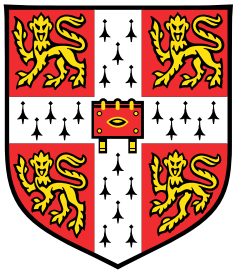
The development of novel methods for the targeting and manipulation of neural circuits *in vivo*



Hassal Lee

MRC - Laboratory of Molecular Biology
University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

Gonville and Caius College August 2019

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Hassal Lee

The development of novel methods for the targeting and manipulation of neural circuits *in vivo*

Hassal Lee

Abstract

Neural networks are at the core of the brain's ability to compute complex responses to our external environment. Clinically, network dysfunction is emerging as a key component of several psychiatric and neurodegenerative disorders such as Alzheimer's disease or schizophrenia. However, our ability to precisely and safely manipulate neural networks for research and deliver network-specific therapy remains limited. To address this problem, our lab recently developed a monosynaptically restricted Self-Inactivating Rabies virus (SiR) which enables the targeting of neural circuits without cytotoxicity. To expand the scope of SiR we further developed the technology in two directions: A) By incorporating the CRISPR/CAS9 gene-editing machinery into the SiR genome to successfully edit endogenous loci *in vitro* and *in vivo*. B) By designing an improved second generation SiR virus (SiR 2.0) which applies the same SiR technology to a challenge rabies strain (CVS-N2C). SiR 2.0 demonstrates increased neurotropism, increased trans-synaptic transfer efficiency and markedly decreased immunogenicity compared to the SiR 1.0 vector.

These advancements expand the scope of SiR viruses to be used in the genome-editing of circuits *in vivo*. A combined SiR 2.0 CAS9 virus, in physiology, allows us to investigate the roles of genes within circuits in the brain function of live animals. For therapy, it paves the way for the rabies virus' potential use to edit disease-related genes in dysfunctional circuits. Despite the circuit-basis of many neurological disorders, existing gene therapy vectors are not circuit specific. In addition, the practical difficulties of delivering therapeutic agents at high doses into the central nervous system exacerbates our inability to achieve high therapeutic loads into affected circuits. In contrast, a SiR 2.0 CAS9 virus would, following injection into peripheral organs, trans-synaptically spread into desired circuits of the central nervous system that are affected in neurological disease (e.g. networks demonstrating pathological protein propagation in neurodegenerative disorders) and edit disease-related genes.

Lastly, our interest in network-level pathological protein propagation also led us to investigate the biology behind this observation. Due to additional evidence that a significant number of other proteins in physiology also show interneuronal movement, we hypothesised that perhaps this is an overlooked phenomena in neurobiology which could have key implications on our knowledge of interneuronal communication and synaptic mechanisms. In the final chapter, I discuss the design and development of novel molecular tools and experimental systems that would enable a deeper study into the mechanisms behind interneuronal protein transfer.

I would like to dedicate this thesis to my loving parents, without whom none of this work would even have begun to be possible . . .

Contributions by Collaborators

The work in this thesis would not have been possible without the help and co-operation of collaborators. A lot of my *in vivo* work was dependent on the efforts of Dr. Ernesto Ciabatti and Dr. Fabio Morgese who were involved in the production of the viral reagents required for these experiments.

Details of the contributions to the work presented in this thesis is as follows:

- Chapter 3:
  - The viruses used in this chapter were produced by Dr. Ernesto Ciabatti (EC) and Dr. Fabio Morgese (FM) as follows:
    - SiR N2C mCHERRY PEST CRE - EC
    - SiR B19 mCHERRY PEST CRE - EC
    - WT B19 nuGFP CRE - EC, FM
    - WT N2C mCHERRY PEST CRE - EC, FM
    - WT B19 GFP - FM
  - The sanger sequencing of individual SiR particles *in vitro* was conducted together with Dr. Ernesto Ciabatti.
- Chapter 4:
  - The viruses SiR-N2C-CAS9-3xFLAG-gRNA-tdTOMATO and SCRAMBLE were produced by Dr. Fabio Morgese.
  - All other viruses in this chapter were produced together by myself and Dr. Ernesto Ciabatti.
  - The two-colour cell reporter system was produced together by myself and Dr. Ernesto Ciabatti.
  - The testing of gRNAs using the two-colour cell reporter system (experiment in Fig 5.13) was conducted by Dr. Ernesto Ciabatti.

Acknowledgements

First I would like to begin by thanking my mentor and supervisor Dr. **Marco Tripodi**. I am very grateful to him for allowing me to work on these very interesting and pertinent research questions, and always appreciate his innovative ideas and vision which have continuously directed my project.

Secondly, I would like to thank Dr. **Ernesto Ciabatti**, with whom I have collaborated for the projects presented in chapters 3 and 4 of this thesis. He poured a lot of hard work into producing the cell lines required to produce these viruses to a high quality. A big thank you also goes to Dr. **Fabio Morgese** who produced some of the viral preparations that I used for this *in vivo* work.

All the other Tripodi Lab members: **Laura, Letizia and Ana** have been a big support to me, and I would like to thank them for bearing with me during my time in the lab with all its ups and downs. I would also like to thank the MRC Laboratory of Molecular Biology. The efficiency and cutting-edge nature of this research institution has always been an inspiration. I would also like to say a big thank you to the head of the neurobiology division, Dr. **Michael Hastings** for his continued care and support throughout my time at the LMB. A special thanks also goes to **Claire Knox, Chloe Watson** and everyone else in the animal facility as well as my fellow PhD student and roommate **Nesem Ozbey**, who's sensitive camaraderie has offered me much support during the final years of my PhD.

Another institution I would like to thank is my college, Gonville and Caius College. As a MB PhD student it is now my 8th year with the college and I will be forever grateful for the inspirational supervisors who taught me during my undergraduate years, especially Dr. **Helen Mott**, Dr. **Tobias Janowitz** and and Prof. **Roger Carpenter** (who's book Neurophysiology will remain forever one of my favourite books). They truly opened my eyes to the biological sciences and without them I would never have had the drive to pursue a PhD in the first place.

And last but not least, I am hugely grateful to both the Rosetrees Trust and the Frank Edward Elmore fund for their generous funding during the period of my PhD.

Table of contents

List of figures

List of tables

List of abbreviations

Roman Symbols

ΔG-Rabies G-deleted Rabies virus

2A 2A peptides, a class of peptides derived from the 2A region of viral genomes that can induce the cleaving of the recombinant protein in the cell

AAV Adeno-associated virus

ANN Artificial Neural Network

BHK Baby hamster kidney cell

CA1 *Cornu Ammonis 1*, a region of the hippocampus

CA2 *Cornu Ammonis 2*, a region of the hippocampus

CA3 *Cornu Ammonis 3*, a region of the hippocampus

CRISPR Clustered regularly interspaced short palindromic repeats

CTB Cholera toxin subunit B

DSB double strand break

EC Enthorhinal cortex

EnvA Envelope protein of the A avian sarcoma and leukosis virus

FACS Fluorescent activated cell sorting

FBS Fetal bovine serum

FLEX CRE-dependent genetic switch

FLP Flippase, a recombinase which targets FRT recognition sites

FRT FLP recognition target

GFP Green fluorescent protein

GO Gene ontology

HEK293T Human embryonic kidney cells 293T

HRP Horseradish peroxidase

Indel Insertions/deletions

IP intraperitoneal

ISGs Interferon-stimulated genes

ITR Inverted terminal repeat

LoxP Locus of X(cross)-over in P1, CRE recognition site

MNP Multi-nucleotide polymorphism

Neuro2A mouse Neuro2A cells

oG Optimised glycoprotein

p.i. Post-infection

PAMPs Pathogen-associated molecular patterns

PBS Phosphate buffered saline

PEST Peptide sequence rich in proline (P), glutamic acid (E), serine (S), and threonine (T)

PFA Paraformaldehyde

PRRs Pattern recognition receptors

PRV Pseudorabies virus

RABV Rabies virus

RFP Red fluorescent protein

RIPA Radioimmunoprecipitation assay buffer

RPM Revolutions per minute

SC subcutaneous

SiR Self-inactivating Rabies virus

SNP Single nucleotide polymorphism

TEV Tobacco etch virus

TEVp Tobacco etch virus protease

TEVs Tobacco etch virus cleavage site

WGA Wheat germ agglutinin

WT Wild-type

Short flippase recognition target sites nomeqref 2.0

Chapter 1

General Introduction

A fundamental aim of neuroscience is to understand what the brain does, and how it does it. Its functions and workings have mystified humans across all times and cultures. One of the oldest known descriptions of the brain comes from a 5000 year old papyrus (1700 BC) of an ancient Egyptian doctor[1]. Although he attributes injuries to an organ in the skull with symptoms such as an inability to speak, walk, or track objects with one's eyes, the first intellectual understanding that the brain is the center of sensation, cognition, emotion and motor control was only made 3000 years later by the Greek philosopher Alcmaeon[2]. Given the highly varied and complex computations a brain can perform, defining a core function of the brain can be difficult. A simplified yet attractive explanation is that the brain is essentially an organ system that converts sensory input from the environment into motor outputs that determine human behavior. Much of the drive of neuroscience research is to try dissect out how the brain manages to carry out such a challenging task.

Anatomically, the human brain consists of approximately 100 billion neurons and each of these neurons, on average, form 7000 synapses with other neurons. A prevalent logic in the field is that if we can understand how these neurons connect and functionally communicate with one another, we can come a step closer to understanding how the brain can perform complex tasks such as make predictions, recollect memories and react to the physical and social world around us. Experimentally, there are many approaches to studying the brain. For example, one can take an anatomical approach and acquire connectivity diagrams of neural circuits, thus revealing the underlying structure of the brain apparatus. This type of information could help uncover how the brain performs complex computations. Alternatively, one can record from or manipulate neurons or groups thereof to establish correlations or causative relationships between the activity of brain units and perception or behavior. From a more global perspective, one can treat the whole brain as an unknown entity and perform psychological experiments to reveal patterns in behavior which can give rise to broader theories of brain function. All of these approaches have been informative and it appears that increasing cross-talk between these different fields will be crucial to furthering our understanding of brain function.

In addition, it is interesting to consider what the pathologies of the brain reveal about the physical substrate of mind and thought. For example, what changes occur in the brain of a hallucinating patient who can hear real and vivid, yet hallucinatory, voices? Such patients, in so many other aspects of brain function, are highly functional. Therefore, these symptoms suggest an extremely specific computational deficit or hyperactivity. This further asks the intriguing question: exactly what physical alteration in the brain gives rise to such a specific dysfunction? There is now accumulating evidence that such deficits exist at the level of

circuits, with many psychiatric disorders showing correlations with altered circuit architecture and connectivity[3] and neurodegenerative diseases revealing network level progression of pathological protein aggregates[4].

