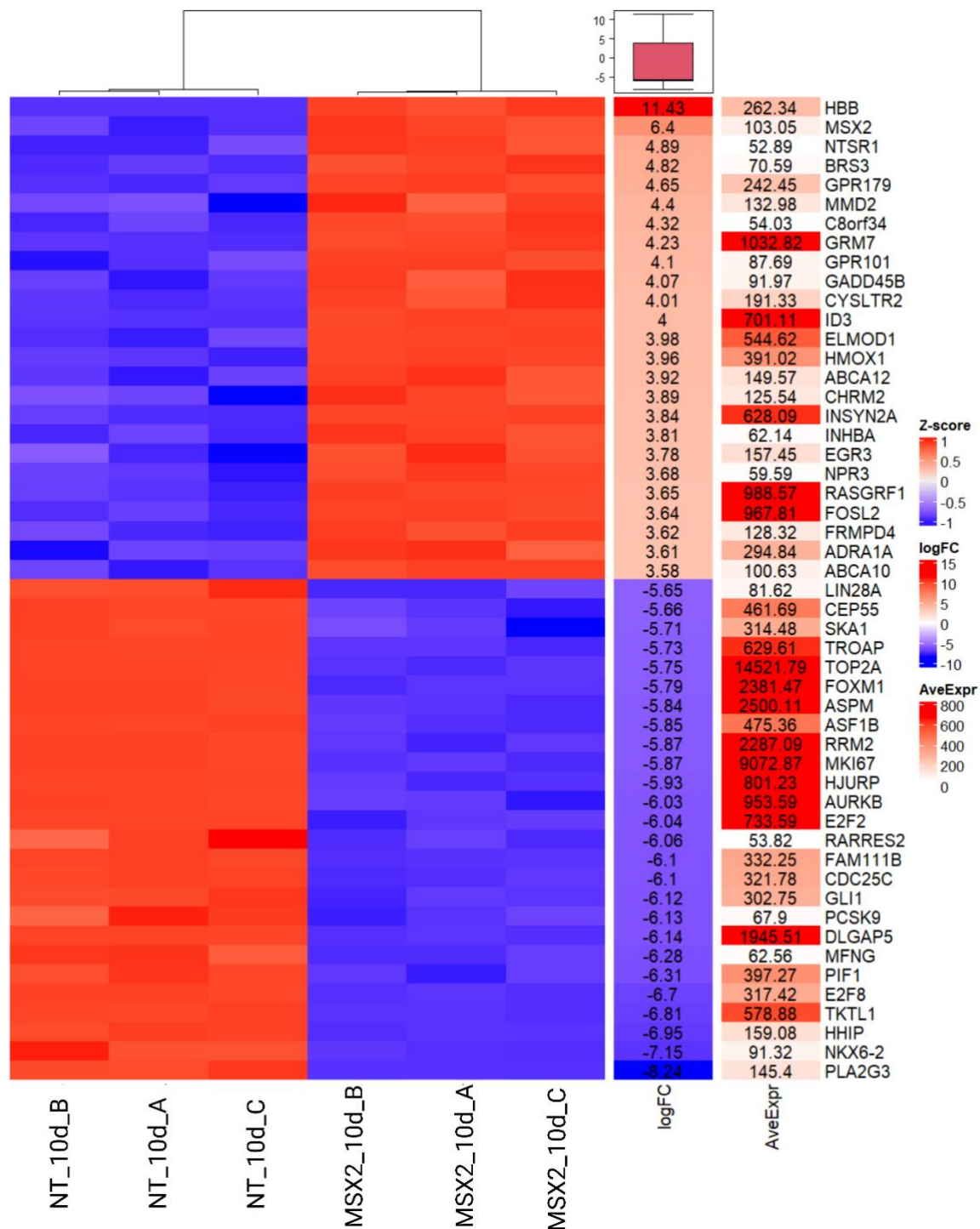


Figure 6.6 Heatmap of the top and bottom 25 differentially expressed genes in the 10 days timepoint analysis. The heatmap shows three variables (left to right). The standard score (Z-score) shows the number of standard deviations (SD) by which each gene differs in expression compared to the average gene expression within each dataset. Red depicts a 1.5 SD increase from the mean and blue depicts -1.5 decrease from the mean. The top two quadrants show the most upregulated genes in the MSX2 dataset relative to the NT dataset, the bottom two show the opposite. The second column shows the Log₂fold change of each gene from the MSX2 triplicates compared to the NT control. The last column shows the Average RNA expression of each gene with a red gradient going from white = 0 gene copies to red = > 400 gene copies.

Gene Set Enrichment Analysis

Whilst investigating individual differentially expressed genes can shed some light on how *MSX2* influences the neuronal transcriptome, the sheer number of differentially expressed genes necessitated a more comprehensive approach to avoid overlooking crucial information. Gene set enrichment analysis (GSEA) was conducted to explore whether the differential gene expression patterns in the 48 hours and 10 days datasets could be associated with various phenotypes. Unlike canonical gene ontology analysis, GSEA does not rely on predetermined cutoffs but instead utilizes the entire dataset (as detailed in the Methods)(Mootha et al., 2003, Subramanian et al., 2005). Given the extensive differential gene expression observed, this approach was deemed more robust.

The GSEA analysis identified two sets of positively and negatively enriched terms for the 48 hours and 10-day datasets. As depicted in **Table 6.1**, the top 20 enriched terms with an adjusted p-value of <0.01 were selected. Both datasets primarily revolved around neuronal functions terms, with a particular focus on synaptic terms: cellular components (CC), molecular function (MF), and biological processes (BP).

Table 6.1 Top 20 positively enriched terms in the 48-hours dataset (top) and 10-days dataset (bottom). The “Source” column records the Gene Ontology (GO) ID for each term. The “Ont” column records the gene ontology category: cellular component (CC), molecular function (MF), biological process (BP). The estimation score (ES) indicates the maximum deviation from 0 and it represents the degree to which a gene set is represented at the extreme of the rank list generated from the 48 hours or 10-days dataset. The normalised estimation score (NES) is equal to the value of each ES divided by the average of the enrichment score of all the dataset permutations performed by the software. Adjusted p-value was < 0.01. Adjusted p-value was based on the Benjamini-Hochberg false discovery method to each NES value, generating an adjusted p-value (padj). In the GSEA report, a padj value of zero (0.0) indicates an actual p value of less than 1/number-of-permutations.

Positively correlated enriched terms in 48-hours dataset (padj <0.01)					
Source	Term	Ont	SIZE	ES	NES
GO:0045211	POSTSYNAPTIC MEMBRANE	CC	273	0.560096	2.439137
GO:0097060	SYNAPTIC MEMBRANE	CC	389	0.537941	2.435568
GO:0030594	NEUROTRANSMITTER RECEPTOR ACTIVITY	MF	96	0.633628	2.425927
GO:0098839	POSTSYNAPTIC DENSITY MEMBRANE	CC	99	0.629402	2.408881
GO:0099634	POSTSYNAPTIC SPECIALIZATION MEMBRANE	CC	122	0.619003	2.396988
GO:0098960	POSTSYNAPTIC NEUROTRANSMITTER RECEPTOR ACTIVITY	MF	69	0.646294	2.340356
GO:0060078	REGULATION OF POSTSYNAPTIC MEMBRANE POTENTIAL	BP	144	0.569464	2.320368

GO:0007215	GLUTAMATE RECEPTOR SIGNALING PATHWAY	BP	50	0.686557	2.30777
GO:0008066	GLUTAMATE RECEPTOR ACTIVITY	MF	27	0.778905	2.305457
GO:0032922	CIRCADIAN REGULATION OF GENE EXPRESSION	BP	68	0.64692	2.303949
Positively correlated enriched terms in 10-days dataset (padj <0.01)					
Source	Term	Ont	SIZE	ES	NES
GO:0045211	POSTSYNAPTIC MEMBRANE	CC	273	0.658409	2.736641
GO:0099634	POSTSYNAPTIC SPECIALIZATION MEMBRANE	CC	122	0.724116	2.731984
GO:0097060	SYNAPTIC MEMBRANE	CC	389	0.624668	2.644233
GO:0030594	NEUROTRANSMITTER RECEPTOR ACTIVITY	MF	96	0.726953	2.642774
GO:0060078	REGULATION OF POSTSYNAPTIC MEMBRANE POTENTIAL	RB	144	0.684838	2.601628
GO:0098960	POSTSYNAPTIC NEUROTRANSMITTER RECEPTOR ACTIVITY	MF	69	0.757885	2.580239
GO:0098839	POSTSYNAPTIC DENSITY MEMBRANE	CC	99	0.714051	2.57071
GO:0005230	EXTRACELLULAR LIGAND GATED MONOATOMIC ION CHANNEL ACTIVITY	MF	70	0.725017	2.491325
GO:0022835	TRANSMITTER GATED CHANNEL ACTIVITY	MF	58	0.753732	2.48757
GO:0007215	GLUTAMATE RECEPTOR SIGNALING PATHWAY	BP	50	0.782838	2.461002

Moreover, GSEA analysis effectively distinguishes between the most upregulated and downregulated terms. **Table 6.2** displays the most downregulated terms for both the 48-hours and 10-days datasets. In both cases, there is a noticeable downregulation of terms related to cellular division and replication.

The observed positive enrichment of synaptic-related terms and the negatively enriched terms on cell replication pathways may be attributed to an involvement from *MSX2* in the differentiation of pluripotent stem cells into mesendoderm, as reported by Wu et al. (2015), resulting in a non-dividing phenotype.

Table 6.2 Top 20 negatively enriched terms in the 48-hours dataset (top) and 10-days dataset (bottom). The “Source” column records the Gene Ontology (GO) ID for each term. The “Ont” column records the gene ontology category: cellular component (CC), molecular function (MF), biological process (BP). The estimation score (ES) indicates the maximum deviation from 0 and it represents the degree to which a gene set is represented at the extreme of the rank list generated from the 48-hours or 10-days dataset. The normalised estimation score (NES) is equal to the value of each ES divided by the average of the enrichment score of all the dataset permutations performed by the software. Adjusted p-value was < 0.01. Adjusted p-value was based on the Benjamini-Hochberg false discovery method to each NES value, generating an adjusted p-value (padj). In the GSEA report, a padj value of zero (0.0) indicates an actual p value of less than 1/number-of-permutations.

Downregulated pathways terms in 48-hours dataset						
Source	Term	Ont	SIZE	ES	NES	
GO:0006261	DNA TEMPLATED DNA REPLICATION	BP	162	-0.71173	-2.78394	
GO:0006270	DNA REPLICATION INITIATION	BP	37	-0.86339	-2.67894	
GO:0005657	REPLICATION FORK	CC	66	-0.7579	-2.60254	
GO:0007059	CHROMOSOME SEGREGATION	BP	434	-0.59065	-2.57606	
GO:0051983	REGULATION OF CHROMOSOME SEGREGATION	BP	131	-0.66368	-2.55261	
HP:0000172	ABNORMAL UVULA MORPHOLOGY	HP	190	-0.63726	-2.53942	
GO:0006260	DNA REPLICATION	BP	277	-0.60572	-2.52054	
GO:0000819	SISTER CHROMATID SEGREGATION	BP	228	-0.61014	-2.51453	
HP:0000048	BIFID SCROTUM	HP	36	-0.81781	-2.50965	
HP:0003510	SEVERE SHORT STATURE	HP	106	-0.67818	-2.50793	
Negatively correlated enriched pathways terms in 10-days dataset						
Source	Term	Ont	SIZE	ES	NES	
GO:0000070	MITOTIC SISTER CHROMATID SEGREGATION	BP	187	-0.79989	-2.54552	
GO:0000819	SISTER CHROMATID SEGREGATION	BP	228	-0.78267	-2.52904	
GO:0098813	NUCLEAR CHROMOSOME SEGREGATION	BP	312	-0.76003	-2.50806	
GO:0051983	REGULATION OF CHROMOSOME SEGREGATION	BP	131	-0.81341	-2.50649	
GO:0051304	CHROMOSOME SEPARATION	BP	81	-0.85679	-2.48062	
GO:0006261	DNA TEMPLATED DNA REPLICATION	BP	162	-0.78327	-2.46552	
GO:0007059	CHROMOSOME SEGREGATION	BP	434	-0.7403	-2.46491	
GO:0140014	MITOTIC NUCLEAR DIVISION	BP	276	-0.75744	-2.45935	
GO:0033045	REGULATION OF SISTER CHROMATID SEGREGATION	BP	104	-0.81351	-2.4293	
GO:0050000	CHROMOSOME LOCALIZATION	BP	116	-0.7923	-2.42222	

GeneFeast to visualise gene enrichment analysis

Gene enrichment analysis offers valuable insights into the terms associated with *MSX2*-modulation of the transcriptome. Specifically, GSEA allows us to focus on the weighted impact of each gene on driving the enrichment of specific GO terms. To summarize and explore similarities between enriched terms, both phenotypically and based on the shared genes driving them, I employed the GeneFeast algorithm (Taylor et al., 2023). I analysed all enriched terms for both datasets, generating “metacommunities” where specific phenotypes and the genes driving them were shared. Given the observed similarities between the 48-hours and 10-days datasets, I concentrated on investigating the latter.

Analysis of the 10-days datasets yielded two metacommunities (metacommunity 01 and 02) for the most upregulated terms and five for the most downregulated terms, of which only two provided terms relevant to this study (metacommunities 03 and 04) (see Methods section).

As depicted in **Table 6.3**, the metacommunities correspond approximately to the main terms discussed earlier. Enriched neuronal terms primarily focus on later-stage neuronal development such as “synaptic membrane” and “synaptic signalling”, while downregulated terms involved “cellular differentiation” and “proliferation”. This analysis further underscores the impact of *MSX2* on neuronal pathways and highlights the modulation of cellular processes related to differentiation and proliferation, corroborating the pattern seen from the top 20 terms initially detailed by the GSEA.

Table 6.3 GeneFeast metacommunities. Metacommunities for the positively (01 and 02) and negatively (03 and 04) terms in the 10-days datasets. Metacommunities were grouped based on the GeneFeast algorithm through analysis of similar genes associated to the terms. The “Ont” column records the gene ontology category: cellular component (CC), molecular function (MF), biological process (BP).

Day 10 most upregulated metacommunities		
Meta Community 01		Ont
	SYNAPTIC MEMBRANE	CC
	MONOATOMIC CATION CHANNEL ACTIVITY	MF
	TRANSPORTER COMPLEX	CC
	REGULATION OF TEMPLATE POTENTIAL	BP
	NEUROTRANSMITTER RECEPTOR ACTIVITY	MF
Meta Community 02		
	REGULATION OF TRANS SYNAPTIC SIGNALING	BP
	VESICLE MEDIATED TRANSPORT IN SYNAPSE	BP
	EXOCYTIC VESICLE	CC
Day 10 most downregulated metacommunities		
Meta Community 03		Ont
	DNA REPLICATION	BP
	CHROMOSOMAL REGION	CC
	MEIOTIC CELL CYCLE	BP
	MIDBODY	CC
	NON MEMBRANE BOUND ORGANELLE ASSEMBLY	BP
	SPINDLE	CC
	CATALYTIC ACTIVITY ACTING ON AGING	MF
	DNA RECOMBINATION	BP
	NUCLEAR CHROMOSOME	CC
Meta Community 04		
	GLIOGENESIS	BP
	NEURAL PRECURSOR CELL PROLIFERATION	BP

Discussion on differential analysis

In this chapter, I have explored the application of KOLFBAT-derived neurons, which express the hSyn-driven CRISPRa-SAM machinery robustly, as a potential model for investigating the impact of possibly neuroprotective genes on the transcriptome. Recent findings from the Buckley lab revealed that silencing the AD-risk associated gene *CLU*, leads to possible neuroprotection against A β -induced stress and to the significant upregulation of *MSX2*. This led me to hypothesize that studying the effects of *MSX2* upregulation on the transcriptome of iPSC-derived neurons could offer valuable insights into its putative neuroprotective effect.

The results from the previous chapter demonstrated noticeable upregulation of *MSX2* 48 hours post lentiviral transduction. However, recognizing the necessity for an adequate time for the *MSX2* transcription factor to influence the transcriptome, I decided to explore the effects of *MSX2* upregulation at both 48 hours and 10 days post lentiviral transduction.

Differential analysis was conducted by comparing RNA sequencing data of the samples transduced with *MSX2* sgRNA or a non-targeting (NT) control. This analysis revealed a substantial number of significantly differentially expressed genes, increasing the complexity of downstream analysis.

However, both datasets indicated that *MSX2* was among the top differentially expressed genes (7.11 and 6.41-Log₂fold), indicating that the CRISPRa-SAM driven upregulation of the gene was successful. Furthermore, *HBB* was found to be the most upregulated gene in both datasets (11.31 and 11.43- Log₂fold). Recent studies have revealed an involvement for *MSX2* in inducing the differentiation of embryonic stem cells into the mesendoderm lineage, acting as a direct target of the BMP signalling pathway, which is already known to regulate this process (Wu et al., 2015). Moreover, *MSX2* has been shown to play an active role in initiating and accelerating the formation of neural crest cells, a cell population crucial for the development of mesenchymal cells (Zhang et al., 2018). However, *MSX2* suppression was also reported to direct embryonic stem cells towards the hematopoietic lineage (Wang et al., 2020). Interestingly, a number of studies reported significant decrease of haemoglobin in AD animal models and patients (Ferrer et al., 2011, Orre et al., 2014) leading to the suggestion that intracellular haemoglobin might play a protective role against hypoxia in AD (Altinoz et al., 2019). While no direct correlation has been established between *HBB* and *MSX2*, further investigation on the role of *HBB* in AD is warranted.

Another gene shared by both datasets is the inhibitor of DNA binding 3 (*ID3*), which exhibited strong differential expression (6.41 and 4- Log₂fold). *ID3* has been reported to be co-induced with

MSX2 by the BMP signalling pathway, potentially explaining its upregulation in the datasets (Hollnagel et al., 1999). Furthermore, although present only in the 48 hours dataset, *ID1* and *ID2* are also involved downstream in the BMP pathway (Yang et al., 2013). The heme oxygenase 1 (*HMOX1*) was also highly upregulated in both datasets, with only a slight decrease in expression from 48 hours to 10 days (4.18 to 3.96-Log₂fold). HO-1, transcribed by *HMOX1*, is involved in heme degradation, exerting important anti-inflammatory and antioxidant roles (Ryter et al., 2006, Jansen and Daiber, 2012). It is also believed to play a role in mesenchymal cell differentiation and hematopoietic stem cell differentiation processes (Szade et al., 2021).

As discussed above, *MSX2* primarily exerts its regulatory roles through gene repression, underscoring the importance of investigating the most differentially downregulated genes (Wu et al., 2015, Wang et al., 2020). Hedgehog interacting protein (*HHIP*), and the manic fringe gene (*MFNG*) were all markedly downregulated in both datasets. While no specific studies have reported a direct association between *MSX2* and *HHIP*, *HHIP* is a marker of quiescent hepatic stellate cells, which diminishes as the cells differentiate into hepatic stellate cells where *MSX2* expression predominates (Choi et al., 2009). This observation may explain the effect seen in the experiment.

Manic fringe (*MFNG*) is one of the three fringe genes involved in the Notch signalling pathway (Yi et al., 2013). While *MFNG* itself does not appear directly associated with *MSX2*, the Notch signalling pathway is closely linked with the BMP pathway. *Notch1*, for instance, has been shown to interact with BMP2 to increase *MSX2* expression, and *MSX2* has been implicated in regulating T-cell differentiation through modulation of the Notch3 pathway (Shimizu et al., 2011, Nagel et al., 2009). This relationship between *MSX2* and the Notch pathway may explain the *MFNG* downregulation. The robust differential expression of these genes, alongside their potential associations with *MSX2* or the pathways it regulates, strongly suggest successful upregulation of *MSX2* and its ability to perturbate the transcriptome.

Based on this, it is worth briefly discussing the phospholipase A2 group III gene (*PLA2G3*), which emerged as the most downregulated gene in the 10-days dataset. Interestingly, *PLA2G3* has been identified as the most overexpressed gene in human neurons AD models under oxidative stress (Martínez-García et al., 2010). In addition, the study demonstrated that *PLA2G3* silencing led to reduced cell death, suggesting a potential role for *PLA2G3* as a minimal but compounding risk factor in Alzheimer's disease (AD) progression. Similarly, research by Yui et al. (2015) reported a significant increase in *PLA2G3* expression in the brains of Alzheimer's disease model mice.

While the investigation of the most differentially genes provided valuable information the number of differentially expressed genes suggested the requirement of a larger scale analysis. For this reason, I performed gene set enrichment analysis, which does not necessitate the setting of predefined thresholds, hypothesizing that it would yield more reliable results given the large set of differentially expressed genes. The analysis unveiled both positively and negatively enriched terms, suggesting an increase in synaptic development pathways and a decrease in cell proliferation pathways, respectively.

To further investigate how the interaction between these pathways and the genes driving them, I utilised the GeneFeast algorithm. Focusing on the 10-days dataset, I identified a number of metacommunities, groups of pathways which shared the genes most associated with them. *NTSR1* (4.89-fold), *CHRM2* (3.89-fold), *KCNK6* (3.79-fold), *ADRA1A* (3.60-fold), and *GABRA4* (3.52 fold) were identified as the most upregulated genes involved in the 01 and 02 metacommunities. Activation of neurotensin receptor 1 (*NTSR1*) has been shown to increase neuronal excitability, thereby improving spatial learning and memory in Alzheimer's disease (AD) mouse models (Xiao et al., 2014). Similarly, the muscarinic acetylcholine receptor M2 (*CHRM2*) has been implicated in learning and memory processes and has been shown to rescue memory loss in AD when activated in mice (Liu et al., 2022a). Notably, the adrenoceptor α 1A has been linked to AD development and is considered a potential target for therapeutic intervention (Hassan et al., 2019). Finally, γ -aminobutyric acid receptor alpha 4 (*GABRA4*), a neurotransmitter receptor involved in neuronal development and memory, has been found to be downregulated in AD patients (Govindpani et al., 2020, Fan et al., 2020).

Conversely, the most downregulated genes in the metacommunities 03 and 04 were *PIF1* (-6.31), *AURKB* (-6.03), *ASPM* (-5.8), *TOP2A* (-5.7), and *NEK2* (-5.6). While *PIF1* is associated with telomere shortening and is considered a risk gene for AD, the other genes do not have a direct implication in AD development and are primarily involved in cellular differentiation and proliferation (Tirozzi et al., 2023).

Closing Remarks

These pathways and the genes driving them provide valuable insights on the possible effect of *MSX2* in the context of neuronal transcriptome, but do not show any clear indication on how *MSX2* can directly provide protection against A β -driven stress.

The transcriptome analysis of *MSX2*-upregulated KOLFBAT neurons serves as an important proof of concept for the application of the KOLFBAT cell line to investigate target gene upregulation in iPSC-derived neurons.

Further investigation on the role of *MSX2* in relation A β can be carried out through the use of KOLFBAT-derived neurons. *MSX2*-upregulated KOLFBAT derived neurons for examples could be exposed to A β to observe any variation in stress resistance compared to WT neurons or even CLU-KO, employed by Robbins et al. (2018) demonstrating the flexibility of this cell line as a model for the study of neurodegeneration.

Chapter 7 Discussion

Integrating synthetic biology tools into pluripotent stem cells can provide novel avenues for investigating and treating complex disorders. The decision-making capability of gene circuits could enable precise regulation of cell behaviour in response to the environment resulting in more sophisticated cell therapies to be developed. The pluripotency of stem cells facilitates the production of various human cell types at scale and *in vitro*, providing relevant models for studying intra and intercellular mechanisms of disease.

Over the past two decades, the synthetic toolkit available for eukaryotic cells has expanded, with CRISPR nucleases playing a central role due to their versatility. Combining Cas molecules with guide RNA sequences enables precise genome targeting, revolutionizing genetic engineering in mammalian cells. The evolution of CRISPR from a nuclease to a transcription factor (CRISPRa/i) allows temporal regulation of target gene expression (Qi et al., 2013, Gilbert et al., 2013, Konermann et al., 2015). Further developments incorporating AND/OR/NOR logic gates demonstrate improved precision.

Neurodegenerative disorders, such as Alzheimer's disease, are complex and suitable for investigation and even potential therapeutic advances using biologically engineered induced pluripotent stem cells (iPSCs).