Despite the emerging importance of circuits in brain function, experimentally studying the roles of neural networks in live animals has proven difficult. Nervous tissue is very dense, heterogeneous and intricately connected making it hard to specifically target or even label selected circuits effectively. Continuous modifications of a naturally trans-synaptic rabies virus has been crucial to facilitate anatomical circuit visualization in mammals. In parallel, there have been increasing improvements of calcium and voltage sensors to record electrical activity from neurons[5]. For functional manipulation, chemogenetic and optogenetic techniques have been developed[6]. In addition, our ability to specifically edit the genome has been revolutionized by the development of CRISPR/CAS9 techniques[7]. These advancements in isolation have been highly transformative and helped to overcome the numerous technical obstacles experimenters face when trying to label, record from or manipulate nervous tissue in live animals. However, we now need an integration of such tools - enabling manipulation not of single neurons, but of neural circuits in live animals as they perform complex brain computations. Neuroscience, like much of science, is a technology-dependent field, where new tools enable the revelation of novel biological principles. More sophisticated methods of recording from and manipulating structural circuits *in vivo* are necessary to help fill the challenging gaps that exist between circuit theories of brain function, and the biological validation of those theories.

In this thesis, I argue the importance of neural circuits in brain function and focus on the development of novel tools to target and probe circuit biology. In my first two experimental chapters I discuss the development of a second generation self-inactivating rabies virus that can be used to perform circuit specific genome editing *in vivo*. Our motivation for this tool development is twofold. First to enable, for the first time, a fusion of the experimental methods available to specifically target circuits, precisely edit the genome of cells, and observe behavioral paradigms. Second, to initiate the development of circuit specific gene therapy vectors for the treatment of neurological and psychiatric disorders. In the final experimental chapter, I work to engineer novel protein tools that enable exploration into the mechanisms behind interneuronal protein transfer, a phenomenon observed both in physiology and neurodegenerative disease. We hope that such tools can enable further understanding of this process, which may have broad and deep implications for the underlying biology of circuit communication and brain function.

1.1 Neural network theory of brain function

For the past century, neuroscience has been dominated by the "neuron doctrine" which theorises that neurons are the fundamental structural and functional unit of the brain. This doctrine was initially heralded by Cajal[8] and Sherrington[9] and this underlying dogma has laid the foundation of modern neuroscience. This idea has prevailed predominantly due to the single cell techniques used to make landmark discoveries. A notable example is the Golgi stain that enabled for the first time the clear visualisation of all the dendritic arborisations and axon processes of a small percentage (around 1%) of neurons in the brain[10]. The anatomical revelations from this novel staining method led Ramon y Cajal to propose that the nervous system was composed of individual cells rather than a reticular network and enabled experimenters to hypothesize about neuronal function (e.g. spatial summation, connective complexity and recurrence). Other notable examples include the development of single-cell recording techniques such as the coupling of a triode thermionic valve amplifier to a capillary electrometer[11] which enabled the sensitive quantitative measurement of electrical activity from single sensory nerve fibres. This technique revealed some key properties of cell-level sensory transduction such as frequency encoding and specificity to particular types of stimuli[11,12]. The introduction of the tungsten microelectrode[13] was then used for single neuron recordings in the visual cortex of unrestrained cats[14]. This enabled the mapping of increasingly complex receptive fields in the visual system and revealed key information about sensory information processing. Such single cell recording techniques have revealed that the electrical activity of sensory neurons correlates to and encodes specific information about sensory stimuli, giving rise to the notion of "receptive fields". Although originally derived from more peripheral sensory fibres, this concept that single cell activity reflects neuronal function and computation has been extended over to higher order neurons which respond selectively to more and more abstract sensory stimuli[15].

Although these studies at the neuron level have been extremely informative, the neuron doctrine is limited in the sense that we are unable to explain all the complex aspects of brain function by only studying neurons as individual entities. The more recently investigated “network theory” leads us to move away from the neuron level analysis and to look at network level functions. At the heart of our motivation to study the nervous system at the level of networks rather than neurons is the principle that the collective concerted activity of individual units can create new emergent properties within the network that cannot be explained by its individual components. In the first subsection, I discuss how developments in the field of artificial intelligence have shown that networks can generate novel emergent

computational properties and perform complex perceptual tasks such as image or voice recognition with unprecedented accuracy. In the second subsection I discuss the biological experimental evidence for the role of network or population level electrical activity in normal brain computations. In the third subsection, I present the final large body of evidence for network theory, which is that network-level dysfunctions can drive specific pathological human behaviours in psychiatric and neurodegenerative disease.

1.1.1 Artificial neural networks as models of the biological brain

Over 2000 years ago, before much was known about the underlying anatomy or hardware of the brain, or even which body organ was the seat of human intelligence, philosophical thinkers such as Aristotle and Plato proposed an associationism theory to explain how humans were able to perform such complex cognitive functions[16]. Associationism was the idea that our ability to learn, perceive, solve problems, imagine and create was explainable by one essential mental process: the ability to form associations between temporally linked phenomena. This theory was expanded and modified through time with a continuous flow of advocates from the Associationist School such as John Locke, David Hume, David Hartley, James Mill, Alexander Bain and Ivan Pavlov. In the late 1800s, the increasing anatomical knowledge that the brain, the organ and hardware of human intelligence, consisted of a mass of densely interconnected cells gave rise to connectionism: an updated associationist theory where the associations enabling all complex brain functions are stored in the connections of simpler units that form interconnected networks[17]. Artificial neural networks(ANNs) are the computational realisation of connectionism.

In recent years, interest in ANNs has increased with the aim of creating connectionist models to solve problems that have proved difficult using conventional computing systems but seemed to be learned with relative ease in biological organisms (e.g. facial recognition). These artificial networks often attempt to mimic biological neural systems and consist of a dense interconnection of simple processing units that mimic neurons. In an ANN, neuronal cell bodies and dendrites are represented as nodes, and the edges connecting the nodes represent axonic projections. The weight of the edges represent the strength of the synaptic connection and the network learns by changing the weights of its connections. The state of each node is characterised by a neuron activation function that can mimic the functional behaviour of neurons. Although there is a huge variety of ANN models, they can in summary be described by:

- the network topology (e.g. number of neurons, number of layers, connections between the layers etc.)
- the neuron activation function used to convert inputs into outputs (e.g. linear vs. non-linear, saturating vs. non-saturating)
- and the learning rules.

The earliest computational model of a biological neuron was proposed by McCulloch and Pitts in 1943[18]. The McCulloch-Pitts (M-P) neuron has a set of binary inputs $I_1, I_2, I_3, ..., I_n$ and one output y. Each input has weight values $W_1, W_2, W_3, ..., W_n$ normalised in the range of either $(0,1)$ or $(-1,1)$. The output y is a linear step function of the weighted sum of the inputs (*Sum*).

$$Sum = \sum_{i=1}^{N} I_i W_i$$

$$y = f(Sum) = 1 \text{ if } Sum \geq T$$

$$y = f(Sum) = 0 \text{ if } Sum < T$$

In essence, the M-P neuron is a linear threshold gate which sums a set of inputs and classifies them into two different classes (an ON or OFF state) by setting a threshold T such that the node fires only if the inputs exceed a threshold value. This design was seen to mimic biological properties such as the summation of pre-synaptic inputs and the all or nothing nature of action potential propagation initiation. Although simple, the M-P neuron can compute many boolean functions (e.g. AND, OR, NOT, NOR etc.) and was shown to have significant computing potential when incorporated into networks. This was a crucial first step for the field of artificial intelligence.

However, a major limitation of the M-P neuron was that its weight and threshold values were fixed. The subsequent development of the M-P neuron, Rosenblatt's perceptron proposed in 1958[19], introduced variable weight values between units and therefore enabled networks of these units to be trained. The convergent learning algorithm incorporated into this computation model again mimicked the biologically driven Hebbian learning rule that "neurons that fire together, wire together."[20]. This development was crucial for artificial networks which needed to be trained to function correctly over time as opposed to conventional computing systems which have fixed instructions to perform specific computations.

Although at the time the computing potentials of the perceptron were criticised, it was recognised that multi-layered perceptrons (rather than individual perceptrons) could have

considerable computing capacity and power[21]. The introduction of backpropagation learning rules facilitated the training of these multiple-layered networks and has yielded increasingly powerful results[22]. In addition, over time there have also been many more activation functions designed for artificial neurons which can be binary or continuous in nature. Artificial neurons with continuous characteristics aim to mimic "pattern" or “frequency" coding of biological neurons and their additional characteristics (e.g. the gradient of a sigmoid activation function) can contribute to the learning of neural networks.

Currently, ANNs are being used to perform many complex cognitive tasks with an unprecedented level of efficiency and accuracy. The study of such ANNs is starting to bridge the fields of biology and psychology, offering insight into the link between the brain and the mind. Whilst neuroscience offers a wealth of knowledge about the biology and molecular chemistry of neurons, it has not yet revealed general theories of brain function. In contrast, psychology and philosophy have generated many theories of brain function but they are yet unable to be tested experimentally. Although ANNs do not embody all the known properties of neurons in the human brain, neural network theory aims not to mimic all the intimate details of a neuron, but rather to simplify neurons as input-output devices which can be connected into various different network arrangements, some of which can perform brain-like tasks. They are essentially network models that mimic important biological connectionist principles and can thus act as a "potential guide to the mind"[23]. For example, ANNs demonstrate the highly complex emergent properties of connectionist systems (e.g. visual recognition and speech recognition) and also reveal how networks can perform such complex tasks, something which is yet difficult to study purely biologically. In other words, ANNs help us generate computational theories of brain function from our present knowledge of biological network architectures, neuronal connectivity and activity tuning through a process of reverse engineering.

1.1.1.1 Artificial neural networks can store distributed memories

Biologically inspired artificial neural network models have clearly demonstrated that networks can generate emergent computational functions that are not present in single units. A seminal example is the Hopfield network which clearly demonstrates that networks can store distributed memories at a population level that can be retrieved in a content-addressable manner[24]. A Hopfield network consists of a dynamic network of N neurons $(V_1, ..., V_i, ..., V_N)$ where:

- each neuron can have one of two states (-1, 1)

- each neuron is symmetrically connected to every other, so $w_{ij} = w_{ji}$ (where the weight from the j'th to the i'th neuron is given by w_{ij})
- and the state of a neuron is updated in each time step by the dynamics $V_i(t+1) = sign(\sum_{j=1}^{N} w_{ij} V_j(t))$

In this model, as the network is updated the total energy of the system decreases with each state flip of a single neuron. As a result, the network can store and retrieve patterns embedded within the entire network as energy minima or "attractors". Although the design and results of this model are not fully biologically accurate (e.g. in large networks of size N, the number of random patterns that can be stored is approximately 0.14 N and around 1/2 of all neurons are active for each stored state), it is a powerful recurrent neural network realisation of associative memory and a clear demonstration that memories can be distributed as patterns of activation states in a population of neurons rather than the code of a single neuron per se. The drosophila mushroom body, which is the learning and memory centre of the insect brain, also shows a highly recurrent network architecture[25,26] which has inspired the design of RNNs for memory generation[27].