Differentiation of stem cells into various somatic cells is accompanied by changing morphology of the genome with sequences becoming epigenetically activated and silenced as the cells reach their final state. This makes gene circuit integration into stem cells challenging as prediction of the behaviour of a circuit in the relevant differentiated cell type can be impacted by its final epigenetic landscape. Similarly, synthetic overexpression of both exogenous and endogenous target genes can impact cell metabolism and misdirect differentiation to unwanted states.

In this project, I focused on developing an A β -activated and CRISPR-driven gene circuit in iPSC-derived neurons, aiming to address these challenges. The results of the study highlight the complexity of integrating finely tuned gene circuits into iPSCs, specifically in the context of controlled expression of sgRNAs as well as the challenges of combining independent components into a unified circuit. As part of the gene circuit optimization, I was able to genetically engineer an iPSC line with a CRISPRa-SAM expression system tailored to neurons, serving as a proof of concept for future gene circuit-fitted iPSC models. Simultaneously, this system demonstrated the potential benefits of engineered iPSCs for the research and selection of potential new therapeutic gene targets for neurodegenerative disorders. As a proof of concept,

I used this cell line to investigate the role of the transcription factor MSX2 in modulating the neuronal transcriptome to investigate a possible neuroprotective role against A β .

In this chapter, I will summarise and discuss the results and provide an in-depth overview of the limitations and future applications of this study, proposing possible solutions for circuit optimization and to further understand the role *MSX2* in neuroprotection using the same KOLFBAT-derived neurons.

Results Summary

The thesis introduced the design of an A β -activated gene circuit in iPSC-derived neurons, relying on three core components: a CMKLR1-Tango receptor sensing extracellular A β , an inducible tTA-based promoter regulating sgRNA expression, and a CRISPRa-SAM machinery driving neuroprotective gene expression. To validate this design, individual investigations of each component were conducted (**Fig. 7.1**).

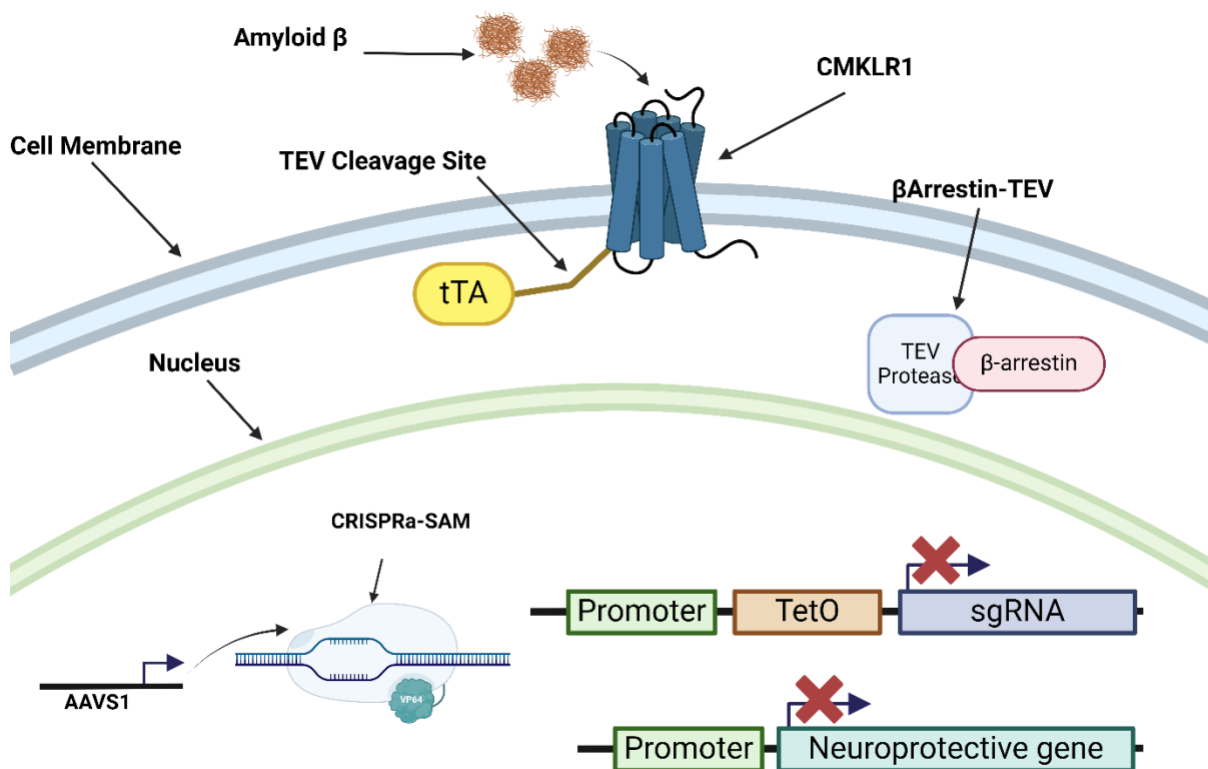


Figure 7.1 Diagram representing the components of the A β -activated gene circuit. CMKLR1-Tango receptor is fused to tTA in the cytoplasm. In the nucleus, sgRNA expression is prevented by an inducible promoter comprising of TetO sequences. The CRISPRa-SAM machinery is constitutively expressed but inactive without the sgRNA. Activation of CMKLR1 by interaction with A β triggers β Arrestin-TEV which frees the tTA. The tTA interacts with the TetO inducing expression of the sgRNA which directs the CRISPRa-SAM to the target gene for expression.

Based on previous reports of CMKLR1-A β interaction, activation of the CMKLR1-Tango system following A β exposure was examined. The system effectively sensed external A β , offering potential applications in A β -activated cell models and therapies.

The activation of the Tango receptor system leads to the release of a tTA molecule, which is designed to induce gene expression under the control of a minimal Pol II promoter combined with TetO sequences. While some studies have successfully developed tTA-repression systems for regulating sgRNA, mechanisms for tTA-induced sgRNA expression have remained elusive. Shaw et al. (2022) introduced the induction system that most closely aligned with the requirements of the circuit discussed here but required exogenous anhydrotetracycline for rtTA activation, underscoring the need for a tTA-inducible sgRNA mechanism to enhance the mammalian gene circuit toolbox. In this study, I demonstrated that embedding sgRNA within two Csy4 cleavage sites, expressed by a minimal Pol II promoter combined with TetO sequences, was the most effective strategy for tTA-inducible sgRNA expression.

With validated components, the CRISPRa-SAM machinery with β A-TEV1 Δ and the CMKLR1-Tango system with the Csy4:sgRNA inducible mechanism were successfully cloned as the BATV2 and CLICK4 targeting vector plasmids respectively. Both these plasmids contain homology arms suitable for their targeted integration into the human safe-harbour loci, *AAVS1* and *CLYBL*. While the BATV2 components functioned as expected, the Csy4:sgRNA mechanism in the CLICK4 plasmid failed to upregulate the gene of interest, although testing was only performed in a non-neuronal cell line. This highlights the challenge of engineering different individual components to behave as a cohesive signalling pathway, but also the need for more physiologically appropriate model systems for simple circuit testing and optimization.

Due to the reported repression of ubiquitously expressed exogenous sequences integrated into the *AAVS1* safe harbour locus during stem cell differentiation (Ordovás et al., 2015, Bhagwan et al., 2020, Klatt et al., 2020, Karbassi et al., 2024), the human synapsin promoter (hSyn) was tested to drive CRISPRa-SAM and β A-TEV1 Δ expression. I hypothesized that an endogenous neuronal promoter would counteract the repression during neuronal differentiation. Integrating BATV2 into iPSCs (KOLFBAT), resulted in basal Cas9 expression increasing during neuronal differentiation, indicating the efficacy of the hSyn promoter in counteracting *AAVS1* locus silencing. This was supported by the side-by-side comparison of the KOLFSAM iPSC line, carrying CAG-driven CRISPRa-SAM within the same locus, which showed decreased Cas9 expression during the same differentiation. As a result, gene upregulation via the integrated CRISPRa-SAM machinery could

readily upregulate target gene expression in differentiated KOLFBAT but not in KOLFSAM cells. This behaviour was reversed when using the systems in the undifferentiated state.

Having demonstrated that KOLFBAT was able to maintain a steady expression of CRISPRa-SAM in its neuronal state, I investigated the effect of *MSX2* upregulation on the neuronal transcriptome. Studies by the Buckley lab, demonstrated that knock-out of the AD risk associated gene *clusterin* (*CLU*), resulted in a neuroprotective phenotype following A β -induced stress (Robbins et al., 2018). Unpublished RNA sequencing data from the same study demonstrated that the *CLU*-KO neurons after A β -induced stress showed significant upregulation of *MSX2*. While no clear indication of the potential neuroprotective effect of *MSX2* on the transcriptome could be definitively inferred, I was able to assess and quantify the perturbation caused by the *MSX2* transcription factor on the transcriptome.

The development and utilization of KOLFBAT for investigating the role of putative therapeutic target genes, such as *MSX2* in Alzheimer's disease, offers valuable insights into the potential of engineered iPSCs for studying and treating neurodegenerative disorders. On the other hand, while the circuit design was ultimately unsuccessful, several reflections on its components can provide useful insights for future applications and studies.

CMKLR1-Tango can be a suitable sensing mechanism for A β

The role of GPCRs as a receptor for A β is increasingly recognized in the context of Alzheimer's disease (Thathiah and De Strooper, 2011). CMKLR1 was found to be co-localised with A β , but its role in AD pathogenesis could not be determined (Peng et al., 2015). Recently, Chen et al. (2022) highlighted an involvement of CMKLR1 in microglia recruitment for A β removal. Their study showed that CMKLR1 activation promotes actin and microtubule remodelling via the p38 MAPK signalling pathway, facilitating microglial activation toward A β deposits. Additionally, the study suggested that chemerin, CMKLR1 main ligand, could reduce A β -induced microglia aggregation. However, while microglial activation aids in A β clearance, it may not directly correlate with improved AD conditions. Deletion of *CMKLR1* was found to increase A β deposits, in line with the above experiments, but also decrease mortality and tau hyperphosphorylation, suggesting a complex role in AD pathology. Therefore, while microglial activation is crucial for A β reduction, chronic inflammation associated with microglial activation may exacerbate AD pathology, consistent with previous studies (Guan et al., 2022). Further investigation into the role of this receptor in different neuronal cell types and its interactions with other factors may provide additional insights into its involvement in AD pathogenesis.

In the context of gene circuits, the validation of the Tango system's interaction with A β opens various applications for the development of gene circuits targeting A β . CMKLR1-Tango can potentially be applied in cell therapies to induce targeted removal of A β or as a tool for drug discovery to investigate compounds that modulate CMKLR1 activation and its inflammatory role. When investigated in stem cells, CMKLR1-Tango could also be utilized to compare CMKLR1 activation effects across different cell lines, such as neurons and microglia, to identify variations in behaviour.

It is worth noting that CMKLR1 is just one of many GPCRs associated with AD (Haque et al., 2018). Receptors like GPR18 and GPR55, also integrated as Tango receptor systems, offer expanded opportunities for AD-focused GPCR drug discovery (Hickman and El Khoury, 2010, McHugh et al., 2010). For example, despite being considered an orphan receptor, GPR18 has been demonstrated to be involved in the resolution phase of inflammatory microglia responses suggesting its potential therapeutic relevance in AD. Similarly, various GPCRs have been strongly associated with tau pathology, including receptors such as the M1 muscarinic acetylcholine receptor, the β 2 adrenergic receptor, as well as the CXCR2 and CCR3 receptors, expanding the range of potential applications for AD modelling (Chidambaram and Chinnathambi, 2020)

However, it is essential to consider the limitations of the Tango system. Firstly, the luciferase assays conducted in this project were not normalized with a secondary luciferase component, such as Renilla. While this approach was consistent with the original data from Barnea et al. (2008), (Kroeze et al., 2015) to ensure that CMKLR1 activation followed the same behaviour as the original experiment, it leaves room for the possibility that the observed signal may be due to unbalanced plasmid efficiency. To address this, it is crucial to normalize future results to eliminate potential biases and ensure accurate interpretation of the data.

GPCRs' signalling pathways, relying on α subunit conformational changes, can lead to system activation without receptor-ligand interaction, known as constitutive activity (Rosenbaum et al., 2009). This poses challenges in drug discovery and gene circuit applications, as false positives and misfiring of Tango receptors can impact system efficiency and cellular control.

To a lesser extent, GPCRs have also been found to be characterized by independent internalization, a process that is usually triggered by arrestin-mediated endocytosis (Koppen and Jakobs, 2004). Unwanted inactivation of the Tango system can pose similar issues to the ones described above, especially when considering it as a component of more complex gene circuits.

Some of these limitations have been addressed in the recent study by Zeghal et al. (2023), which reported the development of a parallel Tango assay, Tango-Trio, utilizing the β -Arrestin 1, 2 and the FYVE internalization domain to distinguish between real activation, constitutive activity and precocious internalization. While this platform enhances the specificity of GPCR-ligand interaction read-out, it may not be as suitable for gene circuit applications given its size and complexity.

Nevertheless, insights from such studies provide valuable information for optimizing the Tango system's core components. For example, the original Tango receptor contained a V2 “tail”, originating from the C-terminus of the AVPR2 receptor. It was introduced into the system due to its affinity with β -Arrestin 2, increasing the recruitment and stabilization of the interaction (Kroeze et al., 2015). However, the study noted how the V2 tail can also be a source of noise by facilitating β -Arrestin recruitment without GPCR activation. Additionally, the study highlighted how the wild-type TEV protease used in the original Tango system can undergo self-cleavage, affecting its activity (Parks et al., 1995). To address this, the study used a TEV variant carrying an S219V-stop mutation, which was reported to be 100-fold more stable than the full-length protease (Kapust et al., 2001). I recognized this potential weakness during the design phase of my experiments and used a different TEV (S153N) variant, known as TEV1 Δ (Sanchez and Ting, 2020) in the construction of the BATV2 plasmid. The TEV1 Δ variant was optimized through directed evolution which was not carried out with the aim of optimization for Tango receptor activation. Consequently, the efficiency of the interaction with the V2 tail containing the TEV cleavage site is unexplored. It would be useful to compare the catalytic activity of TEV S219V and TEV1 Δ in the context of Tango activation to ensure the optimal TEV protease variant was used in my studies.

Finally, Tango-Trio highlighted the variety in arrestin interactions with some GPCRs preferentially interacting with one over the other, and sometimes with neither. This suggests that the system can still be inefficient for a portion of the GPCR family. This is an important aspect to consider when designing a larger platform for the investigation of the role of several GPCRs in the context of disorders such as AD.

Limitations of Pol III-tTA and tRNA-tTA inducible systems

The investigation of several tTA-driven inducible promoters has been crucial for understanding the feasibility and limitations of inducible sgRNA expression. The first strategy combined a Pol III promoter (H1 or U6) with TetO sequences but demonstrated that the addition of tTA silenced gene expression rather than inducing it. Combining a canonical Pol III promoter with TetO has shown that a minimal promoter is crucial for the inducibility of the system, as it requires the presence of

tTA to trigger gene expression effectively. In the canonical tetracycline response element, the minimal promoter was engineered by eliminating the upstream enhancer regions of the constitutive cytomegalovirus (CMV) Pol II promoter (Gossen and Bujard, 1992). The binding of the tTA with its HSV derived VP16 activation domains to the TetO sequences, allows full recruitment of the expression machinery. This promoter has been improved over the years through systematic mutations that have further increased its performance by decreasing background expression while increasing overall expression in combination with tTA (Li et al., 2009, Loew et al., 2010). In the H1 / U6-TetO fusions used in my studies, the H1 / U6 promoter is fully intact and ultimately able to drive sgRNA expression robustly. The inclusion of two TetO sequences, provides a means to prevent expression, with the binding of tTA interfering with the polymerase. Hence the kinetics of induction were completely opposite to my initial assumption. This is further corroborated by the use of VP16, as this transcription factor is known to specifically recruit the trans-activating domain (TAD) which is a very effective scaffold for other Pol II activation transcription factors (Hirai et al., 2010). However, the relationship between TAD and Pol III promoters is much less clear, suggesting that VP16 might not be the best transcription factor for Pol III activation (Graczyk et al., 2015). Based on this, a suitable approach could be the identification and the subsequent fusion to the TetR molecule of transcriptional activators more suited for Pol III promoters, to generate a Pol III -specific tTA.

Similar optimization of the inducible activation of a TetO-tRNA system could be considered. Previous work by Knapp et al. (2019) resulted in a suitable minimal tRNA sequence that allowed catalytic excision of the mature sgRNA from a precursor transcript driven by a Pol II promoter. However, the combination of a TRE promoter with the same tRNA mechanism failed to produce significant upregulation of a gene of interest in my experiments. Potentially, the original tRNA system as reported was not optimized to work with a TRE promoter, resulting in lack of sufficiently processed sgRNA to permit the *CD2* upregulation that constituted the read-out in my experiments. This hypothesis is further corroborated by the extensive engineering carried out by Knapp et al. (2019) to decouple the tRNA promoter activity from the tRNA:sgRNA construct as, in its original form, the presence of the promoter was required to induce the post transcriptional processing of the tRNA. Engineering of the TRE-driven tRNA:sgRNA through systemic mutations could optimize this strategy resulting in a decoupled tRNA:sgRNA regulated by a TRE promoter.

It is also important to recognize that relying solely on a tTA-inducible system may not always be the most effective mechanism within mammalian gene circuits. The basal leakiness of the original TRE promoter has been a common issue reported in the literature (Agha-Mohammadi et

al., 2004). This promoter has undergone various optimizations over the last two decades, resulting in variants such as the TIGHT version used in this project.

However, studies have shown that even the TIGHT promoter, when compared to several different versions of the CMV-TetO promoters, exhibits some level of basal expression (Loew et al., 2010). This phenomenon has also been confirmed in this study. Therefore, while inducible systems like the TIGHT promoter have been improved, its basal expression could still introduce stochasticity, affecting the overall efficiency of a circuit.

In addition, the optimization of circuit components relied upon transient transfection where many copies of the plasmids are present episomally within the cell. The leakiness that prevented conclusions on the functionality of the tested components may be due to the high copy number of episomes. Their behaviour when integrated into the genome at single copy, would have been more relevant to their ultimate role in any integrated gene circuit, but the complexity of gene targeting would have made optimization a very lengthy and resource-intensive process. In this context, it is of note that the integrated luciferase reporter used in HTLA cells did not show high levels of leakiness.

It is also worth considering an alternative iteration of the circuit in which dCas9 is regulated by the TetOFF system rather than the sgRNA. This approach offers several advantages. Firstly, the regulation of Cas9 via Tet systems is well-documented in the literature. For example, González et al. (2014) developed a platform highly relevant to this project, where Cas9 expression was driven by a TRE promoter activated by tTA, with both TRE-Cas9 and tTA integrated into the PPP1R12C safe harbor locus of human iPSCs. This study effectively induced Cas9 expression with minimal basal levels. Similarly, (Chavez et al., 2015) used a TRE-driven dCas9-VP64 and VPR in iPSCs to induce neuronal differentiation.

Beyond the tested expression of Cas9 through a Tet system, this approach could reduce cellular overload by ensuring Cas9 is expressed only when needed, compared to the less demanding sgRNA. Another potential benefit of inducible Cas9 expression is the possibility of overcoming silencing during differentiation. However, Park et al. (2022) found that a Tet-inducible Cas9 integrated into human iPSCs failed to maintain effective Cas9 expression following differentiation into hepatocytes and macrophages, suggesting that inducible expression alone may not be sufficient to prevent silencing.

Given these considerations, redesigning the circuit to incorporate an inducible Cas9 is worth exploring. This approach offers well-optimized regulation of Cas9 through a Tet system,

potentially reducing the number of optimization steps needed to complete the circuit. However, while Cas9 regulation via a TRE promoter has been shown to be more easily controlled, basal expression remains a common issue with Tet-regulated systems (González et al., 2014, Dow et al., 2015, Park et al., 2022). In the context of the circuit, basal dCas9 expression could significantly impact signal transmission and circuit efficiency. Therefore, Csy4-based regulation of sgRNA may remain the preferred method to avoid unintended basal activation of the CRISPRa system.

tTA-Csy4 inducible system is effective but flawed in the context of CLICK4

Considering the basal expression of the TIGHT promoter, exploring alternative inducible systems is essential. The Csy4 system emerged as a promising alternative due to its efficiency in cleaving the sgRNA sequence post-transcriptionally, offering potential advantages over the other strategies explored. In the context of the circuit's design, integrating the Csy4 system introduces an additional AND logic gate. This gate requires both the inducible expression of the Csy4:sgRNA molecule through the interaction of tTA with the TIGHT promoter, and the release of the sgRNA from the Csy4:sgRNA by the Csy4 ribonuclease. I hypothesized that this logic gate would have resulted in a more sophisticated control of the circuit.

However, despite showing inducible sgRNA production, my results revealed some unexpected behaviours of the Csy4 mechanism. The combination of tTA and Csy4 resulted in a lower relative expression of the gene of interest compared to the presence of Csy4 alone, corroborating the leakiness of the TRE promoter. The extent of the basal expression also made the AND gate redundant as one of the two requirements – presence of the tTA – became superfluous due to the leaky expression of the Csy4:sgRNA primary transcript. Furthermore, the integration of the Csy4 mechanism into the CLICK4 plasmid resulted in weaker upregulation of *CD2*. The causes of this phenomenon are still unclear as a degree of upregulation is still seen which matches the activation pattern in the results obtained whilst testing the individual component, albeit at a much lesser scale. A possible explanation could be that the size of the CLICK4 plasmid results in lower transfection efficiency than the single Csy4 component. This, coupled with the requirement to transcribe such a large plasmid, could result in the signal decrease that was observed.

The complexity of developing a gene circuit with a tTA-inducible primary transcript suggests the need for a potential redesign. Simplifying the gene circuit by removing the tTA induction and replacing it with Csy4 itself could offer a more elegant solution and reduce complexity.

This revised configuration, would enable the expression of the primary Csy4:sgRNA transcript using either a constitutive or cell-type specific promoter. Without the Tango receptor activation, Csy4 would not be available to process the primary transcript, thereby preventing CRISPRa-SAM induction.

Upon activation of the Tango system by the GPCR ligand, Csy4 would be liberated by the TEV protease, facilitating the excision of the sgRNA from the primary Csy4:sgRNA transcript. This process would subsequently activate CRISPRa-SAM induction of the target genes (**Fig. 7.2**).

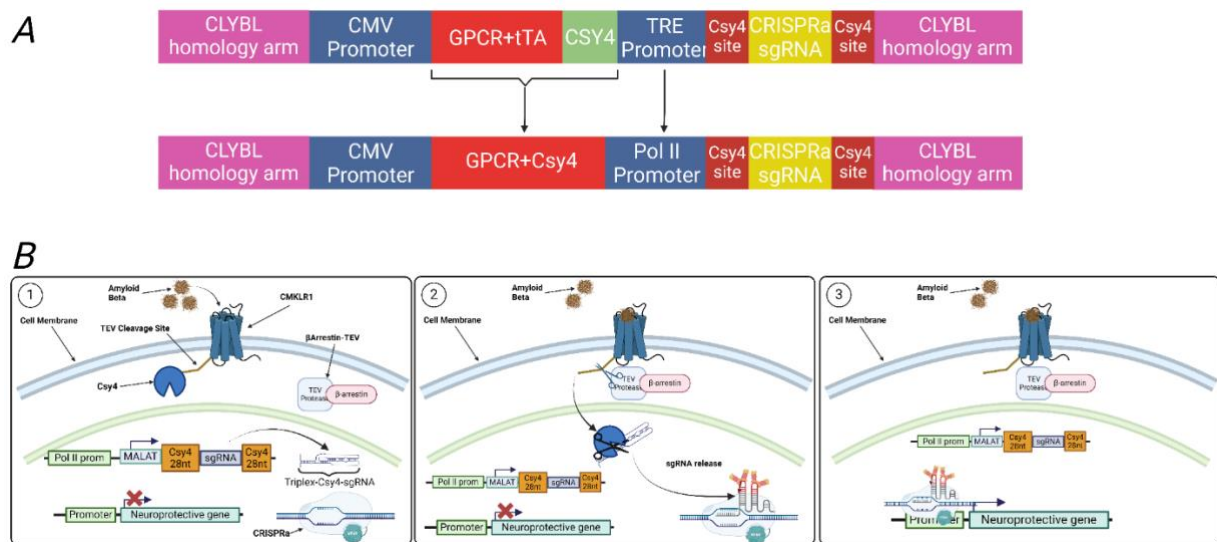


Figure 7.2 Design of gene circuit with Csy4 tethered to CMKLR1-Tango. **A)** Editing of the CLICK4 plasmid required to replace the tTA molecule for Csy4 fused to the V2 tail of the CMKLR1-Tango as well as swap of the TetO promoter for a constitutive or cell specific promoter for the expression of the Csy4:sgRNA. **B)** The function of the updated circuit. The Csy4:sgRNA is constitutively expressed but inactive as the sgRNAs are not released from this primary transcript. Csy4 RNase remains fused to the CMKLR1-Tango until the receptor is activated and the βA-TEV1Δ is triggered. Csy4 RNase cleavage of its sites within the primary Cys4:sgRNA transcript releases the sgRNA which activates the CRISPRa system to induce overexpression of the gene of interest.