1.1.1.2 Deep neural networks can perform increasingly complex cognitive tasks

Rapid developments in the field of artificial intelligence has led to the development of deep neural networks that can learn increasingly complex cognitive tasks[28]. Convolutional neural networks (CNNs) have produced breakthroughs in the processing of visual imagery, whilst recurrent neural networks (RNNs) have shown high performance for tasks involving sequential inputs such as machine language translation[29–31]. These ANNs support biological network theory in two ways: first by demonstrating the impressive emergent computational power of connectionist systems, and second by giving mechanistic insight into how the same cognitive processes are computed by biological neural networks.

In the field of visual recognition, deep neural networks (DNNs) have been increasingly used to classify images with unprecedented accuracy. Notable examples include the best-performing model of the ImageNet Large Scale Visual Recognition Challenge in 2012, which was a deep convolutional network consisting of 650,000 neurons and 8 deep layers (5 convolutional and 3 fully-connected) and achieved a top-5 test error rate of 16%[32]. This was a breakthrough for computer vision at the time, and triggered the increased use of convolutional networks for recognition and detection tasks. Subsequently developed models have improved error rates which have even superseded human performance in image classification, with a 5% top-5 test error rate[33]. This high performance level has led to investigations into whether

the algorithmic representations of DNNs resemble human vision. In order to learn more about exactly how such trained DNNs work, visualisation techniques (both computational and empirical) have been developed to reveal what single neurons at any layer of the network represent[34–36]. Such studies show that DNN architectures learn hierarchical representations which are similar to the hierarchy of low- to high- level visual areas in the human visual pathway[37,38]. Neurons in the DNN represent increasingly complex features at higher levels, with increasingly more intuitively desirable properties such as invariance, class discrimination and compositionality. In addition, a comparison of the neural population representation of the macaque inferotemporal (IT) cortex for object recognition with that of convolutional DNNs has shown that both IT cortex and DNNs create similar representational spaces and that the top-level features of DNN models can predict IT responses, supporting the idea that high-performance DNNs may provide insight into primate visual processing[39].

DNNs have also shown impressive results in the field of language processing and translation. For example, NETtalk, a three layered network of around 300 neurons can learn how to convert english text to speech over the period of half a day[40]. The network model, although very simple, shows many similarities with human learning and memory. For example, spaced learning is more effective for long-term retention than massed practice. The model also learns different phonological rules at different speeds that correlates with sounds that children find difficult to learn to read, such as the soft "c" sound. In addition, the ANN showed a robustness to damage that is typical of biological networks. Its performance degrades gradually with increasing damage to connections within the network, demonstrating that the rules of text pronunciation are learned and distributed into the network, rather than stored in particular units or links. As such, the model illustrates many aspects of human learning and can be used to develop mechanistic theories of complex learning phenomena. The *Drosophila* mushroom body, which is the learning and memory centre of the insect brain, also shows a highly recurrent network architecture[25,26] which has inspired the design of RNNs for memory generation[27].

1.1.2 Neural networks in physiology

In this section I present the biological experimental evidence for neural network theory of brain function. Although the content is by no means comprehensive, here I aim to highlight some pertinent examples of biological evidence supporting the importance of network level activity, not individual neurons, in complex brain function. These examples include our knowledge of neuroanatomy and network structure, the role of population coding in sensory

perception, the ability of neuronal population dynamics to predict behaviours in live animals, and the multifunctionality of single brain units or networks showing the significance of fluid network activity for brain function.

Just as single cell techniques such as the Golgi stain enabled the visualisation of single neurons in the brain and thus paved the beginning of the "neuron doctrine", our understanding of network function has been enabled by novel techniques that allow us to study the brain at a network-wide level. Much of the following evidence for network function in biology has been enabled by the development of novel methods to trans-synaptically trace networks, record or visualise the activity of many individual neurons together at high spatio-temporal resolution in experimental animals (e.g. *in vivo* 2 photon imaging), to manipulate neuronal populations and the increasingly sophisticated methods of functional and anatomical 3D imaging in humans.

1.1.2.1 Structural evidence for distributed networks in brain function

Biological neural circuits, especially in mammals, show dense connectivity[41]. For example, in the mouse cortex, an average pyramidal neuron receives inputs and projects to approximately an order of 10^4 other neurons[42,43]. In the mouse cerebellum, each Purkinje cell receives inputs to the order of 10^6[44]. In addition to dense connectivity, dual whole cell recordings of local inter-pyramidal connections in the cortex show that the majority of these excitatory connections are small in amplitude (0.2–1.0 mV)[45,46]. This pattern of extremely dense but weak excitatory connectivity suggests that each neuron integrates a high number of sub-threshold inputs and projects to many others. In addition, most cortical inhibitory GABAergic neurons (except VIP-expressing interneurons) show an even denser local connectivity that approaches 100% of local targets and appear to connect with as many excitatory neighbouring neurons as possible[47,48]. This dense and unspecific innervating pattern has been described as a "blanket of inhibition" and is hypothesised to play a role in network-level regulation, such as in the oscillatory activity and synchronisation of cortical networks[49]. Such an architecture of extremely dense, sub-threshold and broad connectivity suggests that each individual neuron is highly embedded into a larger network and thus singly contributes to but cannot explain overall circuit function. This sort of architecture consisting of many interconnected individual elements is characteristic of physical systems that generate emergent properties. Such emergent properties are akin to moving images on a TV screen, and cannot be understood by studying each individual element or pixel in isolation[50].

Graph theoretical analysis of connectomic data has revealed that biological neural networks, both human and non-human, have a "small-world architecture"[51]. Small-world

networks are named after the social "small-world phenomenon" of meeting a stranger far away from home, only to then discover that you share a mutual acquaintance[52] and are used to describe networks which show a combination of both clustered and distributed topology. Clustered topology is a characteristic of ordered networks and are measured by parameters such as the clustering coefficient which quantifies the fraction of connections that exist between the nearest neighbours of a node out of the maximum number of possible connections. High clustering is associated with high local information transfer and robustness. On the other hand, distributed topology is characteristic of random networks and can be quantified by measures such as the global efficiency (the inverse of the mean path length between all pairs of nodes of a network). Distributed random networks show a high global efficiency of parallel information transfer. The decreased energetic economy of biological small-world architecture suggests that there is a functional evolutionary purpose to such network organisation[53]. For example, small-world architecture results in a brain consisting of segregated clustered hubs that can perform specialised processes (e.g. motor speech generation or auditory language comprehension) that are then integrated through high efficiency path architectures to perform executive functions that require efficient communication between these hubs (e.g. the ability to understand and generate words to have a normal conversation)[54]. In addition, alterations in the balances between "randomness" and "orderedness" of networks are found to correlate to aberrant cognition in pathology as will be discussed further in chapter 1.2.

1.1.2.2 Population coding in the inferotemporal lobe for visual perception

As highly visual animals, visual processing and recognition forms a large proportion of human sensory cognition and brain function. The inferotemporal (IT) cortex of monkeys and humans is crucial for visual object recognition. The IT cortex is the last purely visual cortical processing area in the ventral visual pathway which begins from V1 and projects through V2, V4 and area TEO of the temporal cortex. Experimental lesions of this area in monkeys results in an inability to learn how to recognise new visual stimuli as well as an inability to remember how to recognise visual stimuli they have seen before[55]. In humans, damage to the IT cortex can also result in general visual agnosia. More rarely, IT damage can also lead to specific agnosias resulting in an inability to recognise particular visual features or concepts such as colours (achromatopsia) or faces (prosopagnosia). In addition to lesion studies, functional imaging with fMRI in humans shows that visual objects activate the IT cortex and that specific regions of the IT cortex are selective to specific classes of visual stimuli, for example the fusiform gyrus for faces and body parts, and the dorsal and anterior regions for recognising eye position.

Single cell recordings of neurons in the IT cortex have enabled interesting insights into how visual pattern recognition is computed. In most of these studies, one finds that the majority of neurons in the IT are not narrowly tuned for a specific visual stimulus or concept but rather respond broadly to most experimentally presented visual stimuli. However, for a lesser proportion of cells, one finds that neurons can show preferential tuning to a particular visual stimulus. In an exemplary study in the macaque IT cortex, it was found that 27% of IT neurons were unresponsive to all presented stimuli, 22% responded equally to all presented stimuli and most of the remaining stimulus-selective cells showed at least a small response to almost every stimulus tested[56]. As a general rule, as you approach the anterior tip of the IT cortex, neurons show preferential responses to increasingly complex visual stimuli, the most complex of which have been identified are faces. For example, studies in the IT cortex of macaques have shown cells firing preferentially to hands[57] or faces[56,58] in the superior temporal sulcus. In humans, single cell recording experiments in the mediotemporal lobe (MTL), a common foci for eplieptic patients, has also revealed that cells in this region can show preferential tuning to very complex visual stimuli and concepts such as the Sydney Opera House or to famous actresses (e.g. the "Jennifer Aniston cell")[59]. These cells fired preferentially not only to photos of an actress' face, but also from the visualisation of the letters of her name.

The identification of such specific selectivity to complex trigger features has led to many theories to explain how the IT cortex computes visual pattern processing. On one extreme, such studies have often been interpreted to support the hypothesis that complex perceptual concepts can be represented by single cells. These perception embodying single neurons have been proposed independently by many and given numerous names: pontifical cells by James[60], gnostic cells by Konorski[61] and grandmother cells or cardinal cells by Barlow[62]. The idea is that there are single cells, such as a "grandmother cell", that only fire selectively to highly specific visual percepts such as one's "grandmother". As the perception of "grandmother" is represented singly by that cell, if it is lost the conceptual perception of "grandmother" is also lost. However, these single cell recording studies cannot truly support a single cell code for visual percepts as they cannot prove that the identified neuron is solely responsible for the recognition of a particular face or concept. In addition, if there truly was a single cell code to represent such complex concepts, statistically the chances of recording from a single neuron and matching it to the specific concept it encodes would be extremely low. In addition, these neurons show preferential tuning, not complete selectivity, to a complex percept. For example in the Quiroga et al. study where a large collection of

pictures of famous individuals and landmarks were used, each unit showed strong responses to about 3% of the stimuli[59].

More convincingly, these single cell recordings support a population code of visual percepts. Most IT neurons are not narrowly tuned for a specific stimulus, and the few that do show preferential tuning still show responses to other non-specific visual stimuli. These findings suggest that visual images will probably activate many cells in the IT cortex and as such, visual concepts in the IT cortex will be represented in the activity of a population of neurons. In addition, further investigation into the properties of "face cells" in macaque monkeys has shown that, rather than encoding a specific person, face neurons show "axis coding"[63]. When images of faces are formatted into a high-dimensional linear space, face cells are found to fire proportionally to the projection of a visualised face onto linear axes within this space which represent specific facial features. This results in a face cell **ensemble** of around 200 cells that can identify any face within the space. The finding of face preferential cells in humans supports a sparse population code for abstract representations[64] and informs us that, within human visual recognition networks, there are nodes that are selective to complex trigger features[65]. These individual neurons participate in many different population assemblies to form a sparse population code which is distributed in the activity of many neurons. Such distributed storage codes theoretically allow for greater coding capacity due to combinatorial multiplication, easier retrieval, greater capacity for generalisation and resistance to noise. As elegantly described by Sherrington, "Where it is a question of 'mind' the nervous system does not integrate itself by centralisation upon one pontifical cell. Rather it elaborates a million-fold democracy whose each unit is a cell"[66].

1.1.2.3 Neural population dynamics can code and predict behavioural outcomes

Decision making is another important aspect of human cognition. In essence, all movements that are made are a consequence of decisions computed by our nervous systems, whether consciously or unconsciously. The biological study of how decisions are made across many different species for various experimental decision making tasks have revealed that behavioural outcomes and reaction times correlate with population-level spatiotemporal sequences of neuronal activation. Such findings have relied on a) improving techniques for simultaneous recording from neural populations and b) improving analytical techniques to extract neural population states and relate them to behaviour. The ability to correlate and even predict behavioural parameters from neural population feature extraction, but not single neuron activity, shows that the processes of decision making and movement generation are population encoded[67].

In the midbody segmental ganglion of the leech central nervous system, simultaneous recording of 143 neurons during swim vs. crawl decision making followed by principal component analysis (PCA) to identify dimensions that separate the two behaviours showed that neural population dynamics could discriminate behavioural outcome earlier in time than single neurons, highlighting the emergent circuit level coding of decision making[68]. In the more complex mouse posterior parietal cortex (PPC), simultaneous two-photon calcium imaging of layer 2/3 neurons during a virtual maze navigation task revealed that whilst individual neurons could not be used to decode the behaviour, distinct spatiotemporal sequences of activity across the population could discern and predict different behavioural outcomes[69].

The recent development of new methods to analyse neuronal population activity of single trials, as opposed to data averaged across many trials, have enabled a deeper illumination of the relationship between neural population activity and behaviour[70,71]. The need for single-trial analyses of population activity is particularly pressing for tasks that require internal processing (e.g. decision making, perception and motor planning). For example, in the mouse parietal cortex, the use of a new clustering algorithm revealed high trial-to-trial variability in a virtual maze navigation task which could be explained by past events such as previous evidence cues, behavioural choices and reward outcomes[72]. This showed that dynamic population properties could represent short-term memory. In addition, single-trial analysis of population dynamics in the monkey premotor cortex have shown that single-trial deviations of neural state space trajectories from the mean neural trajectory for a behavioural outcome can predict behavioural parameters such as hand velocity and reaction time variability between trials for complex reach tasks in monkeys[70,71,73]. Such studies show the subtleties and hence significance of population dynamics for decision making and motor planning.

1.1.2.4 Single neurons and networks can participate in multiple different computations

The lack of strict selectivity of neurons or neuronal populations to specific percepts (as demonstrated by "face cells") or specific behavioural outcomes (as demonstrated by dynamic decision making networks) highlight another principle of brain function which is that of multifunctionality. Just as the same "face cell" can contribute to the perception of different visual objects[59], the same neurons, networks or brain regions can participate flexibly in many different functional groups to code different sensory or behavioural outcomes. This strongly alters the concept of "receptive fields" and demonstrates that neurons may in fact have a much higher dimension than we are currently aware. There is now rapidly accumulating evidence that the previously determined low dimension of "face cells", "command neurons",

"swim interneurons" or "primary sensory cortices" may in fact be a consequence of the simple sensory or behavioural paradigms used in experiments.

In the study of decision making, the discovery of single neurons that are able to trigger complex behaviours on experimental stimulation have historically been defined as being "dedicated" to one specific behaviour. This has led to the hypothesis that there are distinct "dedicated-command" or "decision-making" neurons for each behaviour[74]. However, with time it has been shown that the activity of decision neurons are not behaviour specific and that the stimulation of the same neuron can trigger more than one behaviour. For example, interneurons in the leech nervous system that have previously been labelled as "swim gating" or "swim trigger" interneurons were also excited during whole-body shortening. This finding was generalisable to most neurons in the leech, including high-level decision neurons. In addition, it was also observed that the electrical stimulation of a single neuron (R3b1) in an isolated nerve cord could elicit both swimming and crawling motor patterns showing that decision neurons are not only active during more then one behaviour, they can singly initiate two different behaviours[75]. This multi-functionality of decision neurons suggests that decisions are not conducted by a single cell identity but by a network-level combinatorial code, where the behavioural outcome depends on the combination of decision neurons that are active at a given time.

In the motor system, the same neuron network can elicit multiple different motor behaviours. For example, swim central pattern generator (swim-CPG) interneurons in the leech are active in an oscillatory mode during swimming, but also tonically active during withdrawal[75,76]. The swim-CPG network of the marine mollusk has also been suggested to be multifunctional for both swimming and withdrawal behaviour[77]. More surprisingly, the same swim-CPG network is also able to generate crawling behaviour which requires a starkly different motor output to swimming. Swimming is a rhythmic, muscle-driven brief behaviour whilst crawling is a tonic, cilia-mediated prolonged behaviour which requires connections to different effector neurons[78]. Rodent medullary networks also demonstrate multifunctionality, with the same networks generating different breathing patterns[79] or different responses to food[80]. Thus, across many species, the same motor networks can generate markedly different motor behaviours, highlighting the importance of network dynamics rather than single cell or network identity in behavioural movement generation.

In a more extreme example of multifunctionality, it was recently observed that primary sensory cortical areas in the mouse that are predominantly thought of as visual processing areas are also tuned to mouse facial movements. In this study, two-photon calcium imaging was used to record the activity of $10,000$ neurons in the visual cortex with simultaneous facial

movement monitoring. Analyses revealed that visual cortex activity is highly dimensional, with a third of neural activity being predictable by a multidimensional model of the mouse's facial behaviour. The same neurons were found to represent both sensory stimuli and behavioural variables, showing that neuronal multifunctionality can traverse motor and sensory systems. In addition, subsequent neuropixel probe recording of neurons across 8 other brain regions showed that the encoding of behavioural facial movement was brain-wide[81]. This is supported by anatomical findings of whole-brain connectivity and supports the existence of brain-wide encoding of behavioural variables through widely distributed networks.

1.2 Neural networks in disease

In addition to the computational and physiological evidence presented in the previous sections, another large body of evidence to support the network basis of brain function comes from the study of pathological conditions of the human nervous system.

Psychiatric and neurodegenerative diseases manifest as pathological changes in human cognition and behaviour and present a huge disease burden in modern society. As such there has been intensive research into understanding the aetiology and pathogenesis of these diseases as well as for developing therapeutic agents. However, despite huge efforts, there has been relatively little progress[82]. Even for the cases of hereditary neurodegenerative disorders where causative genes have been identified, we still do not understand the pathophysiology and these diseases are still incurable. This relative lack of progress despite huge inputs has suggested to many the need for new unifying theories or paradigms to approach the study and treatment of these diseases. A paradigm shift that has already begun to take place is that the complex behavioural and cognitive dysfunctions that are observed in these diseases emerge due to altered temporal dynamics and functional connectivity of wide-spread large-scale networks, rather than of micro- or macro- level dysfunctions in circumscribed regions of the brain[53,83].

In fact, this debate on whether higher human cognitive functions are localised to specific brain regions, or distributed widely throughout the brain has been a long topic of debate over the past few centuries. For example, the famous case of Phineas Gage, a railroad worker who survived the impalement of a rod through his left frontal lobe, initiated many discussions in the 19th century regarding the cerebral localisation of brain function. Although the accounts of the symptoms of his almost total left frontal lobe lesion vary significantly, it is most often reported as was described by Harlow, Gage's practising surgeon. Harlow reported that the lesion had almost no effect on Gage's sensory, motor or cognitive function but led to a gross and vulgar personality change[84]. As such, this case is commonly highlighted as a classic example of the localisation of cerebral brain function. In this instance, the left orbitofrontal lobe is described as a specialised region for "personality" or "social behaviour". However, other surgeons at the time such as Bigelow reporting on Gage's condition did not note any significant changes in Gage's mental functions. Rather he stated " . . . *as portions of the brain to which are allotted different functions, passions, etc, were entirely destroyed, phrenology would teach that these functions would be entirely annihilated. This, however, was not the case, for 'the man was still able to walk off, and talked with composure and equanimity of the hole in his head,' and has never been . . . other than a rational man.*[85]". Around

the same time, other debates of a similar vein also include Broca's phrenology-inspired accounts of localised aphasia versus Wernicke's associationist or "disconnection theory" of speech pathologies[54]. As such, the idea that higher cognitive functions such as verbal communication are dependent on network architecture and connectivity has been around since the late 1800s.

Although lesion studies to assess theorised functional localisation have been experimentally commonplace in neuroscience, experimental investigation into connectionist theories of human pathologies have been difficult. Only recently with the advent of improved neuroimaging and electrophysiology techniques to study anatomical and functional connections in living patients, have they been able to be experimentally investigated. Such studies are now resurfacing under a newly defined field of "pathoconnectomics"[83,86]: the idea that altered network architecture and functional connectivity leads to cognitive and behavioural dysfunction. Undoubtedly, the recent technological developments enabling these investigations have contributed to the recent paradigm shift towards connectionist theories regarding network-dysfunction in neurodegenerative and psychiatric disorders.

In this section I first examine the evidence supporting that network-level dysfunctions underlie these pathologies, be it in the context of altered network topology and circuit synchronisation in neuropsychiatric disorders or network-level pathological aggregate protein progression in neurodegenerative disease. I then move on to discuss the therapeutic challenges faced by these pathologies. Especially, I highlight the lack of effective anatomical network targeting in the central nervous system as a key reason why these diseases have yet been so resistant to new therapeutic developments.