This approach finds support in similar studies conducted by Baeumler et al. (2017), Kipniss et al. (2017). Kipniss et al. (2017), integrated the dCas9 from the CRISPRa-VPR system in the Tango construct and proposed a new formulation where the TEV protease was bound to the GPCR and the dCas9 to the βA (named the ChaCha system). ChaCha exhibited higher reporter gene activation and lower leakiness than the GPCR-dCas9 fusion alone. Baeumler et al. (2017) successfully integrated the CRISPRa machinery into the Tango system by fusing the two terminals of the dCas9 from the CRISPRa-SAM system to two different GPCRs to control the activation of CRISPRa through and AND gate. The study demonstrated that not only could dCas9 effectively

recombine, following the activation of the receptors, but also that the relative upregulation of the downstream gene of interest was directly proportional to the concentration of the ligands interacting with the GPCR-Tango, suggesting a potential for quantitative control.

While integrating large size molecules to the GPCR, such as Cas9, may cause unexpected problems, such as conformational changes that might alter the GPCR function, smaller molecules, such as tTA or Csy4 could be more readily fused with the Tango receptor without interfering with its activity. Thus, integrating Csy4 to the GPCR could potentially offer a more effective solution than the current design.

Borrowing from systems engineering, the reliability of a system is equal to the reliability of each component making it (Lusser's Law). Applying this concept to gene circuit engineering, it can be assumed that the elimination of a step such as the tTA-inducible sgRNA expression would ultimately strengthen the reliability of the circuit.

Human synapsin promoter is suitable to overcome AAVS1 silencing in iPSC-derived neurons.

While the complete gene circuit appeared to have been impacted by the behaviour of the CLICK4 plasmid, the BATV2 plasmid was shown to produce the expected results.

Transient transfection of BATV2 demonstrated that the hSyn promoter was able to effectively drive the expression of CRISPRa-SAM and thus the upregulation of *CD2*. However, when compared to a stronger synthetic promoter such as CAG, the level of upregulation was magnitudes lower when CRISPRa-SAM was driven by hSyn. This result was consistent with the expectation that an endogenous promoter such as hSyn would have been less effective in non-neuronal cells.

Integration of BATV2 into the *AAVS1* safe harbour locus of KOLF2-C1 cells yielded promising results. Previous studies reported significant transgene silencing when integrating ubiquitously expressed constructs into the *AAVS1* safe harbour locus during cell differentiation (Bhagwan et al., 2020, Klatt et al., 2020, Karbassi et al., 2024).

For this reason, investigating possible alternative ways of designing constructs for *AAVS1* integration could provide dual benefits. Firstly, determining methods to maintain expression of transgenes integrated into *AAVS1* throughout differentiation would expand accessible safe harbour loci, especially in developing complex gene circuits for stem cells. Secondly, cell type-

specific gene expression would introduce a further layer of control which can be useful when applied in differentiating cells.

Karbassi et al. (2024) reported the silencing of constitutively driven CRISPRa-VPR construct integrated at *AAVS1* during cardiomyocyte differentiation, and were only able to maintain minimal functionality even when employing cardiac or muscle specific promoters. Interestingly, the same study reported that the silencing was specific to dCas9-VPR, with fluorescent reporter cassette driven by the cell-specific promoters maintaining expression through differentiation. The authors conclude that the viral origin of the components of the CRISPRa-VPR system might be triggering a defence mechanism against foreign DNA. The absence of silencing in *AAVS1*-integrated reporter genes such as GFP and mCherry has been reported in other studies, which supports some of the claims regarding the variability of silencing effects (Smith et al., 2007, Ocegüera-Yanez et al., 2016, Karbassi et al., 2024). However, there are contradictory findings as well. For instance, Ordovás et al. (2015) reported the silencing of a GFP reporter gene driven by the *OCT4* promoter in the pluripotent state, which provides conflicting data on this matter.

Additionally, Klatt et al. (2020) demonstrated extensive methylation of *AAVS1* throughout differentiation, providing a potential mechanism underlying the silencing of integrated transgenes. However, they also showed that the use of myeloid-specific promoters could mitigate transgene silencing and allow for expression in the myeloid lineage during differentiation. Similarly, in my studies, hSyn was effective in inducing neuron-specific expression of CRISPRa-SAM when integrated into the *AAVS1* safe harbours locus.

Further investigation into the silencing mechanism of *AAVS1* is warranted to understand why the CRISPR machinery is more prone to silencing compared to other transgenes, and whether *AAVS1* is more susceptible than other genomic regions. One possible explanation is that multiple factors determine the degree of silencing. For instance, the size and nature of the transgene may play a role. CRISPR sequences, comprising large intronless viral sequences, could potentially trigger the human silencing hub (HUSH) complex (Seczynska et al., 2022). Additionally, given that *AAVS1* was originally found as a preferred integration site for adeno-associated viruses (Ramachandra et al., 2011), it may be more susceptible to DNA defence mechanisms when large exogenous sequences are integrated.

A potential approach to explore this hypothesis would involve integrating progressively larger viral and endogenous sequences to determine whether silencing is dependent on the length of the construct. Similarly, investigating the involvement of the HUSH complex could entail performing qPCR on genes expressed as part of the complex, such as *TASOR*, *MPP8*, and *periphilin*, to

ascertain if their increased expression correlates with transgene silencing (Seczynska et al., 2022).

It is also worth noting that a completely different strategy to avoid *AAVS1* silencing would be the integration of the CRISPR machinery into the *CLYBL* safe harbour locus. The Kampmann lab has demonstrated that this is a viable alternative as they generated both CRISPRi and CRISPRa expressing iPSC lines by targeting *CLYBL* (Tian et al., 2019, Tian et al., 2021). In both studies, the CRISPRi and CRISPRa machinery were driven by the CAG promoter and integrated into the *CLYBL* locus at the pluripotent state but shown to maintain expression throughout neuronal differentiation. Furthermore, in both cell lines, the *AAVS1* locus was used for the integration of doxycycline-inducible *NGN2* overexpression cassette, utilised to expedite the neuronal differentiation (Zhang et al., 2013). While these cell lines suggest that *CLYBL* is not affected by the same silencing issues, Karbassi et al. (2024) compared a CAG-driven mCherry construct to a CAG-CRISPRa-mCherry construct integrated into the *CLYBL* locus and noted that the latter was expressed much more weakly than the former. Even with weaker expression, the CAG-CRISPRa-mCherry construct integrated in *CLYBL* was able to induce upregulation of the target gene.

However, the ability to regulate CRISPRa expression in specific cell types can be a valuable tool. This study appears to be the first to illustrate that the hSyn promoter can drive neuron-specific expression of transgenes integrated into the *AAVS1* locus of pluripotent stem cells. Regulating transgene expression based on specific cell types can help reduce metabolic stress during cell differentiation, ensuring that transgenes are activated only when necessary. Additionally, cell-specific transgene activation can enhance the safety of cell therapy by preventing transgene expression if the cell line fails to differentiate correctly. Investigating whether silencing in the *AAVS1* locus is directly linked to the differentiation process would be of interest. Creating a promoter library of constructs integrated at *AAVS1* and examining the behaviour of expression across various cell types would broaden the utility of this system.

Despite the successful generation of KOLFBAT, it's important to acknowledge certain limitations of this study. Firstly, only two independent iPSC clones were characterised. To minimize potential clone-dependent variations, confirming hSyn-driven CRISPRa expression in iPSC-derived neurons across more independent clones is advisable. Secondly, no methylation analysis was conducted in the project. Access to both KOLFSAM and KOLFBAT, with different promoters integrated, presents an opportunity to compare their methylation levels at the *AAVS1* locus.

In addition, while the expression of the β A-TEV1 Δ by the hSyn promoter was found to be effective when testing the BATV2 plasmid in HEK293T cells, basal expression of TIGHT-GFP was later

established to be extensive, suggesting that the results of the β A-TEV1 Δ experiment may not accurately reflect hSyn-driven expression of β A-TEV1 Δ . To address this, a new test using a more reliable reporter plasmid, such as the TetOFF[®] TRE-TIGHT-Luciferase vector, the main control vector provided by Takara with their TetOFF-inducible kit, could provide quantitative results through a luciferase fluorescence assay (Takara [™]). One added complication could be that the hSyn promoter used to drive expression of the β A-TEV1 Δ is not particularly strong in non-neuronal cells, such as the HEK293T cells used in the optimization steps. Thus, it could be crucial to investigate the functionality of the components of the gene circuit in neuronal cells. Finally, if the β A-TEV1 Δ is confirmed not to perform as anticipated, validating the TEVS219V variant used by Kapust et al. (2001) could be beneficial. Furthermore, while hSyn-driven β A-TEV1 Δ was also included in the transgene sequence, its expression across differentiation was not tested. Investigating the behaviour of β A-TEV1 Δ at *AAVS1* could shed light on the silencing of this region, considering that the two different expression cassettes of CRISPRa-SAM and β A-TEV1 Δ may be differentially affected by any potential silencing mechanisms during differentiation.

Finally, previous studies confirmed that exogenous promoters are able to drive the expression of transgenes integrated into the *CLYBL* locus of iPSC, throughout neuronal differentiation, providing support for the choice to employ the CMV to drive the expression of the CMKLR1-Tango and Csy4:sgRNA in the CLICK4 plasmid (Tian et al., 2019, Tian et al., 2021, Karbassi et al., 2024). However, it would be important to consider the replacement of the CMV promoter in the CLICK4 plasmid for an endogenous promoter such as hSyn to ensure that every component of the circuit is only expressed in the appropriate cell type.

Nevertheless, the application of hSyn to drive CRISPRa-SAM in iPSC-derived neurons proved to be an effective approach for testing targeted gene upregulation in a disease-related model, as demonstrated by the upregulation of *MSX2*.

***MSX2* upregulation KOLFBAT-derived neurons**

As proposed by Kampmann (2017), current methods to investigate the pathways involved in cellular neurodegeneration have mostly relied on top-down proteomics and transcriptomics. However, these approaches have struggled to establish the causality of the genes involved in neuronal vulnerability and degeneration. iPSC-based neuronal models capable of selectively activating or silencing targeted genes can offer valuable insights into the genes involved in neurodegeneration and neuroprotection, thereby expanding our current understanding of potential drug targets.

To explore the potential application of the iPSC-based neuronal model capable of selectively expressing CRISPRa-SAM in neurons, I chose to upregulate *MSX2*, a transcription factor involved in several developmental pathways (Yu et al., 2019, Kwang et al., 2002, Chen et al., 2007). Additionally, a recent study by the Buckley lab demonstrated that *clusterin*-KO, one of the most common risk factor genes for AD development, resulted in iPSC-derived neurons better able to withstand A β -induced stress, suggesting that elimination of this gene leads to a neuroprotective phenotype (Robbins et al., 2018). Unpublished data from the same study revealed that *MSX2* was the most upregulated gene in the *clusterin*-KO neurons following A β -induced stress, indicating a potential relationship between *MSX2* and the neuroprotective phenotype.

I transduced *MSX2* sgRNA via lentiviral transduction into day-50 KOLFBAT-derived neurons. The cells were allowed to express and translate *MSX2* for 48 hours and 10 days. RNA sequencing was performed on the *MSX2*-transduced neurons and compared to non-targeting control transduced cells at both time points.

The results revealed an increase in differentially expressed genes from the 48-hour to the 10-day datasets, suggesting that *MSX2* was able to perturb the transcriptome over time. This observation was supported by the significant upregulation of *MSX2*, which was the second and third most upregulated gene at 48 hours and 10 days, respectively. Additionally, genes such as *ID1*, *HMOX*, and *HHIP* showed strong differential expression and were associated with *MSX2*, further indicating that *MSX2* effectively perturbed the transcriptome.

Gene set enrichment analysis was conducted to investigate the most enriched pathways in the 48-hours and 10-days datasets. In both cases, synaptic development terms were found to be the most enriched, while terms associated with DNA replication were downregulated. As discussed previously, the TGF- β signalling pathway and the BMP family associated with it are central regulators of neuronal differentiation (Blank and Karlsson, 2011). *MSX2* has been shown to play a role in several developmental pathways, especially in relation to TGF- β (Lowery and Rosen, 2018). Moreover, *MSX2* has been demonstrated to induce the suppression of *SOX2* thus destabilising pluripotency (Praharaj et al., 2023). Based on this, it is possible that its overexpression can lead to accelerated neuronal maturation, as suggested by the enrichment of synaptic development pathways and the downregulation of DNA replication pathways.

While the enriched pathways provide some insights into the impact of *MSX2* on neuronal development, it's important to acknowledge certain limitations within this study. Firstly, *MSX2* protein quantification was not performed which could have confirmed the effective protein consequences of *MSX2* transcriptional upregulation. Western blot analysis at different time

points could also offer valuable insights into the optimal timing for transcriptomics analysis, aligning with peak *MSX2* protein translation.

Additionally, both differential analyses identified a large number of differentially expressed genes, potentially biasing the gene set enrichment results. This phenomenon could have several underlying causes.

First, *MSX2*'s inherent role in development and differentiation (Kwang et al., 2002, Chen et al., 2007, Han et al., 2007) could be a key factor. Overexpression of *MSX2* in differentiating neurons is likely to cause widespread transcriptome perturbation, leading to the upregulation and downregulation of numerous genes. Studies in yeast have shown that overexpression of transcription factors can result in significant transcriptional changes (Chua et al., 2006, Hackett et al., 2020), supporting this hypothesis. Moreover, the increase in perturbations from the 48-hour to the 10-day time point suggests a cascade effect triggered by *MSX2* overexpression.

However, while *MSX2*'s role as a developmental regulator may explain the number of differentially expressed genes, it does not account for the dramatic LogFold2 changes and extremely low p-values observed. Another possible factor could be the strength of the CRISPRa system used. The human synapsin promoter has been shown to be effective not only for its neuronal specificity but also for its strong expression of exogenous genes (Kügler et al., 2003, McLean et al., 2014). The combination of the human synapsin promoter, the upregulating effect of CRISPRa, and the integration of the system into both copies of *AAVS1* could explain the extreme transcriptome perturbation observed in this study. This explanation could also be supported once again by the increasing number of differentially expressed genes from the 48-hour to the 10day time point, suggesting an additive effect of CRISPRa activation of *MSX2*.

It also highlights two key limitations of the system. First, the only modular component is the targeting sgRNA, which restricts the ability to modify signal strength unless a heterozygous cell line is used, or the system is entirely redesigned. Second, while the integration of the CRISPRa system into a safe harbour locus and the use of the human synapsin promoter were intended to emulate more physiologically relevant behaviour, the extreme transcriptome perturbation observed suggests that this approach may not have achieved the desired physiological relevance. This challenge was compounded by the system's lack of modularity, which prevented fine-tuning of signal strength, such as by reducing CRISPRa transduction.

While these explanations might account for the number and magnitude of differentially expressed genes, it is important to note that the Koppermann lab developed a similar system

where CRISPRa is expressed in iPSC-derived neurons via a CAG promoter integrated into the CLYBL plasmid (Tian et al., 2019, Tian et al., 2021, Leng et al., 2022, Dräger et al., 2022). Despite using a comparable approach by transducing neurons with sgRNA and assessing their genome-wide effects, these studies did not report the same level of impact observed in this study. This discrepancy could be attributed to lower CRISPRa expression efficiency at the CLYBL safe harbour locus (Karbassi et al., 2024), compared to the human synapsin promoter in neurons. However, technical issues may also play a role, such as the use of technical replicates that might amplify natural differences between experimental groups. To address this, biological replicates could be employed—replicating the experiment three times and analysing one triplicate each time to ensure sufficient distinction between sets. Additionally, introducing new KOLFBAT clones for differentiation could introduce variability within the triplicate group.

Despite the challenges posed by the number of differentially expressed genes, the expression patterns of *MSX2* and its associated genes remain clear. Further investigation into the role of these genes in AD neurodegeneration could be valuable. For example, haemoglobin has been associated with AD, with downregulation of intracellular haemoglobin being reported in both AD patients and animal models (Ferrer et al., 2011, Orre et al., 2014). Other studies reported neurotoxicity of extracellular haemoglobin in AD also corroborated by the avidity between A β and haemoglobin (Stone, 2008). This discrepancy in findings led to the hypothesis that intracellular haemoglobin might be beneficial to counteract potential hypoxia occurring during neurodegeneration, but extracellular haemoglobin might exacerbate the disease (Altinoz et al., 2019). Similarly, *PLA2G3*, the most downregulated gene in the 10-day dataset, has been strongly linked to oxidative stress in AD. Additional investigation using the KOLFBAT-derived neurons would be beneficial to further elucidate the role of *HBB* and *PLA2G3* in neurodegeneration and in relation to *MSX2*.

Lastly, while transcriptomics provides valuable insights, it's not the sole method for investigating the role of genes in neurodegenerative disorders. Neuroprotection in *CLU*-KO neurons through exposure to A β_{25} and A β_{42} , was assessed via cell survival and neurite length (Robbins et al., 2018). To further explore the potential role of *MSX2* in neuroprotection, KOLFBAT-derived neurons could be transduced with *MSX2* sgRNA and exposed to A β , followed by survival assays to evaluate whether its upregulation yields phenotypic benefits, possibly compared to the *CLU*-KO neurons already generated by the Buckley lab.

Based on this, the study highlights the utility of CRISPR-based iPSC-derived neurons as valuable models for dissecting genetic mechanisms underlying neurodegeneration and neuroprotection,

offering promising avenues for investigating potential drug targets in neurodegenerative disorders.

The circuit as a model for Alzheimer's and further therapeutic applications

Throughout this project, my goal was to explore the feasibility of engineering iPSC-derived neurons equipped with a gene circuit capable of sensing and responding to specific disease biomarkers, such as extracellular A β . Although the final design and integration of the circuit were not successful, the potential future applications of such engineered cell types are worth considering.

Degenerative disorders like Alzheimer's disease are characterized by a progressively deteriorating cellular environment, leading to a decline in cognitive function and ultimately, death (Knopman et al., 2021). Therapeutics for such disorders may need to target different stages or conditions of the disease and adapt to its evolving nature. In this context, therapeutic cell engineering has emerged as a promising field over the past decade (Mansouri and Fussenegger, 2022). By enabling cells to sense external stimuli and activate gene circuits, we can develop more targeted and personalized treatments (Kitada et al., 2018). A prominent example is the development of CAR-T cells, which have revolutionized cancer treatment (Lim and June, 2017, Bashor et al., 2022).

In the context of Alzheimer's disease, engineered neurons could be designed to dynamically respond to neurodegeneration by sensing exogenous A β and upregulating the expression of neuroprotective genes. For such therapy to be effective, identifying the right neuroprotective genes is crucial. One promising candidate is the transcriptional regulator REST (restrictive element-1 silencing transcription factor), which has been shown to protect against neurodegeneration by repressing apoptotic genes commonly upregulated during AD progression (Lu et al., 2014, Hwang and Zukin, 2018). While REST may protect the engineered neurons, it may not offer benefits to the broader neuronal system of the patient.

Another candidate is nerve growth factor (NGF), a neurotrophic factor known for promoting neurite growth and maintaining neuronal integrity (Aloe et al., 2015). Unlike REST, NGF is produced and released by various cell types, including neurons, making it a suitable candidate for improving the survival of both engineered neurons and neighbouring cells (Huang and Reichardt, 2001, McAllister, 2001). However, it is important to acknowledge the wide range of roles NGF plays in neuronal development and regulation, as its overexpression could potentially have detrimental effects (Aloe et al., 2012).

A final candidate is neprilysin, an enzyme known for its ability to cleave A β (Marr et al., 2003, Iijima-Ando et al., 2008, Takamatsu et al., 2014). Neprilysin's active site is located in the extracellular space, primarily along axons and synapses, where A β degradation occurs (Grimm et al., 2013). An added benefit of neprilysin is its potential to create a self-regulating feedback loop. Engineered neurons could overexpress neprilysin in response to extracellular A β , reducing A β levels and, in turn, modulating neprilysin production.

To enhance the mobility of therapeutic cells and their overall impact on the brain, iPSCs fitted with such gene circuits could be differentiated into various cell types. Microglia, with their dynamic and immune roles in the brain, present a promising option (Lannes et al., 2017). Microglia differentiation protocols are now well-established, eliminating the need for extensive optimization (Lanfer et al., 2022). However, as synapsin does not appear to be expressed in microglia, the gene circuit would need to be modified by substituting the hSyn promoter with a microglia-specific promoter. Several microglia-specific promoters have already been identified. For example, a sequence upstream of the start codon for the exon 1 of the ionized calcium-binding adaptor molecule 1 (IBA1) has been shown to be a strong and specific promoter for microglia (Okada et al., 2022). Similarly, a minimal promoter for the hexosaminidase subunit beta (HEXB) has been found to be highly active in microglia but not in other CNS-specific cell types (Shah et al., 2022).

The combination of neuroprotective genes and alternative cell types presents a promising avenue for potential therapeutic applications. However, beyond the therapeutic strategy itself, the method of delivery is a critical factor that requires careful consideration. Two primary approaches exist for delivering complex gene circuits: gene therapy and cell therapy.

One strategy is to deliver only the gene circuit, which reduces the payload to DNA. Viral delivery has long been considered one of the most effective methods for gene delivery (Mali, 2013). When delivering the gene circuit, three key factors must be considered: the size of the payload, the gene integration method, and cell targeting—particularly the challenge of crossing the blood-brain barrier (BBB). Targeting neurons and achieving delivery across the BBB has historically been difficult. However, recent studies have made significant progress. For instance, Goertsen et al. (2022) identified adeno-associated viruses (AAVs) naturally enriched in the brain, successfully using them for gene delivery in the CNS. Similarly, Huang et al. (2024) developed an AAV that crosses the BBB by targeting the transferrin receptor, which is commonly expressed on the BBB.

These studies suggest that gene delivery to the CNS is viable, but the gene circuit under consideration is approximately 30 Kb in size. Given that the maximum AAV payload is around 5 Kb

(Wu et al., 2010), AAVs are unsuitable for delivering this circuit. Lentiviruses, with a payload capacity of about 10 Kb (Bauler et al., 2020), could allow the circuit to be split into three constructs. However, the increased complexity of delivering multiple constructs and ensuring all components integrate into the same cells makes gene therapy a less effective strategy for this application.

Given the limitations of gene therapy, a cell therapy approach may be more effective. In this scenario, patient cells could be reprogrammed, fitted with the gene circuit, and then delivered into the brain. The CNS is notoriously difficult to target due to the BBB, which limits the access of exogenous molecules (Wu et al., 2023). Interestingly, Fujiwara et al. (2013) demonstrated that surgically transplanting human iPSC-derived neuronal precursor cells improved spatial memory in APP mice models. Similar methods have been shown to be effective in mice, although they tend to be highly invasive (Eckert et al., 2015, Hayashi et al., 2020b).

Recently, intranasal delivery of stem cells has emerged as a less invasive alternative, allowing for repeated dosing (Zhang et al., 2021). Intranasal delivery of mesenchymal cells, first described by Danielyan et al. (2009), relies on the migration of cells from the nasal mucosa into the brain via the olfactory neural pathway, thereby bypassing the BBB. Neural stem cells have been effectively delivered intranasally (Zhang et al., 2021), alleviating symptoms of brain injury and Parkinson's disease in mice (Li et al., 2017, Domínguez-García et al., 2020). In Alzheimer's disease models, intranasal injection of bone marrow mesenchymal cells in APP/PS1 mice not only allowed these cells to migrate to areas of high A β concentration but also suggested their involvement in A β phagocytosis (Danielyan et al., 2014). Similarly, olfactory ecto-mesenchymal cells delivered intranasally improved symptoms of spatial learning and memory deficits in AD rats (Valipour et al., 2024).