1.2.1 Network dysfunction underlies neuropsychiatric and neurodegenerative disorders

1.2.1.1 Psychiatric disorders

A psychiatric disorder is described as a clinically significant behavioural or mental pattern that causes significant distress or impairment of personal functioning. Mental health disorders such as depression, schizophrenia, post-traumatic stress disorder (PTSD), phobias and autism are common in the worldwide population, with one in four people being affected at some point in their lives[87]. Patients suffering from mental disorders exhibit an extremely diverse array of specific cognitive or behavioural symptoms. For example, a patient with agoraphobia feels an overwhelming sense of fear when entering confined public spaces, leading to an inability to use public transport or even leave their own home. On the other hand, an autistic

child will, typically but not always, have difficulties making eye contact and communicating effectively with others but will show obsessive interests in specific objects or hobbies[88]. In recent years, there has been a growing body of evidence to show that these very specific behavioural abnormalities observed in neuropsychiatric disorders have a network basis. They are explainable not by studying individual neurons or circumscribed anatomical regions of the brain, but rather by looking at the dynamics and properties of large-scale distributed networks[89,90].

Developments in magnetic resonance imaging (MRI) have been pivotal in enabling studies into the altered anatomical and functional neural connectivity observed in human psychiatric patients. Most commonly used in these studies are diffusion tensor imaging (DTI), a form of diffusion-weighted MRI that uses the restricted diffusion of water through myelinated nerve fibres to map white matter tracts, and functional MRI (fMRI) which detects the increased oxygenation of active brain areas and so can be used to detect correlated activity of distinct brain regions over time to give a measure of functional connectivity. Such neuroimaging studies in humans typically compare anatomical and functional network connectivity between patients and non-disease control groups. These comparisons have revealed alterations in network connectivity that correlate to disease states, and in often cases, also to symptom severity.

In the case of schizophrenia, over 200 DTI studies have shown that schizophrenic patients have reduced white matter volumes in several brain regions and altered network organisation of cortico-cortical connections[91–93]. Graph theory analyses of both DTI and fMRI derived schizophrenia pathoconnectomic data have shown longer average path lengths and therefore reduced global network efficiency[94,95], supporting the hypothesis that cognitive deficits are a result of altered network architecture and function. Interestingly, different subpopulations of schizophrenic patients showed spatially distinct non-overlapping changes in white matter abnormality. For example, chronic unmedicated patients showed decreased fractional anisotropy in the left inferior longitudinal fasciculus and left inferior frontal-occipital fasciculus that correlated with speed of visual and verbal processing and learning[96]. In contrast, first-episode non-medicated paranoid schizophrenic patients showed fractional anisotropy decreases in tracts only on the right hemisphere[97], highlighting the specificity of network dysfunctions to different clinical behavioural phenotypes.

Neuroimaging studies for other psychiatric disorders have also shown altered connectivity in specific circuits. Patients with autism spectrum disorders display reduced connectivity that is pronounced in circuits specific for mirror neuron systems and social processing, and the degree of altered connectivity correlates to symptom severity[98,99]. Mood disorders such as

major depressive disorder also show decreased integrity of prefrontal intercortical circuits[100] and of the corticospinal tract[101].

In addition to neuroimaging techniques, electrophysiological techniques such as electroencephalography (EEG) and magnetoencephalography (MEG) have also been used to map functional networks in patients. These techniques, when compared to fMRI, have the advantage of directly measuring neuronal activity with a higher temporal resolution (1-100 Hz) which is sufficient to detect both local oscillatory patterns (gamma-band frequencies of 25-200 Hz) and the synchronisation of slower long-range networks (theta, mu, alpha and beta frequencies of 4-25 Hz). Such studies have revealed abnormalities in network oscillation and synchronization in many mental disorders. In schizophrenia, there is a reduction of beta and gamma oscillation and synchronisation during cognitive tasks and during rest[102,103] and an increase in low frequency[104] activity at rest. In autism, mu waves over the motor cortex which are usually suppressed both during the generation and observation of movement show a lack of suppression for observation of movements in autistic patients[105,106]. The degree of mu suppression also correlated to the behavioural assessment of imitation[105], suggesting a dysfunction of the mirror neuron system in this condition and supporting the anatomical findings of neuroimaging studies.

Taken together, such findings implicate altered network architecture and temporal synchronization as a key process of psychiatric pathogenesis and highlight the importance of network activity for normal brain function. They also emphasize that the specific biological defects observed in these complex behavioural conditions can be attributed to circuits, as such informing future novel therapeutic strategies for these debilitating yet currently incurable conditions.

1.2.1.2 Neurodegenerative disorders

Neurodegenerative disorders are characterised by the progressive degeneration of neurons over time that are mirrored by increasing losses of cognitive, behavioural and/or motor function. A large number of post-mortem histopathological studies and neuroimaging studies have revealed that the pathological protein aggregates that characterise these diseases often appear first in an initial seeding site, and then show a stereotyped sequential pattern of spreading through functionally and anatomically interconnected brain regions in human patients. Table 1.1 lists the circuits that are involved in the propagation of protein pathology in Alzheimer's disease (AD), Parkinson's disease (PD), Amyotrophic Lateral Sclerosis (ALS), Huntington's disease (HD) and Frontotemporal lobe dementia (FTLD).

Neurodegenerative Disease	Prion-like Intraneuronal Protein Aggregate	Spatiotemporal Pattern of Spread
Alzheimer's Disease	Tau (neurofibrillary tangles and neuropil threads)	Locus coeruleus →Transentorhinal region (Layer Pre alpha) → Proper entorhinal cortex and CA1 → isocortical association areas[107]
Parkinson's Disease	Alpha Synuclein (Lewy bodies and neurites)	Medulla oblongata/pontine tegmentum and anterior olfactory structures → Substantia nigra other basal mid- and forebrain nuclei → Neocortex[108]
Amyotrophic Lateral Sclerosis	SOD1, TDP-43, FUS	Agranular motor cortex, bulbar and spinal somatomotor neurons → Reticular formation and precerebellar nuclei → Prefrontal neocortex, striatum → Anteromedial temporal lobe including hippocampus[109,110]
Hungtington's Disease	Huntingtin	Cortical aggregates (extensive pyramidal cell loss)[111–113] → Striatum (selective degeneration of medium spiny neurons)[114]
Frontotemporal Lobe Dementia	TDP-43, FUS, Tau	In TDP43 positive behavioural variant frontotemporal dementia: basal forebrain → prefrontal area → motor cortex, bulbarsomatomotor neurons, spinal cord anterior horn→ visual cortex[115]

Table 1.1 The circuit spread of pathological aggregated proteins in human clinical disease.

This pattern of progressive transneuronal degeneration rather than for example, the diffusion of the disease proteins to spatially adjacent cells, or to brain regions connected by blood flow, suggest that the transfer of these pathological proteins occurs through functional synapses. And indeed, increased neuronal synaptic activity is found to drive protein spreading both *in vitro* and *in vivo*[116–118]. In addition, exogenous expression of human mutant P301L Tau in the entorhinal cortex of mice through transgenic means results in the trans-synaptic spread of tau pathology into neurons downstream of EC circuitry including the hippocampus and cingulate cortex[119].

This observation that patients with neurodegenerative diseases produce a stereotyped sequential progression of symptoms that correlates to the stereotyped sequential spread of pathological protein aggregates through large-scale brain networks strongly suggests that increasing network dysfunction is a key component of pathogenesis[120]. This correlation between pathological protein spread through affected circuitry and progressive symptoms can be well demonstrated in ALS.

Patients initially presenting with ALS commonly experience motor symptoms in a specific segment of the neuroaxis. For example, 70% of patients present with symptoms in an asymmetrical limb[121]. ALS affects both upper and lower motor neurons at the same time, causing a mixture of upper motor neuron symptoms (spasticity, poor balance, incoordination) and lower motor neuron symptoms (muscle atrophy and twitching)[121]. With time, patients

experience progressive motor losses into contiguous areas. Typically the disease may begin in one arm, then progress to involve the ipsilateral leg or the contralateral arm, but not the contralateral leg. As motor symptoms progress, patients will also progressively experience cognitive deficits in frontal executive functions and memory[122–125].

This clinical phenotype is mirrored by the histopathological changes observed with disease progression. TDP-43 accumulation first begins in spinal α motorneurons and agranular motor cortex (especially Betz pyramidal neurons projecting through corticospinal and corticobulbar tracts). This reflects the initial presentation of patients with a combination of upper and lower motor neuron symptoms. Interestingly, lateral portions of the spinal cord are affected more than the medial portions, mirroring the observation that distal muscles tend to be affected first clinically in patients. In stage 2 of protein aggregate progression the pathology spreads to the brainstem reticular formation, red nucleus and precerebellar nuclei which receives indirect and direct input from the agranular motor cortex. In stage 3, first the prefrontal and then post central areas of the neocortex become affected as well as the striatum. In stage 4, anteromedial portions of the temporal lobe including the hippocampus are subsequently affected[109,110]. The sequential spreading of protein pathology to prefrontal and hippocampal regions mirrors the eventual development of frontal executive and memory deficits in many patients. Further studies have found that the presence of cognitive decline in patients correlates with the density and extent of pathological aggregates observed in frontotemporal brain regions (especially the cingulate gyrus)[126].

Despite such observed correlations between pathological protein progression and clinical symptoms, it is unclear exactly how these pathological protein aggregates cause network dsyfunction. Although the progressive death of neurons in particular brain regions correlates with specific functional losses, cell death alone cannot explain the complex phenotypes observed in some neurodegenerative disorders. For example, patients with neurodegenerative dementia can have rapid fluctuations of symptoms within a day where they can have "lucid" moments where memories and cognitive functions temporarily return[127]. These rapid fluctuations suggest that phenotypic deficits are not purely determined by the level of neuronal death but by complex changes in network activity that can be influenced by external factors such as the cognitive demands of surrounding environments[128] or altered physiological or internal states. Studies show that pathological proteins can have more subtle effects than cell death such as reducing neuronal excitability, synaptic plasticity or density[129,130]. In some neurodegenerative models, degenerating neurons appear to be a phenotypic manifestation of the "weakest link" within dysfunctional circuits[131] rather than a direct effect of protein accumulation. Taken together, evidence suggests that neurodegenerative pathogenesis involves a

progressive increase in circuit dysfunction which may ultimately lead to cell death and this in turn can explain the long prodromal stages characteristic of neurodegenerative diseases and the rapid fluctuations of cognitive function observed in patients[120,132].

1.2.2 Network-targeted therapy for circuit dysfunctions

There is a great need for developing efficient therapies for neurodegenerative and psychiatric disorders. Even treatments that are partially effective for these diseases can have a significant impact on quality of life of these patients, and the global market for such therapies is extremely large. For example, the leading antidepressants have global sales of 1-3 billion dollars p.a. and even a partially effective Alzheimer's drug is projected to yield 20 billion US dollars per annum. Despite the considerable amount of resources that have gone into developing therapeutic agents for these diseases, there has been relatively little progress in the treatment of these debilitating conditions. This shortfall has been for a variety of reasons including an underlying lack of understanding of the aetiology of most CNS disorders, difficulty in accessing the CNS systemically, and difficulty in producing specific therapeutic agents that can target the CNS selectively enough to evade side effects.

Given the increasing evidence that extensive network dysfunction is a key component of these diseases, we argue that therapeutic agents also need to target the dysfunctional circuits that underlie the symptomology of these conditions. We propose a shift in paradigm for the therapy of neuropsychiatric and neurodegenerative conditions and argue that there is a need for network-specific agents that can anatomically target the circuits implemented in these diseases. Conversely, we also suggest that the lack of efficacy of the current pharmacological and viral therapeutics in development may be due to their lack of circuit-specificity.

In this section I present an example of effective network-based therapy as employed by empirically verified psychotherapeutic techniques in humans and animal models of PTSD. I also highlight the promising potential of viral gene therapy for neurological conditions, especially hereditary ones, and then cover the notable challenges and difficulties faced when trying to develop effective circuit specific vectors for neurological therapy.

1.2.2.1 The network-basis of effective psychotherapy regimens

Psychotherapy is a widely used and accepted method of treatment for many mental disorders[133–135]. The general effect of psychotherapy (estimated for all types of therapy, patient and outcome) is 0.85 standard deviation units compared to untreated controls[133]. When compared to drug therapy, many studies reveal that psychotherapy not only has compa-

rable efficacy to drug treatments but also has longer-lasting enduring effects on patient outcome[133,136]. Despite this clear empirical observation, scientific explanations for the efficacy and endurance of psychotherapy have been lacking. With the development of more sophisticated animal models of psychiatric conditions and the ability to perform live fMRI imaging in human patients, there has been increasing investigation into the underlying neural bases of psychotherapeutic efficacy in order to enhance our understanding of these conditions and improve existing therapies.

fMRI studies in human patients examining the underlying neural basis of psychotherapy in anxiety and mood disorders show that cognitive therapies for anxiety and mood disorders alter cognitive-affective processing networks. For anxiety and related disorders, fMRI studies have revealed that cognitive behavioural therapy increases prefrontal control of subcortical structures such as the amygdala that have shown to be hyperactive in these fear conditions[137]. Similarly, fMRI studies in depressive patients receiving cognitive therapy (CT) have shown subsequently increased functional use of the dorsolateral prefrontal cortex (DLPFC) in cognitive tasks, and decreased amygdala activity in response to emotionally related stimuli[136]. These fMRI findings suggest that cognitive therapies help patients to recruit prefrontal regulatory brain mechanisms and support the underlying aims of CT which is to replace maladaptive thought patterns and automatic emotional reactivity with altered or more controlled cognitive processing. They also highlight the potential advantages of modulating cognitive emotion regulation networks in the treatment of psychiatric conditions.

Animal models of psychiatric conditions have also helped to dissect the underlying neural bases of psychotherapy. For example, Eye Movement Desensitization and Reprocessing (EMDR) psychotherapy is one of the most recommended therapies for post-traumatic stress disorder[138–141]. Post-traumatic stress disorder (PTSD) is a debilitating anxiety disorder that is characterised by the involuntary and vivid re-experience of previous traumatic events. Although many effective therapies for PTSD exist for short-term treatment, PTSD patients often experience relapse[142]. In EMDR, patients are asked to recall traumatic images whilst simultaneously orienting to bilateral sensory input (e.g eye movements or finger tapping). Although a clear mechanistic explanation for EDMR efficacy is yet lacking, it is hypothesised to work by aiding the long-term reprocessing of traumatic memories by sensory stimulation. The long-lasting effect of EDMR in the prevention of fear recovery has been shown in translational mouse PTSD models, enabling further experimental investigation[143]. Fear reduction in mice following alternating bilateral sensory stimulation (ABS) is associated with enhanced superior colliculus-medial dorsal nucleus excitatory transmission, and increased inhibition of the basolateral complex of the amygdala (BLA). In addition, optogenetic

silencing showed that both the SC-MD and MD-BLA pathway were necessary for ABS induced fear-attenuation. In contrast, photostimulation of the SC-MD pathway was sufficient to prevent fear reoccurrence[144]. This study clearly demonstrates the circuit-level effects of EMDR therapy and supports the hypothesis that psychiatric conditions are essentially network-disorders which can be alleviated or cured through targeted modification of these disease circuits.

1.2.2.2 Viral gene therapy for neurodegenerative conditions

A big challenge for effective developments of CNS therapies is the underlying lack of molecular understanding of CNS disease processes, posing a problem for identifying specific molecular targets for therapy. However, in the less common cases of monogenetic hereditary neurodegenerative disorders, the cause of the disease has been clearly identified, offering a clearer target for therapy. For these conditions, therapeutic strategies could include gene replacement for loss-of-function mutations, knock-out of gain-of-function toxic genes or correction of pathogenic mutations by targeted base pair substitutions using more sophisticated CRISPR/CAS9 genome editing techniques. However, even in such cases where the aims of gene therapy are relatively clear, therapeutically efficient execution of these aims has proved difficult.

The two major viral vectors that have been investigated for CNS gene therapy are lentiviruses and recombinant adeno-associated viruses (rAAVs) as they have low immunogenicity, low toxicity and enable stable long-term gene expression in infected cells. They have been used in many clinical trials with varying levels of success. Problems have included host immunity to wild type viruses from which therapeutic vectors are engineered, host immunity to transgenes[145] and difficulties in appropriate CNS targeting[146] and transgene expression levels[147].

In terms of anatomical targeting, many different possible viral delivery methods have been attempted to access affected CNS regions in disease. Local viral targeting in the CNS can be achieved surgically by direct intracerebral stereotaxic injections into affected brain regions. However this method is invasive and most often shows limited efficacy due to its limited spatial coverage. One notably successful use of this method was the direct stereotactic injection of a rAAV type 2 vector expressing the human aromatic L-amino acid decarboxylase (AADC) gene into the putamen of children with inherited AADC deficiency. Here, the putamen was targeted as it is the major site of AADC activity in the brain and this method successfully stimulated motor development in treated children[148]. This is a proof of principle that *in vivo* gene therapy in the brain is feasible and can have notable therapeutic

effects in humans when targeted into a therapeutically relevant brain region in an efficient way. However, the authors of this paper also note the limitations of this topical approach which does not enable broad coverage of the entire putamen or of CNS regions outside of the putamen that also require dopamine and cause non-motor symptoms of ADCC deficiency such as the hippocampus and autonomic system. Thus they also suggest the need for wider delivery of the vector into the brain.

For these reasons, other AAV serotypes (e.g. rh.8, 9 and 10[149]) or engineered coat capsids[150] that show an increased ability to cross the blood brain barrier (BBB) have been pursued extensively for non-invasive, brain-wide systemic CNS gene therapy. However the overall efficiency of BBB penetration is still sub-optimal for safely obtaining effective therapeutic doses in the CNS. The requirement for high systemic viral load delivery poses problems for viral transduction of non-neural peripheral tissues which can cause significant off-target effects as well as the induction of host immune responses. In addition, BBB crossing serotypes can show increased tropism for glial cells as opposed to neurons which are most often the target of gene therapy methods[149].

The other notably effective method of CNS access for gene therapy is the use of transduced hematopoietic progeny cells that can migrate into the CNS. Such an approach is most commonly applied to lysosomal storage diseases, where supraphysiological levels of expression of deficient enzymes by transduced hematopoietic progeny cells that have migrated to CNS tissues can scavenge built up toxic materials. In one promising case, *ex-vivo* lentiviral gene therapy of hematopoietic stem cells followed by infusion back into patients showed promising efficacy in the treatment of metachromatic leukodystrophy (MLD), a neurodegenerative lysosomal storage disease caused by the deficiency of the arylsulfatase A (ARSA) gene. Following gene therapy, patients showed reconstituted ARSA activity in the cerebrospinal fluid and halted disease progression[151]. However, to date, ex-vivo gene therapy has largely shown limited efficacy due to difficulties in achieving stable high supraphysiological levels of deficient enzymes in progeny cells. In addition, such approaches can only replace deficient proteins in transduced hematopoietic progeny cells to modify the environment of the CNS rather than target genes in neurons themselves, which may be necessary for the treatment of other more common acquired and hereditary CNS diseases.

1.2.2.3 Challenges of network-targeting for therapy

Despite the accumulating evidence that network dysfunction is at the heart of many neurological disorders, our ability to safely manipulate networks for therapy is limited. None of the existing viral vectors or therapeutic molecules currently in development for the treatment

of neurological disease are circuit-specific. As such, targeting to affected circuits in disease are attempted non-specifically through topical local injections or brain-wide delivery through the blood or cerebrospinal fluid. However, such methods are yet insufficient to obtain the high therapeutic doses into circuits that are required for effective treatment.

The first challenge that determines the success of any CNS therapy, is the ability to obtain efficient access into the CNS tissue. However, due to the high anatomical protection the CNS receives from the skull, spinal cord and blood brain barrier (BBB), access into the CNS can be difficult. Compared to other peripheral tissues that are separated from the blood by single-cell thick capillaries, the brain is perfused by a specialised three-layered BBB consisting of endothelial cells with tight junctions, astrocyte end-feet and pericytes[152]. As a result, the BBB prevents large or hydrophilic molecules from entering the brain from the blood, making the development of minimally invasive, peripherally delivered therapies that can access the CNS very difficult. Pharmacological compounds that can efficiently penetrate the BBB must be lipophilic or recognised by membrane transport systems. In addition, they also must not be substrates for multi-drug resistance transporters such as the efflux transporter P-glycoprotein that is expressed in the BBB[153,154]. Due to the difficulties faced by traditional pharmacological methods, there has now been an increasing interest in novel delivery systems into the CNS, such as lipid based nanocarrier systems[155] or genetically engineered viral vectors that can cross the BBB[156]. Although non-viral vectors are promising due to their more simple production methods and preferred safety profile, they show relatively low efficacy and require repeated administrations as they only mediate transient effects which increases the risks of triggering immune responses. Therefore, viral gene therapy for long-term stable expression of disease modifying proteins is a highly pursued therapeutic strategy not only for hereditary but also for acquired neurological conditions.

Second, even with efficient access into the CNS, the density and high connectivity of neurons in the CNS makes specific dysfunctional circuits very hard to target selectively *in vivo*. No therapeutic agents in development for pharmacological or viral gene therapy in the CNS are yet trans-synaptic or circuit specific. Although a degree of spatial specificity can be obtained through direct local stereotactic injections of therapeutic agents, they cannot cover the widespread circuit dysfunctions that underlie many neurodegenerative and neuropsychiatric diseases and this is often a reason for the lack of or reduced therapeutic efficacy as discussed in the previous subsection.