While intranasal delivery shows promise, there are important limitations to consider (Villar-Gómez et al., 2022). First, studies indicate that only a fraction of the delivered cells reach the CNS, with most remaining in the nasal turbinates (Galeano et al., 2018). Additionally, many of these studies have been conducted on rodents, whose olfactory systems differ significantly from humans, potentially complicating the translation of these results to human therapy (Villar-Gómez et al., 2022). Notably, human clinical trials are already underway to evaluate the efficacy of intranasal delivery in Alzheimer's patients, with results expected by mid-2025 (ClinicalTrials.gov; Identifiers: [NCT03724136](#), [NCT02795052](#)).

Given the size of the gene circuit and the challenges associated with its delivery, a cell therapy approach coupled with intranasal delivery appears to be a strategy worth serious consideration.

This method could offer a less invasive and potentially more effective way to introduce engineered cells into the CNS for therapeutic purposes.

While the circuit was designed with the hope of providing therapeutic benefits, the complexity of the system, combined with the challenges related to its delivery and application, suggests that alternative uses should be considered. In this regard, the KOLFSAM iPSC line, which expresses the CRISPRa system driven by the hSyn promoter, has shown promise as a tool for investigating transcriptome perturbations in neurons.

A similar approach has been successfully employed by the Kampmann laboratory, which developed an iPSC line where the CRISPRa system, driven by a CAG promoter, was integrated into the CLYBL safe harbour locus. This model was used to explore the role of single genes in neurodegenerative disorders (Tian et al., 2019, Tian et al., 2021, Dräger et al., 2022). While iPSCs have rapidly become a central tool for modelling neurodegeneration, it is important to recognize their limitations. A key challenge is the inability of iPSC-derived neurons to fully mimic the aged neuronal cells found in Alzheimer's disease (AD) patients. Even fully developed iPSC neurons struggle to replicate the age-related aspects of neurodegeneration (Mishra et al., 2023). A recent overexpression of genes associated with aging have been investigated as a possible strategy to develop neurons that better represent the aged neurons. Recent research has explored the overexpression of aging-associated genes, such as synuclein alpha and LMNA, as a strategy to develop neurons that better represent aged cells (Miller et al., 2013, Tagliafierro et al., 2019). Although promising, these approaches have not yet been extensively studied in the context of Alzheimer's (Mishra et al., 2023).

An additional concern with iPSC models is that they fail to recapitulate the complex intercellular interaction that characterize AD. A possible solution to this problem is the use of cerebral organoids to better recapitulate the cellular heterogeneity in AD (Amin and Paşca, 2018). This strategy has proven to not only provide some useful recapitulation of AD but also the combination of different cell types, such as astrocytes and neurons, increase neural maturation (Choi et al., 2014, Park et al., 2018).

However, while organoids address some of the limitations of single-cell-type iPSC models, they are not without their own challenges. Organoids face technical issues such as batch-to-batch variability, reduced longevity, and central necrosis (Mishra et al., 2023). A significant challenge is the lack of vascularization, which can affect the natural development of organoids. Co-culturing neuronal cells with epithelial cells has been explored as a potential solution, with epithelial cells penetrating the organoid but failing to form complete blood vessels (Pham et al., 2018). More

recently, Cakir et al. (2019) made progress in this area by inducing a better structure of endothelial cells, resembling the blood-brain barrier, through the overexpression of the ETV2 transcription factor. Although the study was unable to fully control the developing structure, it represents a significant step forward in overcoming some of the limitations of organoids and enhancing their relevance as models.

Despite the limitations of iPSCs and organoids, they offer significant advantages over animal models in recapitulating the pathophysiology of neurodegeneration. These models provide opportunities for genetic manipulation and high-throughput analysis, making them valuable tools in the drug discovery process for AD (Dawson et al., 2018, Mishra et al., 2023). The rapid advancements in this field also suggest that many of the current limitations of iPSC-based models may be surmountable in the near future, further solidifying their importance in neurodegeneration research.

Conclusion and Future directions

Although the completed gene circuit was not fully effective, the investigation of its individual components provided valuable insights in the design of circuits and their assembly within mammalian cells. Going forward, the integration of the Csy4 ribonuclease into the CMKLR1-Tango receptor could potentially complete the circuit and eliminate the need for a tTA-inducible sgRNA promoter. Streamlining the signal transduction process could enhance the final output of the signal and would hopefully secure the viability of the circuit. Additionally, integrating the BATV2 sequence into the KOLF2-C1 cell line provided valuable insights into the silencing phenomenon affecting transgenes integrated into the *AAVS1* locus. The mechanisms behind this silencing, particularly its heightened effect on CRISPR and the role of methylation or silencing complexes, remains unclear. Further investigation, comparing cell lines such as KOLFBAT and KOLFSAM, could shed light on these mechanisms and provide a more comprehensive understanding of transgene silencing in stem cells, especially during differentiation. Furthermore, investigating the upregulation of *MSX2* using the KOLFBAT cell line hinted at a potential anti-inflammatory role, which could be crucial in investigating neurodegenerative diseases like Alzheimer's (Balusu et al., 2023). Physiological studies validating these results would offer valuable insights.

Based on this project and corroborated by the studies from the Kampmann lab, the development of iPSC-based platforms capable of regulated and targeted gene expression can expand the range of applications for genetic screenings in human neurons, aiding our understanding of pathogenic mechanisms and therapeutic targets in multifactorial disorders such as neurodegeneration.

Whilst the field of synthetic biology and mammalian gene circuits is still in its infancy, the integration of this CRISPRa component into synthetic gene circuits within iPSCs holds promise for further innovative applications.

Bibliography

- AIZENSTEIN, H. J., NEBES, R. D., SAXTON, J. A., PRICE, J. C., MATHIS, C. A., TSOPELAS, N. D., ZIOLKO, S. K., JAMES, J. A., SNITZ, B. E., HOUCK, P. R., BI, W., COHEN, A. D., LOPRESTI, B. J., DEKOSKY, S. T., HALLIGAN, E. M. & KLUNK, W. E. 2008. Frequent Amyloid Deposition Without Significant Cognitive Impairment Among the Elderly. *Archives of Neurology*, 65, 1509-1517.
- ALOE, L., ROCCO, M. L., BALZAMINO, B. O. & MICERA, A. 2015. Nerve Growth Factor: A Focus on Neuroscience and Therapy. *Curr Neuroparmacol*, 13, 294-303.
- ALOE, L., ROCCO, M. L., BIANCHI, P. & MANNI, L. 2012. Nerve growth factor: from the early discoveries to the potential clinical use. *Journal of Translational Medicine*, 10, 239.
- ALONSO, A. D. C., GRUNDKE-IQBAL, I. & IQBAL, K. 1996. Alzheimer's disease hyperphosphorylated tau sequesters normal tau into tangles of filaments and disassembles microtubules. *Nature Medicine*, 2, 783-787.
- ALTINOZ, M. A., GULOKSUZ, S., SCHMIDT-KASTNER, R., KENIS, G., INCE, B. & RUTTEN, B. P. F. 2019. Involvement of hemoglobins in the pathophysiology of Alzheimer's disease. *Experimental Gerontology*, 126, 110680.
- AMIN, N. D. & PAŞCA, S. P. 2018. Building Models of Brain Disorders with Three-Dimensional Organoids. *Neuron*, 100, 389-405.
- ASHBURNER, M., BALL, C. A., BLAKE, J. A., BOTSTEIN, D., BUTLER, H., CHERRY, J. M., DAVIS, A. P., DOLINSKI, K., DWIGHT, S. S., EPPIG, J. T., HARRIS, M. A., HILL, D. P., ISSEL-TARVER, L., KASARSKIS, A., LEWIS, S., MATESE, J. C., RICHARDSON, J. E., RINGWALD, M., RUBIN, G. M. & SHERLOCK, G. 2000. Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25, 25-29.
- AUBREY, BRANDON J., KELLY, GEMMA L., KUEH, ANDREW J., BRENNAN, MARGS S., O'CONNOR, L., MILLA, L., WILCOX, S., TAI, L., STRASSER, A. & HEROLD, MARCO J. 2015. An Inducible Lentiviral Guide RNA Platform Enables the Identification of Tumor-Essential Genes and Tumor-Promoting Mutations In Vivo. *Cell Reports*, 10, 1422-1432.
- AUSLÄNDER, D., EGGERSCHWILER, B., KEMMER, C., GEERING, B., AUSLÄNDER, S. & FUSSENEGGER, M. 2014. A designer cell-based histamine-specific human allergy profiler. *Nature Communications*, 5, 4408.
- AUSLÄNDER, S., AUSLÄNDER, D., MÜLLER, M., WIELAND, M. & FUSSENEGGER, M. 2012. Programmable single-cell mammalian biocomputers. *Nature*, 487, 123-127.
- AUSLANDER, S. & FUSSENEGGER, M. 2016. Engineering Gene Circuits for Mammalian Cell-Based Applications. *Cold Spring Harb Perspect Biol*, 8.
- BAEUMLER, T. A., AHMED, A. A. & FULGA, T. A. 2017. Engineering Synthetic Signaling Pathways with Programmable dCas9-Based Chimeric Receptors. *Cell Rep*, 20, 2639-2653.
- BALUSU, S., HORRÉ, K., THRUPP, N., CRAESSAERTS, K., SNELLINX, A., SERNEELS, L., T'SYEN, D., CHRYSIDOU, I., ARRANZ, A. M., SIERKSMA, A., SIMRÉN, J., KARIKARI, T. K., ZETTERBERG, H., CHEN, W.-T., THAL, D. R., SALTA, E., FIERIS, M. & DE STROOPER, B. 2023. MEG3 activates necroptosis in human neuron xenografts modeling Alzheimer's disease. *Science*, 381, 1176-1182.
- BARNEA, G., STRAPPS, W., HERRADA, G., BERMAN, Y., ONG, J., KLOSS, B., AXEL, R. & LEE, K. J. 2008. The genetic design of signaling cascades to record receptor activation. *Proceedings of the National Academy of Sciences*, 105, 64.
- BASHOR, C. J., HILTON, I. B., BANDUKWALA, H., SMITH, D. M. & VEISEH, O. 2022. Engineering the next generation of cell-based therapeutics. *Nature Reviews Drug Discovery*, 21, 655-675.

- BAULER, M., ROBERTS, J. K., WU, C.-C., FAN, B., FERRARA, F., YIP, B. H., DIAO, S., KIM, Y.-I., MOORE, J., ZHOU, S., WIELGOSZ, M. M., RYU, B. & THROM, R. E. 2020. Production of Lentiviral Vectors Using Suspension Cells Grown in Serum-free Media. *Molecular Therapy Methods & Clinical Development*, 17, 58-68.
- BECIROVIC, E. 2022. Maybe you can turn me on: CRISPRa-based strategies for therapeutic applications. *Cell Mol Life Sci*, 79, 130.
- BECSKEI, A. & SERRANO, L. 2000. Engineering stability in gene networks by autoregulation. *Nature*, 405, 590-593.
- BELL, R. D., SAGARE, A. P., FRIEDMAN, A. E., BEDI, G. S., HOLTZMAN, D. M., DEANE, R. & ZLOKOVIC, B. V. 2007. Transport Pathways for Clearance of Human Alzheimer's Amyloid β -Peptide and Apolipoproteins E and J in the Mouse Central Nervous System. *Journal of Cerebral Blood Flow & Metabolism*, 27, 909-918.
- BENDIXEN, L., JENSEN, T. I. & BAK, R. O. 2023. CRISPR-Cas-mediated transcriptional modulation: The therapeutic promises of CRISPRa and CRISPRi. *Molecular Therapy*, 31, 1920-1937.
- BENNETT, D. A., SCHNEIDER, J. A., ARVANITAKIS, Z., KELLY, J. F., AGGARWAL, N. T., SHAH, R. C. & WILSON, R. S. 2006. Neuropathology of older persons without cognitive impairment from two community-based studies. *Neurology*, 66, 1837-44.
- BHAGWAN, J., COLLINS, E., MOSQUEIRA, D., BAKAR, M., JOHNSON, B., THOMPSON, A., SMITH, J. & DENNING, C. 2020. Variable expression and silencing of CRISPR-Cas9 targeted transgenes identifies the AAVS1 locus as not an entirely safe harbour [version 2; peer review: 2 approved, 1 not approved]. *F1000Research*, 8.
- BINDER, C., CVETKOVSKI, F., SELLBERG, F., BERG, S., VISBAL, H. P., SACHS, D. H., BERGLUND, E. & BERGLUND, D. 2020. CD2 Immunobiology. *Frontiers in Immunology*, 11, 1090.
- BLANK, U. & KARLSSON, S. 2011. The role of Smad signaling in hematopoiesis and translational hematology. *Leukemia*, 25, 1379-1388.
- BLOOM, G. S. 2014. Amyloid- β and Tau: The Trigger and Bullet in Alzheimer Disease Pathogenesis. *JAMA Neurology*, 71, 505-508.
- BOHLEN, C. J., BENNETT, F. C., TUCKER, A. F., COLLINS, H. Y., MULINYAWE, S. B. & BARRES, B. A. 2017. Diverse Requirements for Microglial Survival, Specification, and Function Revealed by Defined-Medium Cultures. *Neuron*, 94, 759-773.e8.
- BÖHM, S., SPLITH, V., RIEDMAYR, L. M., RÖTZER, R. D., GASPARONI, G., NORDSTRÖM, K. J. V., WAGNER, J. E., HINRICHSMEYER, K. S., WALTER, J., WAHL-SCHOTT, C., FENSKE, S., BIEL, M., MICHALAKIS, S. & BECIROVIC, E. 2020. A gene therapy for inherited blindness using dCas9-VPR-mediated transcriptional activation. *Sci Adv*, 6, eaba5614.
- BRAAK, H., ALAFUZOFF, I., ARZBERGER, T., KRETZSCHMAR, H. & DEL TREDICI, K. 2006. Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. *Acta Neuropathol*, 112, 389-404.
- BROPHY, J. A. N. & VOIGT, C. A. 2014. Principles of genetic circuit design. *Nature methods*, 11, 508-520.
- CAKIR, B., XIANG, Y., TANAKA, Y., KURAL, M. H., PARENT, M., KANG, Y.-J., CHAPETON, K., PATTERSON, B., YUAN, Y., HE, C.-S., RAREDON, M. S. B., DENGELI, J., KIM, K.-Y., SUN, P., ZHONG, M., LEE, S., PATRA, P., HYDER, F., NIKLASON, L. E., LEE, S.-H., YOON, Y.-S. & PARK, I.-H. 2019. Engineering of human brain organoids with a functional vascular-like system. *Nature Methods*, 16, 1169-1175.
- CALLAERTS, P., HALDER, G. & GEHRING, W. J. 1997. PAX-6 IN DEVELOPMENT AND EVOLUTION. *Annual Review of Neuroscience*, 20, 483-532.
- CAO, J., WU, L., ZHANG, S. M., LU, M., CHEUNG, W. K., CAI, W., GALE, M., XU, Q. & YAN, Q. 2016. An easy and efficient inducible CRISPR/Cas9 platform with improved specificity for multiple gene targeting. *Nucleic Acids Res*, 44, e149.

- CASCELLA, R., CONTI, S., TATINI, F., EVANGELISTI, E., SCARTABELLI, T., CASAMENTI, F., WILSON, M. R., CHITI, F. & CECCHI, C. 2013. Extracellular chaperones prevent A β 42-induced toxicity in rat brains. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1832, 1217-1226.
- CASTAÑO, J., BUENO, C., JIMÉNEZ-DELGADO, S., ROCA-HO, H., FRAGA, M. F., FERNANDEZ, A. F., NAKANISHI, M., TORRES-RUIZ, R., RODRÍGUEZ-PERALES, S. & MENÉNDEZ, P. 2017. Generation and characterization of a human iPSC cell line expressing inducible Cas9 in the “safe harbor” AAVS1 locus. *Stem Cell Research*, 21, 137-140.
- CERBINI, T., FUNAHASHI, R., LUO, Y., LIU, C., PARK, K., RAO, M., MALIK, N. & ZOU, J. 2015a. Transcription activator-like effector nuclease (TALEN)-mediated CLYBL targeting enables enhanced transgene expression and one-step generation of dual reporter human induced pluripotent stem cell (iPSC) and neural stem cell (NSC) lines. *PLoS One*, 10, e0116032.
- CERBINI, T., LUO, Y., RAO, M. S. & ZOU, J. 2015b. Transfection, selection, and colony-picking of human induced pluripotent stem cells TALEN-targeted with a GFP gene into the AAVS1 safe harbor. *Journal of visualized experiments : JoVE*, 52504.
- CHA, M.-Y., KWON, Y.-W., AHN, H.-S., JEONG, H., LEE, Y. Y., MOON, M., BAIK, S. H., KIM, D. K., SONG, H., YI, E. C., HWANG, D., KIM, H.-S. & MOOK-JUNG, I. 2016. Protein-Induced Pluripotent Stem Cells Ameliorate Cognitive Dysfunction and Reduce A β Deposition in a Mouse Model of Alzheimer's Disease. *Stem Cells Translational Medicine*, 6, 293-305.
- CHAMBERS, S. M., FASANO, C. A., PAPAPETROU, E. P., TOMISHIMA, M., SADELAIN, M. & STUDER, L. 2009. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature Biotechnology*, 27, 275-280.
- CHANG, E. A., JIN, S. W., NAM, M. H. & KIM, S. D. 2019. Human Induced Pluripotent Stem Cells : Clinical Significance and Applications in Neurologic Diseases. *J Korean Neurosurg Soc*, 62, 493-501.
- CHAVEZ, A., SCHEIMAN, J., VORA, S., PRUITT, B. W., TUTTLE, M., E, P. R. I., LIN, S., KIANI, S., GUZMAN, C. D., WIEGAND, D. J., TER-OVANESYAN, D., BRAFF, J. L., DAVIDSOHN, N., HOUSDEN, B. E., PERRIMON, N., WEISS, R., AACH, J., COLLINS, J. J. & CHURCH, G. M. 2015. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods*, 12, 326-8.
- CHAVEZ, A., TUTTLE, M., PRUITT, B. W., EWEN-CAMPEN, B., CHARI, R., TER-OVANESYAN, D., HAQUE, S. J., CECCHI, R. J., KOWAL, E. J. K., BUCHTHAL, J., HOUSDEN, B. E., PERRIMON, N., COLLINS, J. J. & CHURCH, G. 2016. Comparison of Cas9 activators in multiple species. *Nature Methods*, 13, 563-567.
- CHEN, F., SWARTZLANDER, D. B., GHOSH, A., FRYER, J. D., WANG, B. & ZHENG, H. 2021. Clusterin secreted from astrocyte promotes excitatory synaptic transmission and ameliorates Alzheimer's disease neuropathology. *Molecular Neurodegeneration*, 16, 5.
- CHEN, G.-F., XU, T.-H., YAN, Y., ZHOU, Y.-R., JIANG, Y., MELCHER, K. & XU, H. E. 2017. Amyloid beta: structure, biology and structure-based therapeutic development. *Acta Pharmacologica Sinica*, 38, 1205-1235.
- CHEN, M., MAIMAITILI, M., HABEKOST, M., GILL, K. P., MERMET-JORET, N., NABAVI, S., FEBBRARO, F. & DENHAM, M. 2020. Rapid generation of regionally specified CNS neurons by sequential patterning and conversion of human induced pluripotent stem cells. *Stem Cell Research*, 48, 101945.
- CHEN, Y.-H., ISHII, M., SUN, J., SUCOV, H. M. & MAXSON, R. E. 2007. Msx1 and Msx2 regulate survival of secondary heart field precursors and post-migratory proliferation of cardiac neural crest in the outflow tract. *Developmental Biology*, 308, 421-437.
- CHEN, Y., BANDOSZ, P., STOYE, G., LIU, Y., WU, Y., LOBANOV-ROSTOVSKY, S., FRENCH, E., KIVIMAKI, M., LIVINGSTON, G., LIAO, J. & BRUNNER, E. J. 2023. Dementia incidence trend in England and Wales, 2002–19, and projection for dementia burden to

- 2040: analysis of data from the English Longitudinal Study of Ageing. *The Lancet Public Health*, 8, e859-e867.
- CHEN, Y., LIU, Z., GONG, P., ZHANG, H., CHEN, Y., YAO, S., LI, W., ZHANG, Y. & YU, Y. 2022. The Chemerin/CMKLR1 Axis Is Involved in the Recruitment of Microglia to A β Deposition through p38 MAPK Pathway. *International Journal of Molecular Sciences*, 23, 9041.
- CHIDAMBARAM, H. & CHINNATHAMBI, S. 2020. G-Protein Coupled Receptors and Tau-different Roles in Alzheimer's Disease. *Neuroscience*, 438, 198-214.
- CHO, S. W., KIM, S., KIM, J. M. & KIM, J. S. 2013. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol*, 31, 230-2.
- CHOI, S. H., KIM, Y. H., HEBISCH, M., SLIWINSKI, C., LEE, S., D'AVANZO, C., CHEN, H., HOOLI, B., ASSELIN, C., MUFFAT, J., KLEE, J. B., ZHANG, C., WAINGER, B. J., PEITZ, M., KOVACS, D. M., WOOLF, C. J., WAGNER, S. L., TANZI, R. E. & KIM, D. Y. 2014. A three-dimensional human neural cell culture model of Alzheimer's disease. *Nature*, 515, 274-278.
- CHOI, S. S., OMENETTI, A., WITEK, R. P., MOYLAN, C. A., SYN, W.-K., JUNG, Y., YANG, L., SUDAN, D. L., SICKLICK, J. K., MICHELOTTI, G. A., ROJKIND, M. & DIEHL, A. M. 2009. Hedgehog pathway activation and epithelial-to-mesenchymal transitions during myofibroblastic transformation of rat hepatic cells in culture and cirrhosis. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 297, G1093-G1106.
- CHUA, G., MORRIS, Q. D., SOPKO, R., ROBINSON, M. D., RYAN, O., CHAN, E. T., FREY, B. J., ANDREWS, B. J., BOONE, C. & HUGHES, T. R. 2006. Identifying transcription factor functions and targets by phenotypic activation. *Proceedings of the National Academy of Sciences*, 103, 12045-12050.
- CHUNG, J. H., BELL, A. C. & FELSENFELD, G. 1997. Characterization of the chicken β -globin insulator. *Proceedings of the National Academy of Sciences*, 94, 575-580.
- CHYLINSKI, K., HUBMANN, M., HANNA, R. E., YANCHUS, C., MICHLITS, G., UIJTTEWAAL, E. C. H., DOENCH, J., SCHRAMEK, D. & ELLING, U. 2019. CRISPR-Switch regulates sgRNA activity by Cre recombination for sequential editing of two loci. *Nature Communications*, 10, 5454.
- COLEMAN, P. D. & YAO, P. J. 2003. Synaptic slaughter in Alzheimer's disease. *Neurobiology of Aging*, 24, 1023-1027.
- CONG, L., RAN, F. A., COX, D., LIN, S., BARRETTO, R., HABIB, N., HSU, P. D., WU, X., JIANG, W., MARRAFFINI, L. A. & ZHANG, F. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339, 819-23.
- COOPER, Y. A., TEYSSIER, N., DRÄGER, N. M., GUO, Q., DAVIS, J. E., SATTLER, S. M., YANG, Z., PATEL, A., WU, S., KOSURI, S., COPPOLA, G., KAMPMANN, M. & GESCHWIND, D. H. 2022. Functional regulatory variants implicate distinct transcriptional networks in dementia. *Science*, 377, eabi8654.
- CORDER, E. H., SAUNDERS, A. M., STRITTMATTER, W. J., SCHMECHEL, D. E., GASKELL, P. C., SMALL, G. W., ROSES, A. D., HAINES, J. L. & PERICAK-VANCE, M. A. 1993. Gene Dose of Apolipoprotein E Type 4 Allele and the Risk of Alzheimer's Disease in Late Onset Families. *Science*, 261, 921-923.
- CUCCATO, G., GATTA, G. D. & DI BERNARDO, D. 2009. Systems and Synthetic biology: tackling genetic networks and complex diseases. *Heredity*, 102, 527-532.
- DANIELYAN, L., BEER-HAMMER, S., STOLZING, A., SCHÄFER, R., SIEGEL, G., FABIAN, C., KAHLE, P., BIEDERMANN, T., LOURHMATI, A., BUADZE, M., NOVAKOVIC, A., PROKSCH, B., GLEITER, C. H., FREY, W. H. & SCHWAB, M. 2014. Intranasal Delivery of Bone Marrow-Derived Mesenchymal Stem Cells, Macrophages, and Microglia to the Brain in Mouse Models of Alzheimer's and Parkinson's Disease. *Cell Transplantation*, 23, 123-139.
- DANIELYAN, L., SCHÄFER, R., VON AMELN-MAYERHOFER, A., BUADZE, M., GEISLER, J., KLOPFER, T., BURKHARDT, U., PROKSCH, B., VERLEYS DONK, S., AYTURAN, M.,