In contrast, we propose that such widespread brain-wide circuits can uniquely be targeted using single injections of trans-synaptic viral vectors such as the rabies virus which has been used extensively for trans-synaptic circuit tracing in experimental animals. Unlike pharmaco-

logical compounds or AAVs, the rabies virus is highly neurotropic and can propagate through functional synapses even if they are far away from the initial site of delivery. This not only enables long-range circuit targeting in the brain by single injection, but also enables efficient peripheral delivery of the virus into the CNS as the rabies virus can infect neuron terminals in peripheral injection sites. For example, given the involvement of both upper and lower motor neurons in the pathogenesis of ALS, one would ideally want a vector to target both spinal α motorneurons and pyramidal Betz cells in layer V of the motor cortex. A single rabies virus injection into the muscle would enable efficient infection of peripheral motor neuron terminals in the muscle. Subsequently the rabies virus would be transported retrogradely into the spinal motorneuron soma and from there trans-synaptically retrogradely propagate into upper motor neurons in the motor cortex (Fig. 1.1A). In contrast attempts to target both upper and lower motor neurons using topical injections of non-transsynaptic vectors would require two separate invasive injections into the spinal motor neurons and cortical upper motor neurons (Fig. 1.1B). Alternatively, systemic delivery would not enable as efficient delivery of therapeutic agents specifically into the motor network. As such, the rabies virus in theory overcomes two of the major problems hindering CNS therapy development (the BBB and the lack of network specificity).

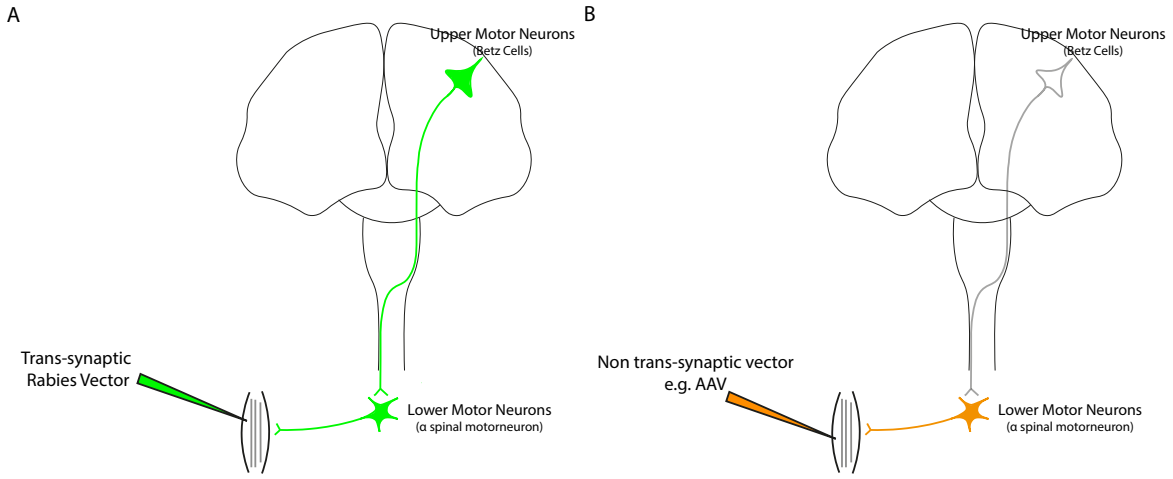


Fig. 1.1 Rabies virus enables non-invasive circuit-specific targeting for therapy *in vivo* (A) Single rabies virus injection into the muscle enables targeting of both lower and upper motor neurons affected in ALS. **(B)** Single injection of non-transsynaptic molecules or vectors can only target lower motor neurons following intramuscular injection.

In order to develop the rabies virus from a experimental tool into a therapeutic vector, it needs to be engineered to have key characteristics such as lack of pathogenicity, low immunogenicity, high neurotropism and efficient execution of permanent gene modification. In the

first experimental chapter of my thesis, I present the design and development of a second generation self-inactivating rabies virus (SiR 2.0) vector that demonstrates increased suitability for therapeutic use (self-inactivation, minimal immunogenicity and increased neurotropism). In the second experimental chapter I present the incorporation of the CRISPR/CAS9 system into the SiR genome, leading to the first production of a circuit-specific gene-modifying viral vector (SiR CAS9 gRNA) that could be used to knock out specific genes within specific circuits *in vivo*.

1.3 A novel mechanism of circuit communication: interneuronal protein transfer

The brain computes its functions through neural circuits. By and large, the focus of neural circuit communication has been defined by synapses and the rapid transfer of electrical signals between one neuron to another through ligand-gated ion channel (LGIC) signalling. LGIC signalling is optimised to transmit quick, transient and membrane-localised changes in the responding cell and so is ideal for the fast transmission of electrical signals from one cell to another. However, interneuronal circuit communication is not defined entirely by the fast transmission of electrical signals from one cell to another. Throughout development and adult life, interneuronal communication must also be able to induce longer-term modifications of synaptic properties which involve changes in gene expression and cell protein composition. Although LGIC signalling, with all its complexities, can induce structural changes in neurons through secondary effects of membrane potential changes, the signalling of longer lasting changes to neuron phenotype (such as synaptic strength, cell tropism or axon growth) are more suited to alternative signalling molecules that can directly alter the cell's transcriptional and translational output. For example, steroid hormones can diffuse through a cell membrane and bind directly to cytoplasmic receptors to form hormone-receptor complexes that directly alter gene expression.

Not surprisingly, neurons are known to be able to communicate by many other signal transduction mechanisms such as G-protein coupled receptor signalling[157], neuromodulators[158] and neurotrophic factors[159]. Recently, there has been a growing body of evidence suggesting that neurons can also communicate with one another through the direct transfer of effector molecules such as RNA, lipids and proteins[160,161]. For example, the commonly observed transsynaptic movement of pathological protein aggregates in neurodegenerative disease begs the question: why do these proteins move interneuronally at all? Are these observations revealing an underlying physiological process by which neurons communicate with one another by methods other than those classically studied? Since the 1980s, the trans-synaptic transfer of proteins has been used as a method of neural network tracing, showing that this process is not purely pathological[162–168]. Recently, there have been increasing discoveries of more physiological examples of trans-synaptic protein and mRNA transfer such as interneuronal transcription factors (e.g. Otx2[169]) and retroviral gag-like proteins (e.g. Arc[170]). Such interneuronal protein transfer is proposed to play a role in the development and maintenance of neuronal populations, cellular network architecture and modulation of synaptic properties and intercellular connectivity[171].

This form of communication is inherently different from conventional signal transduction mechanisms and is hypothetically an extremely efficient method of controlling the concerted change of protein composition in pre and post-synaptic terminals. The advantages over classical signal transduction through secreted messenger molecules include that the transfer of such effector molecules would be immediately spatially specific to desired synaptic connections and that it economically overcomes the difficulties of communicating transcriptional changes between the nucleus of a neuron to its axon or dendritic terminal. In this section of the introduction I will present the accumulating evidence for the transfer of protein between pre- and post-synaptic terminals as a novel method of interneuronal communication. Finally, in the last experimental chapter (Chapter 5), I describe work that has been conducted to develop novel molecular tools and experimental systems that would enable further exploration of the mechanisms behind this observed phenomenon of interneuronal protein transfer.

1.3.1 Pathological protein movement in neurodegenerative disease shows circuit logic

Histopathological studies reveal that many of the protein aggregates which characterise neurodegenerative disease show a stereotypical sequential pattern of spreading that correlates with synaptic connectivity (Table 1.1). Experimental evidence also suggests that the propagation of these proteins may be trans-synaptic. For example, exogenous overexpression of neurodegeneration-associated proteins in animal models leads to propagation of those proteins to synaptically connected brain areas (e.g. pathological human tau (P301L) expression in the enthorinal cortex leads to propagation to the hippocampal formation and cingulate cortex[119,172]). In *in vitro* systems, activity dependent release of tau[117] has been observed whilst addition of botulinum toxin, which inactivates synaptic vesicle fusion, blocks transneuronal spreading of mutant huntingtin protein in mixed organotypic brain slice co-cultures[118].

We propose that the visualisation of the trans-synaptic movement of these disease-causing protein aggregates may be revealing a more general and universal process of interneuronal protein transfer. As the protein aggregates that form in disease are easier to visualise and trace experimentally, much of the research on the movement of these proteins focuses on observing the spread of mutant or aggregated forms. However, we suggest that this ability for transneuronal protein transfer is most probably not limited to pathological processes. For example, wild type tau is released from neurons in an activity dependent manner[117] and wild type tau pathology has been shown to spread faster to more distant brain areas then its mutant counterpart[173].

1.3.2 Traditional protein based circuit tracing methods demonstrate trans-synaptic protein movement

Proteins have historically been used to trace neural circuits. The use of proteins as circuit tracers implies that some proteins are specifically transported through synaptically connected circuits.

For example, plant lectins[162] and wheat germ agglutinin[163,164] were used as traditional circuit tracers in olfactory[165], vestibular, auditory[166] and peripheral sympathetic circuits[167] in a variety of different living organisms. To demonstrate an example, unilateral intranasal irrigation with lectin-HRP leads to positive labelling unilaterally in the olfactory nerve, the olfactory bulb and the terminal forebrain regions of mitral and tufted cells[165]. Bacterial toxins such as cholera toxin and tetanus toxin have also been shown to have trans-neuronal tracing properties in certain experimental conditions *in vivo*[167,168]. Interestingly, the patterns of labelling intensities of primary and second order neurons suggest that different proteins have different mechanisms of trans-synaptic labelling. For example, tetanus toxin shows relatively low levels of labelling in both primary and secondary neurons, suggesting a high efficiency of trans-synaptic transfer. In contrast, WGA results in extremely high labelling intensities of primary neurons but weak labelling of secondary neurons[167]. This suggests that the trans-synaptic transfer efficiency of WGA is low and that its use for trans-synaptic tracing is dependent on the high levels of receptor-mediated endocytic uptake into the primary neuron by binding to sialic acid and N-acetyl glucosaminyl residues on the plasma membrane.

Labelling of proteins by injection of radioactively labelled amino acids have also shown similar results[174,175]. For example, intraocular injection of tritiated amino acids results in the visualisation of radioactive proteins in optic axons, lateral geniculate body and contralateral visual cortex suggesting that these amino acids are assimilated into proteins which are transferred through known visual pathways[174].