- BUNIATIAN, G. H., GLEITER, C. H. & FREY, W. H. 2009. Intranasal delivery of cells to the brain. *European Journal of Cell Biology*, 88, 315-324.
- DAS, A. T., ZHOU, X., METZ, S. W., VINK, M. A. & BERKHOUT, B. 2016. Selecting the optimal Tet-On system for doxycycline-inducible gene expression in transiently transfected and stably transduced mammalian cells. *Biotechnology Journal*, 11, 71-79.
- DAS, A. T., ZHOU, X., VINK, M., KLAVER, B., VERHOEF, K., MARZIO, G. & BERKHOUT, B. 2004. Viral Evolution as a Tool to Improve the Tetracycline-regulated Gene Expression System*. *Journal of Biological Chemistry*, 279, 18776-18782.
- DAWSON, T. M., GOLDE, T. E. & LAGIER-TOURENNE, C. 2018. Animal models of neurodegenerative diseases. *Nature Neuroscience*, 21, 1370-1379.
- DE CHAVES, E. P., NARAYANASWAMI, V., CHRISTOFFERSEN, C. & NIELSEN, L. B. 2008. Apolipoprotein E and cholesterol in aging and disease in the brain. *Future Lipidology*, 3, 505-530.
- DE FELICE, F. G., WU, D., LAMBERT, M. P., FERNANDEZ, S. J., VELASCO, P. T., LACOR, P. N., BIGIO, E. H., JERECIC, J., ACTON, P. J., SHUGHRUE, P. J., CHEN-DODSON, E., KINNEY, G. G. & KLEIN, W. L. 2008. Alzheimer's disease-type neuronal tau hyperphosphorylation induced by A β oligomers. *Neurobiology of Aging*, 29, 1334-1347.
- DE SOLIS, C. A., HO, A., HOLEHONNUR, R. & PLOSKI, J. E. 2016. The Development of a Viral Mediated CRISPR/Cas9 System with Doxycycline Dependent gRNA Expression for Inducible In vitro and In vivo Genome Editing. *Front Mol Neurosci*, 9, 70.
- DE STROOPER, B. & KARRAN, E. 2016. The Cellular Phase of Alzheimer's Disease. *Cell*, 164, 603-15.
- DELTICHEVA, E., CHYLINSKI, K., SHARMA, C. M., GONZALES, K., CHAO, Y., PIRZADA, Z. A., ECKERT, M. R., VOGEL, J. & CHARPENTIER, E. 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, 471, 602-7.
- DEMATTOS, R. B., CIRRITO, J. R., PARSADANIAN, M., MAY, P. C., O'DELL, M. A., TAYLOR, J. W., HARMONY, J. A. K., ARONOW, B. J., BALES, K. R., PAUL, S. M. & HOLTZMAN, D. M. 2004. ApoE and Clusterin Cooperatively Suppress A β Levels and Deposition: Evidence that ApoE Regulates Extracellular A β Metabolism In Vivo. *Neuron*, 41, 193-202.
- DEMATTOS, R. B., O'DELL, M. A., PARSADANIAN, M., TAYLOR, J. W., HARMONY, J. A. K., BALES, K. R., PAUL, S. M., ARONOW, B. J. & HOLTZMAN, D. M. 2002. Clusterin promotes amyloid plaque formation and is critical for neuritic toxicity in a mouse model of Alzheimer's disease. *Proceedings of the National Academy of Sciences*, 99, 10843-10848.
- DOBIN, A., DAVIS, C. A., SCHLESINGER, F., DRENKOW, J., ZALESKI, C., JHA, S., BATUT, P., CHAISSON, M. & GINGERAS, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29, 15-21.
- DOMÍNGUEZ-GARCÍA, S., GERIBALDI-DOLDÁN, N., GÓMEZ-OLIVA, R., RUIZ, F. A., CARRASCAL, L., BOLÍVAR, J., VERÁSTEGUI, C., GARCIA-ALLOZA, M., MACÍAS-SÁNCHEZ, A. J., HERNÁNDEZ-GALÁN, R., NÚÑEZ-ABADES, P. & CASTRO, C. 2020. A novel PKC activating molecule promotes neuroblast differentiation and delivery of newborn neurons in brain injuries. *Cell Death & Disease*, 11, 262.
- DONG, F., XIE, K., CHEN, Y., YANG, Y. & MAO, Y. 2017. Polycistronic tRNA and CRISPR guide-RNA enables highly efficient multiplexed genome engineering in human cells. *Biochemical and Biophysical Research Communications*, 482, 889-895.
- DOW, L. E., FISHER, J., O'ROURKE, K. P., MULEY, A., KASTENHUBER, E. R., LIVSHITS, G., TSCHAHARGANEH, D. F., SOCCI, N. D. & LOWE, S. W. 2015. Inducible in vivo genome editing with CRISPR-Cas9. *Nat Biotechnol*, 33, 390-394.
- DRÄGER, N. M., SATTLER, S. M., HUANG, C. T.-L., TETER, O. M., LENG, K., HASHEMI, S. H., HONG, J., AVILES, G., CLELLAND, C. D., ZHAN, L., UDEOCHU, J. C., KODAMA, L., SINGLETON, A. B., NALLS, M. A., ICHIDA, J., WARD, M. E., FAGHRI, F., GAN, L. &

- KAMPMANN, M. 2022. A CRISPRi/a platform in human iPSC-derived microglia uncovers regulators of disease states. *Nature Neuroscience*, 25, 1149-1162.
- DUVAL, N., DAUBAS, P., BOURCIER DE CARBON, C., ST CLOMENT, C., TINEVEZ, J.-Y., LOPES, M., RIBES, V. & ROBERT, B. 2014. Msx1 and Msx2 act as essential activators of Atoh1 expression in the murine spinal cord. *Development*, 141, 1726-1736.
- EBERT, A. D., LIANG, P. & WU, J. C. 2012. Induced pluripotent stem cells as a disease modeling and drug screening platform. *J Cardiovasc Pharmacol*, 60, 408-16.
- ECKERT, A., HUANG, L., GONZALEZ, R., KIM, H. S., HAMBLIN, M. H. & LEE, J. P. 2015. Bystander Effect Fuels Human Induced Pluripotent Stem Cell-Derived Neural Stem Cells to Quickly Attenuate Early Stage Neurological Deficits After Stroke. *Stem Cells Transl Med*, 4, 841-51.
- ELOWITZ, M. B. & LEIBLER, S. 2000. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403, 335-338.
- ENGLER, C., KANDZIA, R. & MARILLONNET, S. 2008. A one pot, one step, precision cloning method with high throughput capability. *PLoS One*, 3, e3647.
- FAN, C., GAO, Y., LIANG, G., HUANG, L., WANG, J., YANG, X., SHI, Y., DRÄGER, U. C., ZHONG, M., GAO, T. M. & YANG, X. 2020. Transcriptomics of Gabra4 knockout mice reveals common NMDAR pathways underlying autism, memory, and epilepsy. *Mol Autism*, 11, 13.
- FANG, E. F., HOU, Y., PALIKARAS, K., ADRIAANSE, B. A., KERR, J. S., YANG, B., LAUTRUP, S., HASAN-OLIVE, M. M., CAPONIO, D., DAN, X., ROCKTÄSCHEL, P., CROTEAU, D. L., AKBARI, M., GREIG, N. H., FLADBY, T., NILSEN, H., CADER, M. Z., MATTSON, M. P., TAVERNARAKIS, N. & BOHR, V. A. 2019. Mitophagy inhibits amyloid- β and tau pathology and reverses cognitive deficits in models of Alzheimer's disease. *Nature Neuroscience*, 22, 401-412.
- FAROKHIMANESH, S., RAHBARIZADEH, F., RASAEI, M. J., KAMALI, A. & MASHKANI, B. 2010. Hybrid promoters directed tBid gene expression to breast cancer cells by transcriptional targeting. *Biotechnology Progress*, 26, 505-511.
- FERREIRA, R., SKREKAS, C., NIELSEN, J. & DAVID, F. 2018. Multiplexed CRISPR/Cas9 Genome Editing and Gene Regulation Using Csy4 in *Saccharomyces cerevisiae*. *ACS Synthetic Biology*, 7, 10-15.
- FERRER, I., GÓMEZ, A., CARMONA, M., HUESA, G., PORTA, S., RIERA-CODINA, M., BIAGIOLI, M., GUSTINCICH, S. & ASO, E. 2011. Neuronal Hemoglobin is Reduced in Alzheimer's Disease, Argrophilic Grain Disease, Parkinson's Disease, and Dementia with Lewy Bodies. *Journal of Alzheimer's Disease*, 23, 537-550.
- FERRY, Q. R. V., LYUTOVA, R. & FULGA, T. A. 2017. Rational design of inducible CRISPR guide RNAs for de novo assembly of transcriptional programs. *Nature Communications*, 8, 14633.
- FOSTER, E. M., DANGLA-VALLS, A., LOVESTONE, S., RIBE, E. M. & BUCKLEY, N. J. 2019. Clusterin in Alzheimer's Disease: Mechanisms, Genetics, and Lessons From Other Pathologies. *Frontiers in Neuroscience*, 13.
- FUJIWARA, N., SHIMIZU, J., TAKAI, K., ARIMITSU, N., SAITO, A., KONO, T., UMEHARA, T., UEDA, Y., WAKISAKA, S., SUZUKI, T. & SUZUKI, N. 2013. Restoration of spatial memory dysfunction of human APP transgenic mice by transplantation of neuronal precursors derived from human iPS cells. *Neuroscience Letters*, 557, 129-134.
- GALEANO, C., QIU, Z., MISHRA, A., FARNSWORTH, S. L., HEMMI, J. J., MOREIRA, A., EDENHOFFER, P. & HORNSBY, P. J. 2018. The Route by Which Intranasally Delivered Stem Cells Enter the Central Nervous System. *Cell Transplant*, 27, 501-514.
- GAO, Z., HERRERA-CARRILLO, E. & BERKHOUT, B. 2018. RNA Polymerase II Activity of Type 3 Pol III Promoters. *Mol Ther Nucleic Acids*, 12, 135-145.

- GARDNER, T. S., CANTOR, C. R. & COLLINS, J. J. 2000. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 403, 339-342.
- GIANNAKOPOULOS, P., HERRMANN, F. R., BUSSIÈRE, T., BOURAS, C., KÖVARI, E., PERL, D. P., MORRISON, J. H., GOLD, G. & HOF, P. R. 2003. Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. *Neurology*, 60, 1495-1500.
- GILBERT, L. A., HORLBECK, M. A., ADAMSON, B., VILLALTA, J. E., CHEN, Y., WHITEHEAD, E. H., GUIMARAES, C., PANNING, B., PLOEGH, H. L., BASSIK, M. C., QI, L. S., KAMPMANN, M. & WEISSMAN, J. S. 2014. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*, 159, 647-61.
- GILBERT, L. A., LARSON, M. H., MORSUT, L., LIU, Z., BRAR, G. A., TORRES, S. E., STERN-GINOSSAR, N., BRANDMAN, O., WHITEHEAD, E. H., DOUDNA, J. A., LIM, W. A., WEISSMAN, J. S. & QI, L. S. 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, 154, 442-51.
- GOERTSEN, D., FLYTZANIS, N. C., GOEDEN, N., CHUAPOCO, M. R., CUMMINS, A., CHEN, Y., FAN, Y., ZHANG, Q., SHARMA, J., DUAN, Y., WANG, L., FENG, G., CHEN, Y., IP, N. Y., PICKEL, J. & GRADINARU, V. 2022. AAV capsid variants with brain-wide transgene expression and decreased liver targeting after intravenous delivery in mouse and marmoset. *Nature Neuroscience*, 25, 106-115.
- GOLDE, T. E. 2022. Alzheimer's disease – the journey of a healthy brain into organ failure. *Molecular Neurodegeneration*, 17, 18.
- GONZÁLEZ, F., ZHU, Z., SHI, Z.-D., LELLI, K., VERMA, N., LI, QING V. & HUANGFU, D. 2014. An iCRISPR Platform for Rapid, Multiplexable, and Inducible Genome Editing in Human Pluripotent Stem Cells. *Cell Stem Cell*, 15, 215-226.
- GOSSEN, M. & BUJARD, H. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences*, 89, 5547.
- GOSSEN, M. & BUJARD, H. 2001. Tetracyclines in the control of gene expression in eukaryotes. In: NELSON, M., HILLEN, W. & GREENWALD, R. A. (eds.) *Tetracyclines in Biology, Chemistry and Medicine*. Basel: Birkhäuser Basel.
- GOSSEN, M., FREUNDLIEB, S., BENDER, G., MÜLLER, G., HILLEN, W. & BUJARD, H. 1995. Transcriptional Activation by Tetracyclines in Mammalian Cells. *Science*, 268, 1766-1769.
- GOSTIMSKAYA, I. 2022. CRISPR-Cas9: A History of Its Discovery and Ethical Considerations of Its Use in Genome Editing. *Biochemistry (Mosc)*, 87, 777-788.
- GOVINDPANI, K., TURNER, C., WALDVOGEL, H. J., FAULL, R. L. M. & KWAKOWSKY, A. 2020. Impaired Expression of GABA Signaling Components in the Alzheimer's Disease Middle Temporal Gyrus. *Int J Mol Sci*, 21.
- GRACZYK, D., WHITE, R. J. & RYAN, K. M. 2015. Involvement of RNA Polymerase III in Immune Responses. *Mol Cell Biol*, 35, 1848-59.
- GRIMM, M. O., METT, J., STAHLMANN, C. P., HAUPENTHAL, V. J., ZIMMER, V. C. & HARTMANN, T. 2013. Neprilysin and A β Clearance: Impact of the APP Intracellular Domain in NEP Regulation and Implications in Alzheimer's Disease. *Frontiers in Aging Neuroscience*, 5.
- GRONER, A. C., MEYLAN, S., CIUFFI, A., ZANGGER, N., AMBROSINI, G., DÉNERVAUD, N., BUCHER, P. & TRONO, D. 2010. KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS Genet*, 6, e1000869.
- GUAN, Y.-H., ZHANG, L.-J., WANG, S.-Y., DENG, Y.-D., ZHOU, H.-S., CHEN, D.-Q. & ZHANG, L.-C. 2022. The role of microglia in Alzheimer's disease and progress of treatment. *Ibrain*, 8, 37-47.
- HAASS, C. 2004. Take five--BACE and the gamma-secretase quartet conduct Alzheimer's amyloid beta-peptide generation. *Embo j*, 23, 483-8.

- HACKETT, S. R., BALTZ, E. A., CORAM, M., WRANIK, B. J., KIM, G., BAKER, A., FAN, M., HENDRICKSON, D. G., BERNDL, M. & MCISAAC, R. S. 2020. Learning causal networks using inducible transcription factors and transcriptome-wide time series. *Molecular Systems Biology*, 16, e9174.
- HAGINO-YAMAGISHI, K., SAIJOH, Y., IKEDA, M., ICHIKAWA, M., MINAMIKAWA-TACHINO, R. & HAMADA, H. 1997. Predominant expression of Brn-2 in the postmitotic neurons of the developing mouse neocortex. *Brain Research*, 752, 261-268.
- HAN, J., ISHII, M., BRINGAS, P., MAAS, R. L., MAXSON, R. E. & CHAI, Y. 2007. Concerted action of Msx1 and Msx2 in regulating cranial neural crest cell differentiation during frontal bone development. *Mechanisms of Development*, 124, 729-745.
- HAQUE, M. E., KIM, I.-S., JAKARIA, M., AKTHER, M. & CHOI, D.-K. 2018. Importance of GPCR-Mediated Microglial Activation in Alzheimer's Disease. *Frontiers in cellular neuroscience*, 12, 258-258.
- HARDY, J. 2003. The relationship between amyloid and tau. *Journal of Molecular Neuroscience*, 20, 203-206.
- HARDY, J. 2006. Alzheimer's disease: the amyloid cascade hypothesis: an update and reappraisal. *J Alzheimers Dis*, 9, 151-3.
- HARDY, J. A. & HIGGINS, G. A. 1992. Alzheimer's Disease: The Amyloid Cascade Hypothesis. *Science*, 256, 184-185.
- HARROW, J., FRANKISH, A., GONZALEZ, J. M., TAPANARI, E., DIEKHANS, M., KOKOCINSKI, F., AKEN, B. L., BARRELL, D., ZADISSA, A., SEARLE, S., BARNES, I., BIGNELL, A., BOYCHENKO, V., HUNT, T., KAY, M., MUKHERJEE, G., RAJAN, J., DESPACIO-REYES, G., SAUNDERS, G., STEWARD, C., HARTE, R., LIN, M., HOWALD, C., TANZER, A., DERRIEN, T., CHRAST, J., WALTERS, N., BALASUBRAMANIAN, S., PEI, B., TRESS, M., RODRIGUEZ, J. M., EZKURDIA, I., VAN BAREN, J., BRENT, M., HAUSSLER, D., KELLIS, M., VALENCIA, A., REYMOND, A., GERSTEIN, M., GUIGÓ, R. & HUBBARD, T. J. 2012. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res*, 22, 1760-74.
- HASSAN, M., RAZA, H., ABBASI, M. A., MOUSTAFA, A. A. & SEO, S.-Y. 2019. The exploration of novel Alzheimer's therapeutic agents from the pool of FDA approved medicines using drug repositioning, enzyme inhibition and kinetic mechanism approaches. *Biomedicine & Pharmacotherapy*, 109, 2513-2526.
- HAURWITZ, R. E., JINEK, M., WIEDENHEFT, B., ZHOU, K. & DOUDNA, J. A. 2010. Sequence- and Structure-Specific RNA Processing by a CRISPR Endonuclease. *Science*, 329, 1355-1358.
- HAYASHI, H., KUBO, Y., IZUMIDA, M. & MATSUYAMA, T. 2020a. Efficient viral delivery of Cas9 into human safe harbor. *Scientific Reports*, 10, 21474.
- HAYASHI, Y., LIN, H.-T., LEE, C.-C. & TSAI, K.-J. 2020b. Effects of neural stem cell transplantation in Alzheimer's disease models. *Journal of Biomedical Science*, 27, 29.
- HE, H.-Q. & YE, R. D. 2017. The Formyl Peptide Receptors: Diversity of Ligands and Mechanism for Recognition. *Molecules*, 22, 455.
- HEIDERSBACH, A. J., DORIGHI, K. M., GOMEZ, J. A., JACOBI, A. M. & HALEY, B. 2023. A versatile, high-efficiency platform for CRISPR-based gene activation. *Nature Communications*, 14, 902.
- HENRIKSEN, J. R., LØKKE, C., HAMMERØ, M., GEERTS, D., VERSTEEG, R., FLAEGSTAD, T. & EINVIK, C. 2007. Comparison of RNAi efficiency mediated by tetracycline-responsive H1 and U6 promoter variants in mammalian cell lines. *Nucleic acids research*, 35, e67-e67.
- HEROLD, M. J., VAN DEN BRANDT, J., SEIBLER, J. & REICHARDT, H. M. 2008. Inducible and reversible gene silencing by stable integration of an shRNA-encoding lentivirus in transgenic rats. *Proceedings of the National Academy of Sciences*, 105, 18507.

- HEVNER, R. F. 2019. Intermediate progenitors and Tbr2 in cortical development. *Journal of anatomy*, 235, 616-625.
- HICKMAN, S. E. & EL KHOURY, J. 2010. Mechanisms of mononuclear phagocyte recruitment in Alzheimer's disease. *CNS Neurol Disord Drug Targets*, 9, 168-73.
- HILTON, I. B., D'IPPOLITO, A. M., VOCKLEY, C. M., THAKORE, P. I., CRAWFORD, G. E., REDDY, T. E. & GERSBACH, C. A. 2015. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature Biotechnology*, 33, 510-517.
- HIRAI, H., TANI, T. & KIKYO, N. 2010. Structure and functions of powerful transactivators: VP16, MyoD and FoxA. *Int J Dev Biol*, 54, 1589-96.
- HITOSHI, N., KEN-ICHI, Y. & JUN-ICHI, M. 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene*, 108, 193-199.
- HOBOM, B. 1980. [Gene surgery: on the threshold of synthetic biology]. *Med Klin*, 75, 834-41.
- HOCKEMEYER, D. & JAENISCH, R. 2016. Induced Pluripotent Stem Cells Meet Genome Editing. *Cell Stem Cell*, 18, 573-586.
- HOLLNAGEL, A., OEHLMANN, V., HEYMER, J., RÜTHER, U. & NORDHEIM, A. 1999. Id Genes Are Direct Targets of Bone Morphogenetic Protein Induction in Embryonic Stem Cells*. *Journal of Biological Chemistry*, 274, 19838-19845.
- HUANG, E. J. & REICHARDT, L. F. 2001. Neurotrophins: roles in neuronal development and function. *Annu Rev Neurosci*, 24, 677-736.
- HUANG, Q., CHAN, K. Y., WU, J., BOTTICELLO-ROMERO, N. R., ZHENG, Q., LOU, S., KEYES, C., SVANBERGSSON, A., JOHNSTON, J., MILLS, A., LIN, C.-Y., BRAUER, P. P., CLOUSE, G., PACOURET, S., HARVEY, J. W., BEDDOW, T., HURLEY, J. K., TOBEY, I. G., POWELL, M., CHEN, A. T., BARRY, A. J., EID, F.-E., CHAN, Y. A. & DEVERMAN, B. E. 2024. An AAV capsid reprogrammed to bind human transferrin receptor mediates brain-wide gene delivery. *Science*, 384, 1220-1227.
- HULME, A. J., MAKSOOR, S., ST-CLAIR GLOVER, M., MIELLET, S. & DOTTORI, M. 2022. Making neurons, made easy: The use of Neurogenin-2 in neuronal differentiation. *Stem Cell Reports*, 17, 14-34.
- HWANG, J.-Y. & ZUKIN, R. S. 2018. REST, a master transcriptional regulator in neurodegenerative disease. *Current Opinion in Neurobiology*, 48, 193-200.
- HWANG, W. Y., FU, Y., REYON, D., MAEDER, M. L., TSAI, S. Q., SANDER, J. D., PETERSON, R. T., YEH, J. R. & JOUNG, J. K. 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol*, 31, 227-9.
- IJIMA-ANDO, K., HEARN, S. A., GRANGER, L., SHENTON, C., GATT, A., CHIANG, H.-C., HAKKER, I., ZHONG, Y. & IJIMA, K. 2008. Overexpression of neprilysin reduces alzheimer amyloid-beta42 (Abeta42)-induced neuron loss and intraneuronal Abeta42 deposits but causes a reduction in cAMP-responsive element-binding protein-mediated transcription, age-dependent axon pathology, and premature death in Drosophila. *The Journal of biological chemistry*, 283, 19066-19076.
- JACKSON, J. T., NUTT, S. L. & MCCORMACK, M. P. 2023. The Haematopoietically-expressed homeobox transcription factor: roles in development, physiology and disease. *Frontiers in Immunology*, 14.
- JANSEN, T. & DAIBER, A. 2012. Direct Antioxidant Properties of Bilirubin and Biliverdin. Is there a Role for Biliverdin Reductase? *Frontiers in Pharmacology*, 3.
- JEAN-CHARLES, P. Y., KAUR, S. & SHENOY, S. K. 2017. G Protein-Coupled Receptor Signaling Through β -Arrestin-Dependent Mechanisms. *J Cardiovasc Pharmacol*, 70, 142-158.
- JINEK, M., CHYLINSKI, K., FONFARA, I., HAUER, M., DOUDNA, J. A. & CHARPENTIER, E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337, 816-21.

- JINEK, M., EAST, A., CHENG, A., LIN, S., MA, E. & DOUDNA, J. 2013. RNA-programmed genome editing in human cells. *eLife*, 2, e00471.
- JOHNSON, G. V. W. & STOOTHOFF, W. H. 2004. Tau phosphorylation in neuronal cell function and dysfunction. *Journal of Cell Science*, 117, 5721-5729.
- JOVANOVIC, V. M., SALT, A., TILLEMANN, H., ZEGA, K., JUKIC, M. M., ZOU, H., FRIEDEL, R. H., PRAKASH, N., BLAESS, S., EDENHOFER, F. & BRODSKI, C. 2018. BMP/SMAD Pathway Promotes Neurogenesis of Midbrain Dopaminergic Neurons In Vivo and in Human Induced Pluripotent and Neural Stem Cells. *J Neurosci*, 38, 1662-1676.
- KAMPMANN, M. 2017. A CRISPR Approach to Neurodegenerative Diseases. *Trends Mol Med*, 23, 483-485.
- KAMPMANN, M. 2018. CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. *ACS Chemical Biology*, 13, 406-416.
- KANAFI, M. M. & TAVALLAEI, M. 2022. Overview of advances in CRISPR/deadCas9 technology and its applications in human diseases. *Gene*, 830, 146518.
- KAPUST, R. B., TÖZSÉR, J., FOX, J. D., ANDERSON, D. E., CHERRY, S., COPELAND, T. D. & WAUGH, D. S. 2001. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Engineering, Design and Selection*, 14, 993-1000.
- KARBASSI, E., PADGETT, R., BERTERO, A., REINECKE, H., KLAIMAN, J. M., YANG, X., HAUSCHKA, S. D. & MURRY, C. E. 2024. Targeted CRISPR activation is functional in engineered human pluripotent stem cells but undergoes silencing after differentiation into cardiomyocytes and endothelium. *Cellular and Molecular Life Sciences*, 81, 95.
- KARRA, D. & DAHM, R. 2010. Transfection Techniques for Neuronal Cells. *The Journal of Neuroscience*, 30, 6171-6177.
- KAZI, A., XIANG, S., YANG, H., DELITTO, D., TREVINO, J., JIANG, R. H. Y., AYAZ, M., LAWRENCE, H. R., KENNEDY, P. & SEBTI, S. M. 2018. GSK3 suppression upregulates β -catenin and c-Myc to abrogate KRas-dependent tumors. *Nature Communications*, 9, 5154.
- KEARNS, N. A., PHAM, H., TABAK, B., GENGA, R. M., SILVERSTEIN, N. J., GARBER, M. & MAEHR, R. 2015. Functional annotation of native enhancers with a Cas9–histone demethylase fusion. *Nature Methods*, 12, 401-403.
- KELKAR, A., ZHU, Y., GROTH, T., STOLFA, G., STABLEWSKI, A. B., SINGHI, N., NEMETH, M. & NEELAMEGHAM, S. 2020. Doxycycline-Dependent Self-Inactivation of CRISPR-Cas9 to Temporally Regulate On- and Off-Target Editing. *Molecular Therapy*, 28, 29-41.
- KEMALADEWI, D. U., BASSI, P. S., ERWOOD, S., AL-BASHA, D., GAWLIK, K. I., LINDSAY, K., HYATT, E., KEMBER, R., PLACE, K. M., MARKS, R. M., DURBEEJ, M., PRESCOTT, S. A., IVAKINE, E. A. & COHN, R. D. 2019. A mutation-independent approach for muscular dystrophy via upregulation of a modifier gene. *Nature*, 572, 125-130.
- KEPP, K. P., ROBAKIS, N. K., HØILUND-CARLSEN, P. F., SENSI, S. L. & VISSEL, B. 2023. The amyloid cascade hypothesis: an updated critical review. *Brain*, 146, 3969-3990.
- KILLIAN, T., DICKOPF, S., HAAS, A. K., KIRSTENPFAD, C., MAYER, K. & BRINKMANN, U. 2017. Disruption of diphthamide synthesis genes and resulting toxin resistance as a robust technology for quantifying and optimizing CRISPR/Cas9-mediated gene editing. *Scientific Reports*, 7, 15480.
- KIM, D.-S., LEE, J. S., LEEM, J. W., HUH, Y. J., KIM, J. Y., KIM, H.-S., PARK, I.-H., DALEY, G. Q., HWANG, D.-Y. & KIM, D.-W. 2010. Robust Enhancement of Neural Differentiation from Human ES and iPS Cells Regardless of their Innate Difference in Differentiation Propensity. *Stem Cell Reviews and Reports*, 6, 270-281.
- KIPNISS, N. H., DINGAL, P. C. D. P., ABBOTT, T. R., GAO, Y., WANG, H., DOMINGUEZ, A. A., LABANIEH, L. & QI, L. S. 2017. Engineering cell sensing and responses using a GPCR-coupled CRISPR-Cas system. *Nature Communications*, 8, 2212.

- KIS, Z., PEREIRA, H. S. A., HOMMA, T., PEDRIGI, R. M. & KRAMS, R. 2015. Mammalian synthetic biology: emerging medical applications. *Journal of The Royal Society Interface*, 12, 20141000.
- KITADA, T., DIANDRETH, B., TEAGUE, B. & WEISS, R. 2018. Programming gene and engineered-cell therapies with synthetic biology. *Science*, 359, eaad1067.
- KLAJN, A., DRAKULIC, D., TOSIC, M., PAVKOVIC, Z., SCHWIRTLICH, M. & STEVANOVIC, M. 2014. SOX2 overexpression affects neural differentiation of human pluripotent NT2/D1 cells. *Biochemistry (Mosc)*, 79, 1172-82.
- KLATT, D., CHENG, E., HOFFMANN, K., SANTILLI, G., THRASHER, A. J., BRENDEN, C. & SCHAMBACH, A. 2020. Differential Transgene Silencing of Myeloid-Specific Promoters in the AAVS1 Safe Harbor Locus of Induced Pluripotent Stem Cell-Derived Myeloid Cells. *Human Gene Therapy*, 31, 199-210.
- KLOSKE, C. M. & WILCOCK, D. M. 2020. The Important Interface Between Apolipoprotein E and Neuroinflammation in Alzheimer's Disease. *Frontiers in Immunology*, 11.
- KNAPP, D. J. H. F., MICHAELS, Y. S., JAMILLY, M., FERRY, Q. R. V., BARBOSA, H., MILNE, T. A. & FULGA, T. A. 2019. Decoupling tRNA promoter and processing activities enables specific Pol-II Cas9 guide RNA expression. *Nature Communications*, 10, 1490.
- KNOPMAN, D. S., AMIEVA, H., PETERSEN, R. C., CHÉTELAT, G., HOLTZMAN, D. M., HYMAN, B. T., NIXON, R. A. & JONES, D. T. 2021. Alzheimer disease. *Nature Reviews Disease Primers*, 7, 33.
- KONDO, T., ASAI, M., TSUKITA, K., KUTOKU, Y., OHSAWA, Y., SUNADA, Y., IMAMURA, K., EGAWA, N., YAHATA, N., OKITA, K., TAKAHASHI, K., ASAKA, I., AOI, T., WATANABE, A., WATANABE, K., KADOYA, C., NAKANO, R., WATANABE, D., MARUYAMA, K., HORI, O., HIBINO, S., CHOSHI, T., NAKAHATA, T., HIOKI, H., KANEKO, T., NAITOH, M., YOSHIKAWA, K., YAMAWAKI, S., SUZUKI, S., HATA, R., UENO, S.-I., SEKI, T., KOBAYASHI, K., TODA, T., MURAKAMI, K., IRIE, K., KLEIN, WILLIAM L., MORI, H., ASADA, T., TAKAHASHI, R., IWATA, N., YAMANAKA, S. & INOUE, H. 2013. Modeling Alzheimer's Disease with iPSCs Reveals Stress Phenotypes Associated with Intracellular A β and Differential Drug Responsiveness. *Cell Stem Cell*, 12, 487-496.
- KONERMANN, S., BRIGHAM, M. D., TREVINO, A. E., JOUNG, J., ABUDAYYEH, O. O., BARCENA, C., HSU, P. D., HABIB, N., GOOTENBERG, J. S., NISHIMASU, H., NUREKI, O. & ZHANG, F. 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, 517, 583-588.
- KOPPEN, C. J. V. & JAKOBS, K. H. 2004. Arrestin-Independent Internalization of G Protein-Coupled Receptors. *Molecular Pharmacology*, 66, 365-367.
- KOR, S. D., CHOWDHURY, N., KEOT, A. K., YOGENDRA, K., CHIKKAPUTTAIAH, C. & SUDHAKAR REDDY, P. 2023. RNA Pol III promoters—key players in precisely targeted plant genome editing. *Frontiers in Genetics*, 13.
- KOTIN, R. M., LINDEN, R. M. & BERNIS, K. I. 1992. Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. *The EMBO Journal*, 11, 5071-5078.
- KRAMER, B. P., VIRETTA, A. U., BABA, M. D.-E., AUBEL, D., WEBER, W. & FUSSENEGGER, M. 2004. An engineered epigenetic transgene switch in mammalian cells. *Nature Biotechnology*, 22, 867-870.
- KROEZE, W. K., SASSANO, M. F., HUANG, X.-P., LANSU, K., MCCORVY, J. D., GIGUÈRE, P. M., SCIAKY, N. & ROTH, B. L. 2015. PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nature Structural & Molecular Biology*, 22, 362-369.
- KÜGLER, S., KILIC, E. & BÄHR, M. 2003. Human synapsin 1 gene promoter confers highly neuron-specific long-term transgene expression from an adenoviral vector in the adult rat brain depending on the transduced area. *Gene Ther*, 10, 337-47.

- KUMAR, N., STANFORD, W., DE SOLIS, C., ARADHANA, ABRAHAM, N. D., DAO, T.-M. J., THASEEN, S., SAIRAVI, A., GONZALEZ, C. U. & PLOSKI, J. E. 2018. The Development of an AAV-Based CRISPR SaCas9 Genome Editing System That Can Be Delivered to Neurons in vivo and Regulated via Doxycycline and Cre-Recombinase. *Frontiers in Molecular Neuroscience*, 11.
- KURATA, M., WOLF, N. K., LAHR, W. S., WEG, M. T., KLUESNER, M. G., LEE, S., HUI, K., SHIRAIWA, M., WEBBER, B. R. & MORIARTY, B. S. 2018. Highly multiplexed genome engineering using CRISPR/Cas9 gRNA arrays. *PLOS ONE*, 13, e0198714.
- KWANG, S. J., BRUGGER, S. M., LAZIK, A., MERRILL, A. E., WU, L.-Y., LIU, Y.-H., ISHII, M., SANGIORGI, F. O., RAUCHMAN, M., SUCOV, H. M., MAAS, R. L. & MAXSON, R. E., JR. 2002. Msx2 is an immediate downstream effector of Pax3 in the development of the murine cardiac neural crest. *Development*, 129, 527-538.
- LAMBERT, J.-C., IBRAHIM-VERBAAS, C. A., HAROLD, D., NAJ, A. C., SIMS, R., BELLENGUEZ, C., JUN, G., DESTEFANO, A. L., BIS, J. C., BEECHAM, G. W., GRENIER-BOLEY, B., RUSSO, G., THORNTON-WELLS, T. A., JONES, N., SMITH, A. V., CHOURAKI, V., THOMAS, C., IKRAM, M. A., ZELENKA, D., VARDARAJAN, B. N., KAMATANI, Y., LIN, C.-F., GERRISH, A., SCHMIDT, H., KUNKLE, B., DUNSTAN, M. L., RUIZ, A., BIHOREAU, M.-T., CHOI, S.-H., REITZ, C., PASQUIER, F., HOLLINGWORTH, P., RAMIREZ, A., HANON, O., FITZPATRICK, A. L., BUXBAUM, J. D., CAMPION, D., CRANE, P. K., BALDWIN, C., BECKER, T., GUDNASON, V., CRUCHAGA, C., CRAIG, D., AMIN, N., BERR, C., LOPEZ, O. L., DE JAGER, P. L., DERAMECOURT, V., JOHNSTON, J. A., EVANS, D., LOVESTONE, S., LETENNEUR, L., MORÓN, F. J., RUBINSZTEIN, D. C., EIRIKSDOTTIR, G., SLEEGERS, K., GOATE, A. M., FIÉVET, N., HUENTELMAN, M. J., GILL, M., BROWN, K., KAMBOH, M. I., KELLER, L., BARBERGER-GATEAU, P., MCGUINNESS, B., LARSON, E. B., GREEN, R., MYERS, A. J., DUFOUIL, C., TODD, S., WALLON, D., LOVE, S., ROGAEVA, E., GALLACHER, J., ST GEORGE-HYSLOP, P., CLARIMON, J., LLEO, A., BAYER, A., TSUANG, D. W., YU, L., TSOLAKI, M., BOSSÙ, P., SPALLETTA, G., PROITSI, P., COLLINGE, J., SORBI, S., SANCHEZ-GARCIA, F., FOX, N. C., HARDY, J., NARANJO, M. C. D., BOSCO, P., CLARKE, R., BRAYNE, C., GALIMBERTI, D., MANCUSO, M., MATTHEWS, F., MOEBUS, S., MECOCCHI, P., DEL ZOMPO, M., MAIER, W., et al. 2013. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature Genetics*, 45, 1452-1458.
- LANFER, J., KAINDL, J., KRUMM, L., GONZALEZ ACERA, M., NEURATH, M., REGENSBURGER, M., KRACH, F. & WINNER, B. 2022. Efficient and Easy Conversion of Human iPSCs into Functional Induced Microglia-like Cells. *Int J Mol Sci*, 23.
- LANNES, N., EPPLER, E., ETEMAD, S., YOTOVSKI, P. & FILGUEIRA, L. 2017. Microglia at center stage: a comprehensive review about the versatile and unique residential macrophages of the central nervous system. *Oncotarget*, 8, 114393-114413.
- LE RHUN, A., ESCALERA-MAURER, A., BRATOVIČ, M. & CHARPENTIER, E. 2019. CRISPR-Cas in *Streptococcus pyogenes*. *RNA Biology*, 16, 380-389.
- LEE, A. S., TANG, C., RAO, M. S., WEISSMAN, I. L. & WU, J. C. 2013. Tumorigenicity as a clinical hurdle for pluripotent stem cell therapies. *Nature Medicine*, 19, 998-1004.
- LENG, K., ROSE, I. V. L., KIM, H., XIA, W., ROMERO-FERNANDEZ, W., ROONEY, B., KOONTZ, M., LI, E., AO, Y., WANG, S., KRAWCZYK, M., TCW, J., GOATE, A., ZHANG, Y., ULLIAN, E. M., SOFRONIEW, M. V., FANCY, S. P. J., SCHRAG, M. S., LIPPMANN, E. S. & KAMPMANN, M. 2022. CRISPRi screens in human iPSC-derived astrocytes elucidate regulators of distinct inflammatory reactive states. *Nature Neuroscience*, 25, 1528-1542.
- LI, C., HIRSCH, M., CARTER, P., ASOKAN, A., ZHOU, X., WU, Z. & SAMULSKI, R. J. 2009. A small regulatory element from chromosome 19 enhances liver-specific gene expression. *Gene Therapy*, 16, 43-51.

- LI, K., LIU, Y., CAO, H., ZHANG, Y., GU, Z., LIU, X., YU, A., KAPHLE, P., DICKERSON, K. E., NI, M. & XU, J. 2020. Interrogation of enhancer function by enhancer-targeting CRISPR epigenetic editing. *Nature Communications*, 11, 485.
- LI, Y.-H., YU, J.-W., XI, J.-Y., YU, W.-B., LIU, J.-C., WANG, Q., SONG, L.-J., FENG, L., YAN, Y.-P., ZHANG, G.-X., XIAO, B.-G. & MA, C.-G. 2017. Fasudil Enhances Therapeutic Efficacy of Neural Stem Cells in the Mouse Model of MPTP-Induced Parkinson's Disease. *Molecular Neurobiology*, 54, 5400-5413.
- LIANG, J.-W., FANG, Z.-Y., HUANG, Y., LIUYANG, Z.-Y., ZHANG, X.-L., WANG, J.-L., WEI, H., WANG, J.-Z., WANG, X.-C., ZENG, J. & LIU, R. 2018. Application of Weighted Gene Co-Expression Network Analysis to Explore the Key Genes in Alzheimer's Disease. *Journal of Alzheimer's Disease*, 65, 1353-1364.
- LIAO, Y., SMYTH, G. K. & SHI, W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30, 923-930.
- LIM, W. A. & JUNE, C. H. 2017. The Principles of Engineering Immune Cells to Treat Cancer. *Cell*, 168, 724-740.
- LIN, Y. T., SEO, J., GAO, F., FELDMAN, H. M., WEN, H. L., PENNEY, J., CAM, H. P., GJONESKA, E., RAJA, W. K., CHENG, J., RUEDA, R., KRITSKIY, O., ABDURROB, F., PENG, Z., MILO, B., YU, C. J., ELMSAOURI, S., DEY, D., KO, T., YANKNER, B. A. & TSAI, L. H. 2018. APOE4 Causes Widespread Molecular and Cellular Alterations Associated with Alzheimer's Disease Phenotypes in Human iPSC-Derived Brain Cell Types. *Neuron*, 98, 1294.
- LIU, C., HUANG, M., HAN, C., LI, H., WANG, J., HUANG, Y., CHEN, Y., ZHU, J., FU, G., YU, H., LEI, Z. & CHU, X. 2021. A narrative review of the roles of muscle segment homeobox transcription factor family in cancer. *Annals of Translational Medicine*, 9, 810.
- LIU, W., LI, J., YANG, M., KE, X., DAI, Y., LIN, H., WANG, S., CHEN, L. & TAO, J. 2022a. Chemical genetic activation of the cholinergic basal forebrain hippocampal circuit rescues memory loss in Alzheimer's disease. *Alzheimer's Research & Therapy*, 14, 53.
- LIU, X., CHE, R., LIANG, W., ZHANG, Y., WU, L., HAN, C., LU, H., SONG, W., WU, Y. & WANG, Z. 2022b. Clusterin transduces Alzheimer-risk signals to amyloidogenesis. *Signal Transduction and Targeted Therapy*, 7, 325.
- LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2- $\Delta\Delta$ CT Method. *Methods*, 25, 402-408.
- LOEW, R., HEINZ, N., HAMPF, M., BUJARD, H. & GOSSEN, M. 2010. Improved Tet-responsive promoters with minimized background expression. *BMC Biotechnology*, 10, 81.
- LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15, 550.
- LOWERY, J. W. & ROSEN, V. 2018. Bone Morphogenetic Protein-Based Therapeutic Approaches. *Cold Spring Harb Perspect Biol*, 10.
- LU, C., SHI, X., ALLEN, A., BAEZ-NIETO, D., NIKISH, A., SANJANA, N. E. & PAN, J. Q. 2019. Overexpression of NEUROG2 and NEUROG1 in human embryonic stem cells produces a network of excitatory and inhibitory neurons. *Faseb j*, 33, 5287-5299.
- LU, T., ARON, L., ZULLO, J., PAN, Y., KIM, H., CHEN, Y., YANG, T.-H., KIM, H.-M., DRAKE, D., LIU, X. S., BENNETT, D. A., COLAIÁCOVO, M. P. & YANKNER, B. A. 2014. REST and stress resistance in ageing and Alzheimer's disease. *Nature*, 507, 448-454.
- LUND, R. J., NÄRVÄ, E. & LAHESMAA, R. 2012. Genetic and epigenetic stability of human pluripotent stem cells. *Nature Reviews Genetics*, 13, 732-744.
- LUNDIN, A., PORRITT, M. J., JAISWAL, H., SEELIGER, F., JOHANSSON, C., BIDAR, A. W., BADERTSCHER, L., WIMBERGER, S., DAVIES, E. J., HARDAKER, E., MARTINS, C. P., JAMES, E., ADMYRE, T., TAHERI-GHAHFAROKHI, A., BRADLEY, J., SCHANTZ, A., ALAEIMAHABADI, B., CLAUSEN, M., XU, X., MAYR, L. M., NITSCH, R., BOHLOOLY-Y, M., BARRY, S. T. & MARESCA, M. 2020. Development of an ObLiGaRe Doxycycline Inducible

- Cas9 system for pre-clinical cancer drug discovery. *Nature communications*, 11, 4903-4903.
- LUO, H., JING, B., XIA, Y., ZHANG, Y., HU, M., CAI, H., TONG, Y., ZHOU, L., YANG, L., YANG, J., LEI, H., XU, H., LIU, C. & WU, Y. 2019. WP1130 reveals USP24 as a novel target in T-cell acute lymphoblastic leukemia. *Cancer Cell International*, 19, 56.
- MALI, P., YANG, L., ESVELT, K. M., AACH, J., GUELL, M., DICARLO, J. E., NORVILLE, J. E. & CHURCH, G. M. 2013. RNA-guided human genome engineering via Cas9. *Science*, 339, 823-6.
- MALI, S. 2013. Delivery systems for gene therapy. *Indian J Hum Genet*, 19, 3-8.
- MALONEY, J. A., BAINBRIDGE, T., GUSTAFSON, A., ZHANG, S., KYAUK, R., STEINER, P., VAN DER BRUG, M., LIU, Y., ERNST, J. A., WATTS, R. J. & ATWAL, J. K. 2014. Molecular Mechanisms of Alzheimer Disease Protection by the A673T Allele of Amyloid Precursor Protein*. *Journal of Biological Chemistry*, 289, 30990-31000.
- MANSOURI, M. & FUSSENEGGER, M. 2022. Therapeutic cell engineering: designing programmable synthetic genetic circuits in mammalian cells. *Protein & Cell*, 13, 476-489.
- MAREI, H. E., KHAN, M. U. A. & HASAN, A. 2023. Potential use of iPSCs for disease modeling, drug screening, and cell-based therapy for Alzheimer's disease. *Cellular & Molecular Biology Letters*, 28, 98.
- MARR, R. A., ROCKENSTEIN, E., MUKHERJEE, A., KINDY, M. S., HERSH, L. B., GAGE, F. H., VERMA, I. M. & MASLIAH, E. 2003. Neprilysin Gene Transfer Reduces Human Amyloid Pathology in Transgenic Mice. *The Journal of Neuroscience*, 23, 1992.
- MARTÍNEZ-GARCÍA, A., SASTRE, I., RECUERO, M., ALDUDO, J., VILELLA, E., MATEO, I., SÁNCHEZ-JUAN, P., VARGAS, T., CARRO, E., BERMEJO-PAREJA, F., RODRÍGUEZ-RODRÍGUEZ, E., COMBARROS, O., ROSICH-ESTRAGO, M., FRANK, A., VALDIVIESO, F. & BULLIDO, M. J. 2010. PLA2G3, a Gene Involved in Oxidative Stress Induced Death, is Associated with Alzheimer's Disease. *Journal of Alzheimer's Disease*, 22, 1181-1187.
- MAY, P. C., LAMPERT-ETCHELLS, M., JOHNSON, S. A., POIRIER, J., MASTERS, J. N. & FINCH, C. E. 1990. Dynamics of gene expression for a hippocampal glycoprotein elevated in Alzheimer's disease and in response to experimental lesions in rat. *Neuron*, 5, 831-839.
- MCALLISTER, A. K. 2001. Neurotrophins and neuronal differentiation in the central nervous system. *Cellular and Molecular Life Sciences CMLS*, 58, 1054-1060.
- MCCARTY, N. S., GRAHAM, A. E., STUDENÁ, L. & LEDESMA-AMARO, R. 2020. Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nature Communications*, 11, 1281.
- MCHUGH, D., HU, S. S. J., RIMMERMAN, N., JUKNAT, A., VOGEL, Z., WALKER, J. M. & BRADSHAW, H. B. 2010. N-arachidonoyl glycine, an abundant endogenous lipid, potently drives directed cellular migration through GPR18, the putative abnormal cannabidiol receptor. *BMC Neuroscience*, 11, 44.
- MCLEAN, J. R., SMITH, G. A., ROCHA, E. M., HAYES, M. A., BEAGAN, J. A., HALLETT, P. J. & ISACSON, O. 2014. Widespread neuron-specific transgene expression in brain and spinal cord following synapsin promoter-driven AAV9 neonatal intracerebroventricular injection. *Neuroscience Letters*, 576, 73-78.
- MCNAMARA, M. J., GOMEZ-ISLA, T. & HYMAN, B. T. 1998. Apolipoprotein E Genotype and Deposits of Aβ40 and Aβ42 in Alzheimer Disease. *Archives of Neurology*, 55, 1001-1004.
- MEFFERD, A. L., KORNEPATI, A. V., BOGERD, H. P., KENNEDY, E. M. & CULLEN, B. R. 2015. Expression of CRISPR/Cas single guide RNAs using small tRNA promoters. *Rna*, 21, 1683-9.
- MENEZES, J. & LUSKIN, M. 1994. Expression of neuron-specific tubulin defines a novel population in the proliferative layers of the developing telencephalon. *The Journal of Neuroscience*, 14, 5399-5416.

- MERINO-ZAMORANO, C., FERNÁNDEZ-DE RETANA, S., MONTAÑOLA, A., BATLLE, A., SAINT-POL, J., MYSIOREK, C., GOSSELET, F., MONTANER, J. & HERNÁNDEZ-GUILLAMON, M. 2016. Modulation of Amyloid- β 1–40 Transport by ApoA1 and ApoJ Across an in vitro Model of the Blood-Brain Barrier. *Journal of Alzheimer's Disease*, 53, 677-691.
- MEYER, K., FELDMAN, H. M., LU, T., DRAKE, D., LIM, E. T., LING, K.-H., BISHOP, N. A., PAN, Y., SEO, J., LIN, Y.-T., SU, S. C., CHURCH, G. M., TSAI, L.-H. & YANKNER, B. A. 2019. REST and Neural Gene Network Dysregulation in iPSC Models of Alzheimer's Disease. *Cell Reports*, 26, 1112-1127.e9.
- MILLER, JUSTINE D., GANAT, YOSIF M., KISHINEVSKY, S., BOWMAN, ROBERT L., LIU, B., TU, EDMUND Y., MANDAL, P. K., VERA, E., SHIM, J.-W., KRIKS, S., TALDONE, T., FUSAKI, N., TOMISHIMA, MARK J., KRAINIC, D., MILNER, TERESA A., ROSSI, DERRICK J. & STUDER, L. 2013. Human iPSC-Based Modeling of Late-Onset Disease via Progerin-Induced Aging. *Cell Stem Cell*, 13, 691-705.
- MINERS, J. S., CLARKE, P. & LOVE, S. 2017. Clusterin levels are increased in Alzheimer's disease and influence the regional distribution of A β . *Brain Pathol*, 27, 305-313.
- MISHRA, S., VAN DEURSEN, C., ROBERT, A., VAN DER KANT, R. & YOUNG, J. E. 2023. Chapter 11 - Advantages and limitations of hiPSC-derived neurons for the study of neurodegeneration. In: BUTTERMORE, E. D. (ed.) *Phenotyping of Human iPSC-derived Neurons*. Academic Press.
- MIYAMOTO, T., RAZAVI, S., DEROSE, R. & INOUE, T. 2013. Synthesizing Biomolecule-Based Boolean Logic Gates. *ACS Synthetic Biology*, 2, 72-82.
- MO, H., KIM, J., KIM, J. Y., KIM, J. W., HAN, H., CHOI, S. H., RIM, Y. A. & JU, J. H. 2023. Intranasal administration of induced pluripotent stem cell-derived cortical neural stem cell-secretome as a treatment option for Alzheimer's disease. *Translational Neurodegeneration*, 12, 50.
- MONOD, J. & JACOB, F. 1961.
- MOOTHA, V. K., LINDGREN, C. M., ERIKSSON, K.-F., SUBRAMANIAN, A., SIHAG, S., LEHAR, J., PUIGSERVER, P., CARLSSON, E., RIDDERSTRÅLE, M., LAURILA, E., HOUSTIS, N., DALY, M. J., PATTERSON, N., MESIROV, J. P., GOLUB, T. R., TAMAYO, P., SPIEGELMAN, B., LANDER, E. S., HIRSCHHORN, J. N., ALTSHULER, D. & GROOP, L. C. 2003. PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics*, 34, 267-273.
- MORIYAMA, H., MORIYAMA, M., SAWARAGI, K., OKURA, H., ICHINOSE, A., MATSUYAMA, A. & HAYAKAWA, T. 2013. Tightly Regulated and Homogeneous Transgene Expression in Human Adipose-Derived Mesenchymal Stem Cells by Lentivirus with Tet-Off System. *PLOS ONE*, 8, e66274.
- MULDER, S. D., NIELSEN, H. M., BLANKENSTEIN, M. A., EIKELENBOOM, P. & VEERHUIS, R. 2014. Apolipoproteins E and J interfere with amyloid-beta uptake by primary human astrocytes and microglia in vitro. *Glia*, 62, 493-503.
- MURATORE, C. R., RICE, H. C., SRIKANTH, P., CALLAHAN, D. G., SHIN, T., BENJAMIN, L. N. P., WALSH, D. M., SELKOE, D. J. & YOUNG-PEARSE, T. L. 2014. The familial Alzheimer's disease APPV717I mutation alters APP processing and Tau expression in iPSC-derived neurons. *Human Molecular Genetics*, 23, 3523-3536.
- NAGEL, S., VENTURINI, L., PRZYBYLSKI, G. K., GRABARCZYK, P., MEYER, C., KAUFMANN, M., BATTMER, K., SCHMIDT, C. A., DREXLER, H. G., SCHERR, M. & MACLEOD, R. A. 2009. NK-like homeodomain proteins activate NOTCH3-signaling in leukemic T-cells. *BMC Cancer*, 9, 371.
- NALDI, M., FIORI, J., PISTOLOZZI, M., DRAKE, A. F., BERTUCCI, C., WU, R., MLYNARCZYK, K., FILIPEK, S., DE SIMONE, A. & ANDRISANO, V. 2012. Amyloid β -peptide 25-35 self-assembly and its inhibition: a model undecapeptide system to gain atomistic and

- secondary structure details of the Alzheimer's disease process and treatment. *ACS chemical neuroscience*, 3, 952-962.
- NANDY, K., BABU, D., RANI, S., JOSHI, G., IJEE, S., GEORGE, A., PALANI, D., PREMKUMAR, C., RAJESH, P., VIJAYANAND, S., DAVID, E., MURUGESAN, M. & VELAYUDHAN, S. R. 2023. Efficient gene editing in induced pluripotent stem cells enabled by an inducible adenine base editor with tunable expression. *Scientific Reports*, 13, 21953.
- NARAYAN, P., ORTE, A., CLARKE, R. W., BOLOGNESI, B., HOOK, S., GANZINGER, K. A., MEEHAN, S., WILSON, M. R., DOBSON, C. M. & KLENERMAN, D. 2012. The extracellular chaperone clusterin sequesters oligomeric forms of the amyloid- β 1–40 peptide. *Nature Structural & Molecular Biology*, 19, 79-83.
- NEWBERRY, E. P., LATIFI, T., BATTAILE, J. T. & TOWLER, D. A. 1997. Structure–Function Analysis of Msx2-Mediated Transcriptional Suppression. *Biochemistry*, 36, 10451-10462.
- NGUYEN, D. P., MIYAOKA, Y., GILBERT, L. A., MAYERL, S. J., LEE, B. H., WEISSMAN, J. S., CONKLIN, B. R. & WELLS, J. A. 2016. Ligand-binding domains of nuclear receptors facilitate tight control of split CRISPR activity. *Nature Communications*, 7, 12009.
- NICHOLAS, CORY R., CHEN, J., TANG, Y., SOUTHWELL, DEREK G., CHALMERS, N., VOGT, D., ARNOLD, CHRISTINE M., CHEN, Y.-JIUN J., STANLEY, EDOUARD G., ELEFANTY, ANDREW G., SASAI, Y., ALVAREZ-BUYLLA, A., RUBENSTEIN, JOHN L. R. & KRIEGSTEIN, ARNOLD R. 2013. Functional Maturation of hPSC-Derived Forebrain Interneurons Requires an Extended Timeline and Mimics Human Neural Development. *Cell Stem Cell*, 12, 573-586.
- NIELSEN, H. M., MULDER, S. D., BELIËN, J. A. M., MUSTERS, R. J. P., EIKELENBOOM, P. & VEERHUIS, R. 2010. Astrocytic A β 1-42 uptake is determined by A β -aggregation state and the presence of amyloid-associated proteins. *Glia*, 58, 1235-1246.
- NISSIM, L., PERLI, SAMUEL D., FRIDKIN, A., PEREZ-PINERA, P. & LU, TIMOTHY K. 2014. Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells. *Molecular Cell*, 54, 698-710.
- NIWA, H., YAMAMURA, K. & MIYAZAKI, J. 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene*, 108, 193-199.
- NUUTINEN, T., HUUSKONEN, J., SUURONEN, T., OJALA, J., MIETTINEN, R. & SALMINEN, A. 2007. Amyloid- β 1–42 induced endocytosis and clusterin/apoJ protein accumulation in cultured human astrocytes. *Neurochemistry International*, 50, 540-547.
- O'BRIEN, R. J. & WONG, P. C. 2011. Amyloid precursor protein processing and Alzheimer's disease. *Annu Rev Neurosci*, 34, 185-204.
- OCEGUERA-YANEZ, F., KIM, S.-I., MATSUMOTO, T., TAN, G. W., XIANG, L., HATANI, T., KONDO, T., IKEYA, M., YOSHIDA, Y., INOUE, H. & WOLTJEN, K. 2016. Engineering the AAVS1 locus for consistent and scalable transgene expression in human iPSCs and their differentiated derivatives. *Methods*, 101, 43-55.
- OH, Y. & JANG, J. 2019. Directed Differentiation of Pluripotent Stem Cells by Transcription Factors. *Mol Cells*, 42, 200-209.
- OKADA, Y., HOSOI, N., MATSUZAKI, Y., FUKAI, Y., HIRAGA, A., NAKAI, J., NITTA, K., SHINOHARA, Y., KONNO, A. & HIRAI, H. 2022. Development of microglia-targeting adeno-associated viral vectors as tools to study microglial behavior in vivo. *Communications Biology*, 5, 1224.
- ORDOVÁS, L., BOON, R., PISTONI, M., CHEN, Y., WOLFS, E., GUO, W., SAMBATHKUMAR, R., BOBIS-WOZOWICZ, S., HELSEN, N., VANHOVE, J., BERCKMANS, P., CAI, Q., VANUYTSEL, K., EGGERMONT, K., VANSLEMBROUCK, V., SCHMIDT, B. Z., RAITANO, S., VAN DEN BOSCH, L., NAHMIAS, Y., CATHOMEN, T., STRUYS, T. & VERFAILLIE, C. M. 2015. Efficient Recombinase-Mediated Cassette Exchange in hPSCs to Study the Hepatocyte Lineage Reveals AAVS1 Locus-Mediated Transgene Inhibition. *Stem Cell Reports*, 5, 918-931.

- ORRE, M., KAMPHUIS, W., OSBORN, L. M., MELIEF, J., KOOIJMAN, L., HUITINGA, I., KLOOSTER, J., BOSSERS, K. & HOL, E. M. 2014. Acute isolation and transcriptome characterization of cortical astrocytes and microglia from young and aged mice. *Neurobiology of Aging*, 35, 1-14.
- OVCHINNIKOV, D. A., KORN, O., VIRSHUP, I., WELLS, C. A. & WOLVETANG, E. J. 2018. The Impact of APP on Alzheimer-like Pathogenesis and Gene Expression in Down Syndrome iPSC-Derived Neurons. *Stem Cell Reports*, 11, 32-42.
- OVCHINNIKOV, D. A., TITMARSH, D. M., FORTUNA, P. R. J., HIDALGO, A., ALHARBI, S., WHITWORTH, D. J., COOPER-WHITE, J. J. & WOLVETANG, E. J. 2014. Transgenic human ES and iPS reporter cell lines for identification and selection of pluripotent stem cells in vitro. *Stem Cell Research*, 13, 251-261.
- PAQUET, D., KWART, D., CHEN, A., SPROUL, A., JACOB, S., TEO, S., OLSEN, K. M., GREGG, A., NOGGLE, S. & TESSIER-LAVIGNE, M. 2016. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature*, 533, 125-129.
- PARK, J., WETZEL, I., MARRIOTT, I., DRÉAU, D., D'AVANZO, C., KIM, D. Y., TANZI, R. E. & CHO, H. 2018. A 3D human triculture system modeling neurodegeneration and neuroinflammation in Alzheimer's disease. *Nature Neuroscience*, 21, 941-951.
- PARK, S. B., UCHIDA, T., TILSON, S., HU, Z., MA, C. D., LEEK, M., EICHNER, M., HONG, S. G. & LIANG, T. J. 2022. A dual conditional CRISPR-Cas9 system to activate gene editing and reduce off-target effects in human stem cells. *Molecular Therapy - Nucleic Acids*, 28, 656-669.
- PARKS, T. D., HOWARD, E. D., WOLPERT, T. J., ARP, D. J. & DOUGHERTY, W. G. 1995. Expression and Purification of a Recombinant Tobacco Etch Virus Nla Proteinase: Biochemical Analyses of the Full-Length and a Naturally Occurring Truncated Proteinase Form. *Virology*, 210, 194-201.
- PAVANI, G. & AMENDOLA, M. 2021. Targeted Gene Delivery: Where to Land. *Frontiers in Genome Editing*, 2.
- PEABODY, D. S. 1993. The RNA binding site of bacteriophage MS2 coat protein. *The EMBO Journal*, 12, 595-600.
- PENG, L., YU, Y., LIU, J., LI, S., HE, H., CHENG, N. & YE, R. D. 2015. The chemerin receptor CMKLR1 is a functional receptor for amyloid-beta peptide. *J Alzheimers Dis*, 43, 227-42.
- PENNEY, J., RALVENIUS, W. T. & TSAI, L.-H. 2020. Modeling Alzheimer's disease with iPSC-derived brain cells. *Molecular Psychiatry*, 25, 148-167.
- PHAM, M. T., POLLOCK, K. M., ROSE, M. D., CARY, W. A., STEWART, H. R., ZHOU, P., NOLTA, J. A. & WALDAU, B. 2018. Generation of human vascularized brain organoids. *NeuroReport*, 29.
- PIRSICOVEANU, D. F. V., PIRICI, I., TUDORICA, V., BALSEANU, T. A., ALBU, V. C., BONDARI, S., BUMBEA, A. M. & PIRSCOVEANU, M. 2017. Tau protein in neurodegenerative diseases - a review. *Rom J Morphol Embryol*, 58, 1141-1150.
- POOLER, A. M., POLYDORO, M., MAURY, E. A., NICHOLLS, S. B., REDDY, S. M., WEGMANN, S., WILLIAM, C., SAQRAN, L., CAGSAL-GETKIN, O., PITSTICK, R., BEIER, D. R., CARLSON, G. A., SPIRES-JONES, T. L. & HYMAN, B. T. 2015. Amyloid accelerates tau propagation and toxicity in a model of early Alzheimer's disease. *Acta Neuropathologica Communications*, 3, 14.
- PORT, F. & BULLOCK, S. L. 2016. Augmenting CRISPR applications in Drosophila with tRNA-flanked sgRNAs. *Nature Methods*, 13, 852-854.
- PORTEUS, M. 2016. Genome Editing: A New Approach to Human Therapeutics. *Annu Rev Pharmacol Toxicol*, 56, 163-90.
- PRAHARAJ, P. P., PATRA, S., MISHRA, S. R., MUKHOPADHYAY, S., KLIONSKY, D. J., PATIL, S. & BHUTIA, S. K. 2023. CLU (clusterin) promotes mitophagic degradation of MSX2 through

- an AKT-DNM1L/Drp1 axis to maintain SOX2-mediated stemness in oral cancer stem cells. *Autophagy*, 19, 2196-2216.
- PURNICK, P. E. M. & WEISS, R. 2009. The second wave of synthetic biology: from modules to systems. *Nature Reviews Molecular Cell Biology*, 10, 410-422.
- QI, L., HAURWITZ, R. E., SHAO, W., DOUDNA, J. A. & ARKIN, A. P. 2012. RNA processing enables predictable programming of gene expression. *Nature Biotechnology*, 30, 1002-1006.
- QI, LEI S., LARSON, MATTHEW H., GILBERT, LUKE A., DOUDNA, JENNIFER A., WEISSMAN, JONATHAN S., ARKIN, ADAM P. & LIM, WENDELL A. 2013. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*, 152, 1173-1183.
- QI, W., ZHU, T., TIAN, Z., LI, C., ZHANG, W. & SONG, R. 2016. High-efficiency CRISPR/Cas9 multiplex gene editing using the glycine tRNA-processing system-based strategy in maize. *BMC Biotechnology*, 16, 58.
- RAINA, A. K., HOCHMAN, A., ICKES, H., ZHU, X., OGAWA, O., CASH, A. D., SHIMOHAMA, S., PERRY, G. & SMITH, M. A. 2003. Apoptotic promoters and inhibitors in Alzheimer's disease: Who wins out? *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 27, 251-254.
- RAMACHANDRA, C. J., SHAHBAZI, M., KWANG, T. W., CHOUDHURY, Y., BAK, X. Y., YANG, J. & WANG, S. 2011. Efficient recombinase-mediated cassette exchange at the AAVS1 locus in human embryonic stem cells using baculoviral vectors. *Nucleic Acids Res*, 39, e107.
- REBUZZINI, P., ZUCCOTTI, M., REDI, C. A. & GARAGNA, S. 2016. Achilles' heel of pluripotent stem cells: genetic, genomic and epigenetic variations during prolonged culture. *Cellular and Molecular Life Sciences*, 73, 2453-2466.
- REIMAN, E. M., ARBOLEDA-VELASQUEZ, J. F., QUIROZ, Y. T., HUENTELMAN, M. J., BEACH, T. G., CASELLI, R. J., CHEN, Y., SU, Y., MYERS, A. J., HARDY, J., PAUL VONSATTEL, J., YOUNKIN, S. G., BENNETT, D. A., DE JAGER, P. L., LARSON, E. B., CRANE, P. K., KEENE, C. D., KAMBOH, M. I., KOFLER, J. K., DUQUE, L., GILBERT, J. R., GWIRTSMAN, H. E., BUXBAUM, J. D., DICKSON, D. W., FROSCHE, M. P., GHETTI, B. F., LUNETTA, K. L., WANG, L.-S., HYMAN, B. T., KUKULL, W. A., FOROUD, T., HAINES, J. L., MAYEUX, R. P., PERICAK-VANCE, M. A., SCHNEIDER, J. A., TROJANOWSKI, J. Q., FARRER, L. A., SCHELLENBERG, G. D., BEECHAM, G. W., MONTINE, T. J., JUN, G. R., ABNER, E., ADAMS, P. M., ALBERT, M. S., ALBIN, R. L., APOSTOLOVA, L. G., ARNOLD, S. E., ASTHANA, S., ATWOOD, C. S., BALDWIN, C. T., BARBER, R. C., BARNES, L. L., BARRAL, S., BECKER, J. T., BEEKLY, D., BIGIO, E. H., BIRD, T. D., BLACKER, D., BOEVE, B. F., BOWEN, J. D., BOXER, A., BURKE, J. R., BURNS, J. M., CAIRNS, N. J., CANTWELL, L. B., CAO, C., CARLSON, C. S., CARLSSON, C. M., CARNEY, R. M., CARRASQUILLO, M. M., CHUI, H. C., CRIBBS, D. H., CROCCO, E. A., CRUCHAGA, C., DECARLI, C., DICK, M., DOODY, R. S., DUARA, R., ERTEKIN-TANER, N., EVANS, D. A., FABER, K. M., FAIRCHILD, T. J., FALLON, K. B., FARDO, D. W., FARLOW, M. R., FERRIS, S., GALASKO, D. R., GEARING, M., GESCHWIND, D. H., GHISAYS, V., GOATE, A. M., GRAFF-RADFORD, N. R., GREEN, R. C., GROWDON, J. H., HAKONARSON, H., HAMILTON, R. L., HAMILTON-NELSON, K. L., HARRELL, L. E., HONIG, L. S., HUEBINGER, R. M., et al. 2020. Exceptionally low likelihood of Alzheimer's dementia in APOE2 homozygotes from a 5,000-person neuropathological study. *Nature Communications*, 11, 667.
- REYNOLDS, B. A. & WEISS, S. 1992. Generation of Neurons and Astrocytes from Isolated Cells of the Adult Mammalian Central Nervous System. *Science*, 255, 1707-1710.
- ROBBINS, J. P., PERFECT, L., RIBE, E. M., MARESCA, M., DANGLA-VALLS, A., FOSTER, E. M., KILLICK, R., NOWOSIAD, P., REID, M. J., POLIT, L. D., NEVADO, A. J., EBNER, D., BOHLOOLY-Y, M., BUCKLEY, N., PANGALOS, M. N., PRICE, J. & LOVESTONE, S. 2018. Clusterin Is Required for β -Amyloid Toxicity in Human iPSC-Derived Neurons. *Frontiers in Neuroscience*, 12.

- ROSA, F., DHINGRA, A., UYSAL, B., MENDIS, G. D. C., LOEFFLER, H., ELSEIN, G., MUELLER, S., SCHWARZ, N., CASTILLO-LIZARDO, M., CUDDY, C., BECKER, F., HEUTINK, P., REID, C. A., PETROU, S., LERCHE, H. & MALJEVIC, S. 2020. In Vitro Differentiated Human Stem Cell-Derived Neurons Reproduce Synaptic Synchronicity Arising during Neurodevelopment. *Stem Cell Reports*, 15, 22-37.
- ROSENBAUM, D. M., RASMUSSEN, S. G. F. & KOBILKA, B. K. 2009. The structure and function of G-protein-coupled receptors. *Nature*, 459, 356-363.
- RÖSSGER, K., CHARPIN-EL-HAMRI, G. & FUSSENEGGER, M. 2013. A closed-loop synthetic gene circuit for the treatment of diet-induced obesity in mice. *Nature Communications*, 4, 2825.
- ROTH, B. L. 2013. Impossible or Merely Difficult? Two Grand Challenges from a Biologist's Perspective. *ACS Medicinal Chemistry Letters*, 4, 316-318.
- RYDING, A. D., SHARP, M. G. & MULLINS, J. J. 2001. Conditional transgenic technologies. *Journal of Endocrinology*, 171, 1-14.
- RYTER, S. W., ALAM, J. & CHOI, A. M. K. 2006. Heme Oxygenase-1/Carbon Monoxide: From Basic Science to Therapeutic Applications. *Physiological Reviews*, 86, 583-650.
- SALEHI, A., DELCROIX, J.-D. & MOBLEY, W. C. 2003. Traffic at the intersection of neurotrophic factor signaling and neurodegeneration. *Trends in Neurosciences*, 26, 73-80.
- SANCHEZ, M. I. & TING, A. Y. 2020. Directed evolution improves the catalytic efficiency of TEV protease. *Nat Methods*, 17, 167-174.
- SAUNDERS, A. M., ROSES, A. D., PERICAK-VANCE, M. A., DOLE, K. C., STRITTMATTER, W. J., SCHMECHEL, D. E., SZYMANSKI, M. H., MCCOWN, N., MANWARING, M. G., SCHMADER, K., BREITNER, J. C. S., GOLDGABER, D., BENSON, M. D., GOLDFARB, L. & BROWN, W. T. 1993. Apolipoprotein E ϵ 4 allele distributions in late-onset Alzheimer's disease and in other amyloid-forming diseases. *The Lancet*, 342, 710-711.
- SCHNEIDER, L. S. 2022. Aducanumab Trials EMERGE But Don't ENGAGE. *The Journal of Prevention of Alzheimer's Disease*, 9, 193-196.
- SECZYNSKA, M., BLOOR, S., CUESTA, S. M. & LEHNER, P. J. 2022. Genome surveillance by HUSH-mediated silencing of intronless mobile elements. *Nature*, 601, 440-445.
- SELKOE, D. J. & HARDY, J. 2016. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol Med*, 8, 595-608.
- SHAH, S., WONG, L. M., ELLIS, K., BODNAR, B., SARIBAS, S., TING, J., WEI, Z., TANG, Y., WANG, X., WANG, H., LING, B., MARGOLIS, D. M., GARCIA, J. V., HU, W. & JIANG, G. 2022. Microglia-Specific Promoter Activities of HEXB Gene. *Frontiers in Cellular Neuroscience*, 16.
- SHAW, W. M., STUDENÁ, L., ROY, K., HAPETA, P., MCCARTY, N. S., GRAHAM, A. E., ELLIS, T. & LEDESMA-AMARO, R. 2022. Inducible expression of large gRNA arrays for multiplexed CRISPRai applications. *Nature Communications*, 13, 4984.
- SHI, Y., KIRWAN, P. & LIVESEY, F. J. 2012. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nature Protocols*, 7, 1836-1846.
- SHIMIZU, T., TANAKA, T., ISO, T., MATSUI, H., OYAMA, Y., KAWAI-KOWASE, K., ARAI, M. & KURABAYASHI, M. 2011. Notch Signaling Pathway Enhances Bone Morphogenetic Protein 2 (BMP2) Responsiveness of Msx2 Gene to Induce Osteogenic Differentiation and Mineralization of Vascular Smooth Muscle Cells*. *Journal of Biological Chemistry*, 286, 19138-19148.
- SHOJI, E., WOLTJEN, K. & SAKURAI, H. 2016. Directed Myogenic Differentiation of Human Induced Pluripotent Stem Cells. In: NAGY, A. & TURKSEN, K. (eds.) *Patient-Specific Induced Pluripotent Stem Cell Models: Generation and Characterization*. New York, NY: Springer New York.
- SIMS, J. R., ZIMMER, J. A., EVANS, C. D., LU, M., ARDAYFIO, P., SPARKS, J., WESSELS, A. M., SHCHERBININ, S., WANG, H., MONKUL NERY, E. S., COLLINS, E. C., SOLOMON, P.,

- SALLOWAY, S., APOSTOLOVA, L. G., HANSSON, O., RITCHIE, C., BROOKS, D. A., MINTUN, M. & SKOVRONSKY, D. M. 2023. Donanemab in Early Symptomatic Alzheimer Disease: The TRAILBLAZER-ALZ 2 Randomized Clinical Trial. *Jama*, 330, 512-527.
- SLOAN, S. A., DARMANIS, S., HUBER, N., KHAN, T. A., BIREY, F., CANEDA, C., REIMER, R., QUAKE, S. R., BARRES, B. A. & PAŞCA, S. P. 2017. Human Astrocyte Maturation Captured in 3D Cerebral Cortical Spheroids Derived from Pluripotent Stem Cells. *Neuron*, 95, 779-790.e6.
- SMITH, A. K. & KLIMOV, D. K. 2019. De novo aggregation of Alzheimer's A β 25-35 peptides in a lipid bilayer. *Scientific Reports*, 9, 7161.
- SMITH, J. R., MAGUIRE, S., DAVIS, L. A., ALEXANDER, M., YANG, F., CHANDRAN, S., FFRENCH-CONSTANT, C. & PEDERSEN, R. A. 2007. Robust, Persistent Transgene Expression in Human Embryonic Stem Cells Is Achieved with AAVS1-Targeted Integration. *Stem Cells*, 26, 496-504.
- SMITH, J. R., VALLIER, L., LUPO, G., ALEXANDER, M., HARRIS, W. A. & PEDERSEN, R. A. 2008. Inhibition of Activin/Nodal signaling promotes specification of human embryonic stem cells into neuroectoderm. *Developmental Biology*, 313, 107-117.
- STINSON, B. M. & LOPARO, J. J. 2021. Repair of DNA Double-Strand Breaks by the Nonhomologous End Joining Pathway. *Annu Rev Biochem*, 90, 137-164.
- STONE, J. 2008. What initiates the formation of senile plaques? The origin of Alzheimer-like dementias in capillary haemorrhages. *Medical Hypotheses*, 71, 347-359.
- SUBRAMANIAN, A., TAMAYO, P., MOOTHA, V. K., MUKHERJEE, S., EBERT, B. L., GILLETTE, M. A., PAULOVICH, A., POMEROY, S. L., GOLUB, T. R., LANDER, E. S. & MESIROV, J. P. 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102, 15545-15550.
- SUN, N., PETIWALA, S., WANG, R., LU, C., HU, M., GHOSH, S., HAO, Y., MILLER, C. P. & CHUNG, N. 2019. Development of drug-inducible CRISPR-Cas9 systems for large-scale functional screening. *BMC Genomics*, 20, 225.
- SUPHARATTANASITTHI, W., CARLSSON, E., SHARIF, U. & PARAOAN, L. 2019. CRISPR/Cas9-mediated one step bi-allelic change of genomic DNA in iPSCs and human RPE cells in vitro with dual antibiotic selection. *Scientific Reports*, 9, 174.
- SZADE, A., SZADE, K., MAHDI, M. & JÓZKOWICZ, A. 2021. The role of heme oxygenase-1 in hematopoietic system and its microenvironment. *Cellular and Molecular Life Sciences*, 78, 4639-4651.
- TAGLIAFIERRO, L., ZAMORA, M. E. & CHIBA-FALEK, O. 2019. Multiplication of the SNCA locus exacerbates neuronal nuclear aging. *Human Molecular Genetics*, 28, 407-421.
- TAKAHASHI, K., TANABE, K., OHNUKI, M., NARITA, M., ICHISAKA, T., TOMODA, K. & YAMANAKA, S. 2007. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*, 131, 861-872.
- TAKAHASHI, K. & YAMANAKA, S. 2006. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, 126, 663-676.
- TAKAMATSU, K., IKEDA, T., HARUTA, M., MATSUMURA, K., OGI, Y., NAKAGATA, N., UCHINO, M., ANDO, Y., NISHIMURA, Y. & SENJU, S. 2014. Degradation of amyloid beta by human induced pluripotent stem cell-derived macrophages expressing Neprilysin-2. *Stem Cell Research*, 13, 442-453.
- TANENBAUM, M. E., GILBERT, L. A., QI, L. S., WEISSMAN, J. S. & VALE, R. D. 2014. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*, 159, 635-46.
- TAYLOR, A., MACAULAY, V. M., MAURYA, A. K., MIOSSEC, M. J. & BUFFA, F. M. 2023. GeneFEAST: the pivotal, gene-centric step in functional enrichment analysis interpretation.

Available: <https://ui.adsabs.harvard.edu/abs/2023arXiv230900061T> [Accessed August 01, 2023].

- TCW, J. & GOATE, A. M. 2017. Genetics of β -Amyloid Precursor Protein in Alzheimer's Disease. *Cold Spring Harb Perspect Med*, 7.
- TEIXEIRA, A. P. & FUSSENEGGER, M. 2023. Synthetic Gene Circuits for Regulation of Next-Generation Cell-Based Therapeutics. *Advanced Science*, n/a, 2309088.
- THATHIAH, A. & DE STROOPER, B. 2011. The role of G protein-coupled receptors in the pathology of Alzheimer's disease. *Nature Reviews Neuroscience*, 12, 73-87.
- TIAN, R., ABARIENTOS, A., HONG, J., HASHEMI, S. H., YAN, R., DRÄGER, N., LENG, K., NALLS, M. A., SINGLETON, A. B., XU, K., FAGHRI, F. & KAMPMANN, M. 2021. Genome-wide CRISPRi/a screens in human neurons link lysosomal failure to ferroptosis. *Nature Neuroscience*, 24, 1020-1034.
- TIAN, R., GACHECHILADZE, M. A., LUDWIG, C. H., LAURIE, M. T., HONG, J. Y., NATHANIEL, D., PRABHU, A. V., FERNANDOPULLE, M. S., PATEL, R., ABSHARI, M., WARD, M. E. & KAMPMANN, M. 2019. CRISPR Interference-Based Platform for Multimodal Genetic Screens in Human iPSC-Derived Neurons. *Neuron*, 104, 239-255.e12.
- TIROZZI, A., QUICCIONE, M. S., CERLETTI, C., DONATI, M. B., DE GAETANO, G., IACOVIELLO, L. & GIALLUISI, A. 2023. A Multi-Trait Association Analysis of Brain Disorders and Platelet Traits Identifies Novel Susceptibility Loci for Major Depression, Alzheimer's and Parkinson's Disease. *Cells* [Online], 12.
- TOGO, K., FUKUSUMI, H., SHOFUDA, T., OHNISHI, H., YAMAZAKI, H., HAYASHI, M. K., KAWASAKI, N., TAKEI, N., NAKAZAWA, T., SAITO, Y., BABA, K., HASHIMOTO, H., SEKINO, Y., SHIRAO, T., MOCHIZUKI, H. & KANEMURA, Y. 2021. Postsynaptic structure formation of human iPS cell-derived neurons takes longer than presynaptic formation during neural differentiation in vitro. *Molecular Brain*, 14, 149.
- URLINGER, S., BARON, U., THELLMANN, M., HASAN, M. T., BUJARD, H. & HILLEN, W. 2000. Exploring the sequence space for tetracycline-dependent transcriptional activators: Novel mutations yield expanded range and sensitivity. *Proceedings of the National Academy of Sciences*, 97, 7963-7968.
- VALENZUELA, D., ECONOMIDES, A., ROJAS, E., LAMB, T., NUNEZ, L., JONES, P., LP, N., R ESPINOSA, R., BRANNAN, C. & GILBERT, D. 1995. Identification of mammalian noggin and its expression in the adult nervous system. *The Journal of Neuroscience*, 15, 6077-6084.
- VALIPOUR, B., SIMORGH, S., MIRSALEHI, M., MORADI, S., TAGHIZADEH-HESARY, F., SEIDKHANI, E., AKBARNEJAD, Z. & ALIZADEH, R. 2024. Improvement of spatial learning and memory deficits by intranasal administration of human olfactory ecto-mesenchymal stem cells in an Alzheimer's disease rat model. *Brain Research*, 1828, 148764.
- VAN DYCK, C. H., SWANSON, C. J., AISEN, P., BATEMAN, R. J., CHEN, C., GEE, M., KANEKIYO, M., LI, D., REYDERMAN, L., COHEN, S., FROELICH, L., KATAYAMA, S., SABBAGH, M., VELLAS, B., WATSON, D., DHADDA, S., IRIZARRY, M., KRAMER, L. D. & IWATSUBO, T. 2023. Lecanemab in Early Alzheimer's Disease. *N Engl J Med*, 388, 9-21.
- VETRIVEL, K. S., ZHANG, Y.-W., XU, H. & THINAKARAN, G. 2006. Pathological and physiological functions of presenilins. *Molecular Neurodegeneration*, 1, 4.
- VILLAR-GÓMEZ, N., OJEDA-HERNANDEZ, D. D., LÓPEZ-MUGURUZA, E., GARCÍA-FLORES, S., BONEL-GARCÍA, N., BENITO-MARTÍN, M. S., SELMA-CALVO, B., CANALES-AGUIRRE, A. A., MATEOS-DÍAZ, J. C., MONTERO-ESCRIBANO, P., MATIAS-GUIU, J. A., MATÍAS-GUIU, J. & GÓMEZ-PINEDO, U. 2022. Nose-to-Brain: The Next Step for Stem Cell and Biomaterial Therapy in Neurological Disorders. *Cells*, 11.

- WADHWANI, A. R., AFFANEH, A., VAN GULDEN, S. & KESSLER, J. A. 2019. Neuronal apolipoprotein E4 increases cell death and phosphorylated tau release in alzheimer disease. *Ann Neurol*, 85, 726-739.
- WALKER, L. C. 2020. A β Plaques. *Free Neuropathol*, 1.
- WANG, G., YANG, L., GRISHIN, D., RIOS, X., YE, L. Y., HU, Y., LI, K., ZHANG, D., CHURCH, G. M. & PU, W. T. 2017. Efficient, footprint-free human iPSC genome editing by consolidation of Cas9/CRISPR and piggyBac technologies. *Nature Protocols*, 12, 88-103.
- WANG, H., GUO, R., DU, Z., BAI, L., LI, L., CUI, J., LI, W., HOFFMAN, A. R. & HU, J.-F. 2018. Epigenetic Targeting of Granulin in Hepatoma Cells by Synthetic CRISPR dCas9 Epi-suppressors. *Molecular Therapy - Nucleic Acids*, 11, 23-33.
- WANG, H., WANG, M., WANG, Y., WEN, Y., CHEN, X., WU, D., SU, P., ZHOU, W., SHI, L. & ZHOU, J. 2020. MSX2 suppression through inhibition of TGF β signaling enhances hematopoietic differentiation of human embryonic stem cells. *Stem Cell Research & Therapy*, 11, 147.
- WEI, Y., QIU, Y., CHEN, Y., LIU, G., ZHANG, Y., XU, L. & DING, Q. 2017. CRISPR/Cas9 with single guide RNA expression driven by small tRNA promoters showed reduced editing efficiency compared to a U6 promoter. *Rna*, 23, 1-5.
- WILUSZ, J. E., JNBAPTISTE, C. K., LU, L. Y., KUHN, C. D., JOSHUA-TOR, L. & SHARP, P. A. 2012. A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev*, 26, 2392-407.
- WITTAMER, V. R., FRANSSSEN, J.-D., VULCANO, M., MIRJOLET, J.-F. O., LE POUL, E., MIGEOTTE, I., BRÉZILLON, S. P., TYLDESLEY, R., BLANPAIN, C. D., DETHEUX, M., MANTOVANI, A., SOZZANI, S., VASSART, G., PARMENTIER, M. & COMMUNI, D. 2003. Specific Recruitment of Antigen-presenting Cells by Chemerin, a Novel Processed Ligand from Human Inflammatory Fluids. *Journal of Experimental Medicine*, 198, 977-985.
- WOJTAS, A. M., KANG, S. S., OLLEY, B. M., GATHERER, M., SHINOHARA, M., LOZANO, P. A., LIU, C.-C., KURTI, A., BAKER, K. E., DICKSON, D. W., YUE, M., PETRUCCELLI, L., BU, G., CARARE, R. O. & FRYER, J. D. 2017. Loss of clusterin shifts amyloid deposition to the cerebrovasculature via disruption of perivascular drainage pathways. *Proceedings of the National Academy of Sciences*, 114, E6962-E6971.
- WOLFE, B. L. & TREJO, J. 2007. Clathrin-Dependent Mechanisms of G Protein-coupled Receptor Endocytosis. *Traffic*, 8, 462-470.
- WOODRUFF, G., YOUNG, JESSICA E., MARTINEZ, FERNANDO J., BUEN, F., GORE, A., KINAGA, J., LI, Z., YUAN, SHAUNA H., ZHANG, K. & GOLDSTEIN, LAWRENCE S. B. 2013. The Presenilin-1 Δ E9 Mutation Results in Reduced γ -Secretase Activity, but Not Total Loss of PS1 Function, in Isogenic Human Stem Cells. *Cell Reports*, 5, 974-985.
- WU, D., CHEN, Q., CHEN, X., HAN, F., CHEN, Z. & WANG, Y. 2023. The blood–brain barrier: structure, regulation, and drug delivery. *Signal Transduction and Targeted Therapy*, 8, 217.
- WU, Q., ZHANG, L., SU, P., LEI, X., LIU, X., WANG, H., LU, L., BAI, Y., XIONG, T., LI, D., ZHU, Z., DUAN, E., JIANG, E., FENG, S., HAN, M., XU, Y., WANG, F. & ZHOU, J. 2015. MSX2 mediates entry of human pluripotent stem cells into mesendoderm by simultaneously suppressing SOX2 and activating NODAL signaling. *Cell Research*, 25, 1314-1332.
- WU, Z., YANG, H. & COLOSI, P. 2010. Effect of genome size on AAV vector packaging. *Mol Ther*, 18, 80-6.
- XIAO, Z., CILZ, N. I., KURADA, L., HU, B., YANG, C., WADA, E., COMBS, C. K., PORTER, J. E., LESAGE, F. & LEI, S. 2014. Activation of Neurotensin Receptor 1 Facilitates Neuronal Excitability and Spatial Learning and Memory in the Entorhinal Cortex: Beneficial Actions in an Alzheimer's Disease Model. *The Journal of Neuroscience*, 34, 7027-7042.
- XIE, K., MINKENBERG, B. & YANG, Y. 2015. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proceedings of the National Academy of Sciences*, 112, 3570-3575.

- XU, L., ZHAO, L., GAO, Y., XU, J. & HAN, R. 2016. Empower multiplex cell and tissue-specific CRISPR-mediated gene manipulation with self-cleaving ribozymes and tRNA. *Nucleic Acids Research*, 45, e28-e28.
- YANG, J., LI, X., LI, Y., SOUTHWOOD, M., YE, L., LONG, L., AL-LAMKI, R. S. & MORRELL, N. W. 2013. Id proteins are critical downstream effectors of BMP signaling in human pulmonary arterial smooth muscle cells. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 305, L312-L321.
- YERBURY, J. J. & WILSON, M. R. 2010. Extracellular chaperones modulate the effects of Alzheimer's patient cerebrospinal fluid on A β 1-42 toxicity and uptake. *Cell Stress and Chaperones*, 15, 115-121.
- YI, F., AMARASINGHE, B. & DANG, T. P. 2013. Manic fringe inhibits tumor growth by suppressing Notch3 degradation in lung cancer. *Am J Cancer Res*, 3, 490-9.
- YIN, J.-A., FRICK, L., SCHEIDMANN, M. C., LIU, T., TREVISAN, C., DHINGRA, A., SPINELLI, A., WU, Y., YAO, L., VENA, D. L., KNAPP, B., CECCO, E. D., GING, K., ARMANI, A., OAKELEY, E., NIGSCH, F., JENZER, J., HAEGELE, J., PIKUSA, M., TÄGER, J., RODRIGUEZ-NIETO, S., GUO, J., BOURIS, V., RIBEIRO, R., BARONI, F., BEDI, M. S., BERRY, S., LOSA, M., HORNEMANN, S., KAMPMANN, M., PELKMANS, L., HOEPFNER, D., HEUTINK, P. & AGUZZI, A. 2023. Robust and Versatile Arrayed Libraries for Human Genome-Wide CRISPR Activation, Deletion and Silencing. *bioRxiv*, 2022.05.25.493370.
- YOSHIHARA, M., HAYASHIZAKI, Y. & MURAKAWA, Y. 2017. Genomic Instability of iPSCs: Challenges Towards Their Clinical Applications. *Stem Cell Reviews and Reports*, 13, 7-16.
- YOUNG, J. E., FONG, L. K., FRANKOWSKI, H., PETSKE, G. A., SMALL, S. A. & GOLDSTEIN, L. S. B. 2018. Stabilizing the Retromer Complex in a Human Stem Cell Model of Alzheimer's Disease Reduces TAU Phosphorylation Independently of Amyloid Precursor Protein. *Stem Cell Reports*, 10, 1046-1058.
- YU, B., JIN, Y., SHEN, Y., YANG, Y., WANG, G., ZHU, H., YU, Y. & WANG, J. 2019. Loss of homeoprotein Msx1 and Msx2 leading to athletic and kinematic impairment related to the increasing neural excitability of neurons in aberrant neocortex in mice. *Biochemical and Biophysical Research Communications*, 516, 229-235.
- YU, J., VODYANIK, M. A., SMUGA-OTTO, K., ANTOSIEWICZ-BOURGET, J., FRANE, J. L., TIAN, S., NIE, J., JONSDOTTIR, G. A., RUOTTI, V., STEWART, R., SLUKVIN, I. I. & THOMSON, J. A. 2007. Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science*, 318, 1917-1920.
- YU, Y. & YE, R. D. 2015. Microglial A β Receptors in Alzheimer's Disease. *Cellular and Molecular Neurobiology*, 35, 71-83.
- YUE, W., LI, Y., ZHANG, T., JIANG, M., QIAN, Y., ZHANG, M., SHENG, N., FENG, S., TANG, K., YU, X., SHU, Y., YUE, C. & JING, N. 2015. ESC-Derived Basal Forebrain Cholinergic Neurons Ameliorate the Cognitive Symptoms Associated with Alzheimer's Disease in Mouse Models. *Stem Cell Reports*, 5, 776-790.
- YUI, D., NISHIDA, Y., NISHINA, T., MOGUSHI, K., TAJIRI, M., ISHIBASHI, S., AJIOKA, I., ISHIKAWA, K., MIZUSAWA, H., MURAYAMA, S. & YOKOTA, T. 2015. Enhanced Phospholipase A2 Group 3 Expression by Oxidative Stress Decreases the Insulin-Degrading Enzyme. *PLOS ONE*, 10, e0143518.
- ZANDL-LANG, M., FANAEE-DANESH, E., SUN, Y., ALBRECHER, N. M., GALI, C. C., ČANČAR, I., KOBER, A., TAM-AMERSDORFER, C., STRACKE, A., STORCK, S. M., SAEED, A., STEFULJ, J., PIETRZIK, C. U., WILSON, M. R., BJÖRKHEM, I. & PANZENBOECK, U. 2018. Regulatory effects of simvastatin and apoJ on APP processing and amyloid- β clearance in blood-brain barrier endothelial cells. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1863, 40-60.

- ZEGHAL, M., LAROCHE, G., FREITAS, J. D., WANG, R. & GIGUÈRE, P. M. 2023. Profiling of basal and ligand-dependent GPCR activities by means of a polyvalent cell-based high-throughput platform. *Nature Communications*, 14, 3684.
- ZHANG, C., TSOI, R. & YOU, L. 2016a. Addressing biological uncertainties in engineering gene circuits. *Integr Biol (Camb)*, 8, 456-64.
- ZHANG, H., LIN, A., GONG, P., CHEN, Y., YE, R. D., QIAN, F., ZHANG, Y. & YU, Y. 2020. The Chemokine-like Receptor 1 Deficiency Improves Cognitive Deficits of AD Mice and Attenuates Tau Hyperphosphorylation via Regulating Tau Seeding. *J Neurosci*, 40, 6991-7007.
- ZHANG, L., WANG, H., LIU, C., WU, Q., SU, P., WU, D., GUO, J., ZHOU, W., XU, Y., SHI, L. & ZHOU, J. 2018. MSX2 Initiates and Accelerates Mesenchymal Stem/Stromal Cell Specification of hPSCs by Regulating TWIST1 and PRAME. *Stem Cell Reports*, 11, 497-513.
- ZHANG, W., XUE, D., YIN, H., WANG, S., LI, C., CHEN, E., HU, D., TAO, Y., YU, J., ZHENG, Q., GAO, X. & PAN, Z. 2016b. Overexpression of HSPA1A enhances the osteogenic differentiation of bone marrow mesenchymal stem cells via activation of the Wnt/ β -catenin signaling pathway. *Scientific Reports*, 6, 27622.
- ZHANG, X., LI, Y., XU, H. & ZHANG, Y. W. 2014. The γ -secretase complex: from structure to function. *Front Cell Neurosci*, 8, 427.
- ZHANG, Y.-T., HE, K.-J., ZHANG, J.-B., MA, Q.-H., WANG, F. & LIU, C.-F. 2021. Advances in intranasal application of stem cells in the treatment of central nervous system diseases. *Stem Cell Research & Therapy*, 12, 210.
- ZHANG, Y., PAK, C., HAN, Y., AHLENIUS, H., ZHANG, Z., CHANDA, S., MARRO, S., PATZKE, C., ACUNA, C., COVY, J., XU, W., YANG, N., DANKO, T., CHEN, L., WERNIG, M. & SÜDHOF, THOMAS C. 2013. Rapid Single-Step Induction of Functional Neurons from Human Pluripotent Stem Cells. *Neuron*, 78, 785-798.
- ZHANG, Y., TIAN, Q., ZHANG, Q., ZHOU, X., LIU, S. & WANG, J.-Z. 2009. Hyperphosphorylation of microtubule-associated tau protein plays dual role in neurodegeneration and neuroprotection. *Pathophysiology*, 16, 311-316.
- ZHAO, J., DENG, Y., JIANG, Z. & QING, H. 2016. G Protein-Coupled Receptors (GPCRs) in Alzheimer's Disease: A Focus on BACE1 Related GPCRs. *Frontiers in aging neuroscience*, 8, 58-58.
- ZHONG, M. Z., PENG, T., DUARTE, M. L., WANG, M. & CAI, D. 2024. Updates on mouse models of Alzheimer's disease. *Molecular Neurodegeneration*, 19, 23.
- ZHOU, X., VINK, M., KLAVER, B., BERKHOUT, B. & DAS, A. T. 2006. Optimization of the Tet-On system for regulated gene expression through viral evolution. *Gene Therapy*, 13, 1382-1390.
- ZHUANG, F., NGUYEN, M. P., SHULER, C. & LIU, Y.-H. 2009. Analysis of Msx1 and Msx2 transactivation function in the context of the heat shock 70 (Hspa1b) gene promoter. *Biochemical and Biophysical Research Communications*, 381, 241-246.