1.3.3 Physiological examples of functional interneuronal protein movement in development and adult physiology

Neuronal transcription factors such as homeoproteins encode positional information along various axes of neural structures throughout development[176]. Homeoproteins can also be involved in the regulation of adult physiology (e.g. Engrailed is expressed in adult aminergic nuclei[177]). Interestingly, these homeoproteins, which play such pivotal roles in the transfer of positional information, can also be transferred between cells by atypical secretion and internalisation methods[169]. Their DNA binding homeodomains contain highly conserved

and distinct internalisation and secretion signals which have been confirmed through deletion studies and gain of function studies using synthetic peptides[178,179]. Thus, homeoproteins can spread transneuronally, bind directly to DNA/RNA to control transcription/translation[180] and thus regulate neuronal functions such as neurotrophism[179], neural morphogenesis[181], axonal arborisation and target choice[182]. For example, the trans-synaptic transfer of Otx2 is physiologically implicated *in vivo* to activate postnatal plasticity in visual pathways[169].

Another more recent discovery of a physiological neuronal protein showing a clear ability for trans-synaptic transfer was that of Mammalian Activity-Regulated Cytoskeleton-Associated protein (Arc). Arc is an activity-regulated immediate early gene (IEG) which mediates many aspects of synaptic maturation and plasticity, learning and memory[183]. For example, immediately after a rat is exposed for 5 minutes to a novel environment, there is a brief burst of Arc transcription in memory consolidation networks which quickly returns to baseline levels (nuclear signals return to baseline within 16 minutes whilst cytoplasmic and dendritic mRNA induction decreases more slowly over a time-scale of 20-60minutes)[184]. Mutations in this gene have also been implicated in neurodevelopmental disorders such as autism[185] and schizophrenia[186]. However, despite its clear importance in many aspects of brain function, the molecular functions and mechanisms of Arc are still unclear. Recently it was uncovered that the Arc protein, which contains a retroviral group antigen (Gag)-like domain, could self-assemble into virus-like capsids that can package RNAs and are transferred trans-synaptically through extracellular vesicles and endocytosed by receiving partners[170,187]. This trans-synaptic transfer is dependent on the 3'UTR of Arc mRNA, and it was also shown that pre-synaptic Arc mRNA with an intact 3'UTR is necessary for synaptic expansion and plasticity across the *Drosophila* neuromuscular junction[170].

1.3.4 Remaining questions for interneuronal protein transfer as a potential interneuronal communication mechanism

Despite the rapidly growing body of evidence to support transneuronal protein transfer, we still have very little understanding of its mechanistic details, control and functional implications. A particular area of controversy is that, despite much observational and indirect experimental support, there is still no direct evidence that clearly shows that these proteins travel trans-synaptically. In addition, many other fundamental questions remain unanswered. For example, how universal is this process? How many different proteins are being transferred transneuronally and what are their identities? When and where are these proteins transferred in development and adulthood? Does this process play a functional role

in normal circuit function or malfunction? Answering these questions is of fundamental importance as they address a key molecular aspect of interneuronal communication and functional neural circuitry. In addition, a greater understanding of protein transfer may aid in the identification of new potential drug targets to prevent pathological protein propagation in neurodegenerative disease or to control circuit activity in psychiatric conditions.

1.4 Aims of the thesis

In the introduction of this thesis, I have argued for the importance of neural networks in brain function. I have highlighted that despite much progress over the past century, we still face many technical difficulties when trying to experimentally manipulate or visualise network structure and function in live animals. This has resulted in gaps between the cognitive theories of brain function and the biological investigations of these theories. In the three following experimental chapters of my thesis I present work which aims to develop novel circuit-specific experimental tools that can help to answer some of the important remaining questions regarding circuit function. I also highlight the potential uses of these network specific tools for potential gene therapy in the treatment of neurological disease.

1. In the first experimental chapter, I describe the design, production and *in vivo* characterisation of a novel SiR 2.0 circuit-targeting viral vector. The design of this vector combines the advantageous qualities of two previously existing trans-synaptic rabies vectors to create a new improved trans-synaptic vector that: a) self-inactivates, thus enabling life-long investigation into labelled circuits b) is minimally immunogenic during its short period of transcriptional activity before self-inactivation and c) shows higher efficiency of trans-synaptic targeting of peripheral circuits.

2. In the second experimental chapter, I aim to functionally incorporate into the SiR 2.0 genome the components of the CRISPR-Cas9 genome editing technology. The motivation is to create a circuit specific viral vector that can efficiently and permanently modify endogenous genes of host cells *in vivo*. This would enable long-term experiments to establish causality between the role of specific genes within circuits and the computation and generation of behaviours in live animals. It also has the promising potential to be developed as a gene-therapy vector that can efficiently target large-scale diseased networks in neurodegenerative and psychiatric conditions - something which is yet hard to achieve with the existing non-circuit specific vectors.

3. In the third and final experimental chapter, I aim to develop novel methods to facilitate further mechanistic investigation into the observed phenomenon of interneuronal protein transfer. First, I investigate the efficacy and suitability of different experimental systems which can be used to study protein transfer in neural networks *in vitro* and *in vivo*. Second, I aim to develop an unbiased top-down proteomic approach to visualise and identify novel proteins showing inter-neuronal movement. Third, I aim to develop a small split-CRE based protein tag that would enable the amplification of interneuronal protein movement signals whilst also genetically modifying protein-receiving neurons.

Chapter 2

Materials and Methods

nomenclatureDESeq2Differential gene expression analysis based on the negative binomial distribution, version 2

2.1 Mice

All animal procedures were conducted in accordance with the UK Animals (Scientific procedures) Act 1986 and European Community Council Directive on Animal Care. All mice were bred and maintained in pathogen and opportunistic agents-free conditions and monitored quarterly in the ARES Animal Facility in the Babraham campus. A complete list of the wild-type (WT) and transgenic lines used in this thesis, with information of expression patterns and origins can be found in Table 2.1. All transgenic mice were isogenic in a C57BL/6 background. Biopsies for the genotype screening were taken by the LMB and ARES biological service groups and the genotyping reactions performed by the LMB genotyping facility.

All experimental procedures were performed in the LMB Animal Facility. Animals were group-housed in a 12 hours light/dark cycle (7 a.m. to 7 p.m.) with food and water *ad libitum*. Experiments in adult animals were performed on mice aged between 8 and 12 weeks. Intramuscular injections were carried out in pups aged P3-P5.

Line Name	Expression Pattern	Source
C57BL/6J	-	The Jackson Laboratories: 000644
CD1 (ICR)	-	Charles River: 022
Rosa-LoxP-STOP-LoxP -tdTomato	CRE-dependent expression of the tdTomato reporter gene	The Jackson Laboratories: 007914
Rosa-LoxP-STOP-LoxP -YFP	CRE-dependent expression of the YFP reporter gene	The Jackson Laboratories: 006148

Table 2.1 List of mouse lines

2.2 Cell lines

HEK293T, BHK-21 and mouse Neuro2A cells were purchased from ATTC. All cell lines were maintained in Dulbecco's modified Eagle's medium (DMEM, Gibco) supplemented with 10% fetal bovine serum (FBS, Gibco) and 100units/mL penicillin and 100μg/mL Streptomycin unless otherwise specified. The cells were maintained in a humidified incubator with 5% CO_2

and were passaged three times a week. For MutRS or BioID2 studies, AnI (Jena Bioscience) or biotin was added at the stated concentrations to the culture medium respectively.

2.3 Plasmid design and construction

All plasmids produced in this thesis were generated by gibson assembly or by use of classical restriction enzymes, backbone dephosphorylation and DNA ligation (T4 DNA ligase, NEB).

The pAAV vectors used in this thesis were generated by gibson assembly to insert sequences between the inverted terminal repeat (ITR) sequences of a rAAV plasmid backbone originally supplied by Dr. Silvia Arber. Expression constructs for cell mixing experiments were cloned into the pCAG-Brainbow 3.0 backbone (plasmid #45176, Addgene). Lentiviral vectors (pLenti) were derived from an empty third generation lentivirus transfer vector plasmid supplied by Dr. Marco Brancaccio ("361 polylinker", originally pCCL-SIN-18PPT.Pgk.EGFP-WPRE). The backbone was opened with XbaI and KpnI followed by gibson assembly with PCR amplified promoters (CMV or CAMKII) and different inserts (e.g. iCRE, GFP, P2A-human wild type 0N4R Tau). Inserts were designed with 30 base pair overlapping regions for efficient assembly.

The F2, F6 and F9 linkers were synthesised using gBlock Gene Fragments (Integrated DNA Technologies) and inserted by Gibson Assembly.

Coding sequences for the following genes were PCR amplified from gifted plasmid vectors:

- Otx2 was a gift from Malin Parmar (Addgene plasmid #35001)
- WGA was a gift from Kwang-Soo Kim (Addgene plasmid #80337)
- MutRS was a gift from David Tirrell (Addgene plasmid #63177)
- BioID2 was a gift from Kyle Roux (Addgene plasmid #74224)
- GFP nanobody (GBP) was a gift from Joe Watson[188]

2.4 Viral production

2.4.1 Lentivirus production

Human HEK 293T cells were seeded onto 10 cm plates and transfected with polyethylenimine (PEI) the following day with 3 μg ENV plasmid (VSV-G), 5 μg Packaging plasmid (pMDLg/pRRE), 2.5 μg of pRSV-REV and 16-18 μg of transfer plasmid per plate using

polyethylenimine (PEI). Five hours later, transfection medium was removed and replaced with fresh medium. After fourteen hours, medium was removed once more and replaced with 5mL fresh medium to begin the collection of lentivector. The cell supernatants were collected after 24 hours and filtered with a 40 μm filter to remove cellular debris (STERIFLIP millipore units). The supernatants were either used directly to transduce cells *in vitro* or purified by ultracentrifugation at 50, 000 RCF for 2h 30min at 4°C on a sucrose cushion for *in vivo* injections. The viral pellets were resuspended in PBS and stored at -80°C until further use. Productions were functionally titered by performing *in vitro* infections on HEK293T cells, with serial dilutions of viral preparations.

2.4.2 AAV production

For each viral preparation, human HEK 293T cells were seeded onto 10 dishes of 15cm plates. At 80% confluency, cells were transfected with AAV helper plasmid (Rep/Cap), AAV vector (with expression cassette) and pHGTI-adenol (to provide adenoviral helper function) using PEI. 60h after transfection, cells were dislodged from the plate and collected by centrifugation at 2000RPM for 10 minutes. The cell pellet was then washed with 15mL PBS and resuspended in 10mL lysis buffer (150mM NaCl, 20mM tris pH 8.0, filter sterilized and stored at 4ºC). Cells were subsequently freeze-thawed three times using dry ice/ethanol and a 37'C water bath. Cell lysates were dounce homogenized 20 times to destroy cell nuclei and put into a new falcon tube. $MgCl_2$ and benzonase (Sigma) was added to a final concentration of 1mM and 250U/mL respectively and the mixture was incubated at 37ºC for 10-15 min to dissolve genomic DNA/protein aggregates formed during freeze-thaw cycles. Cell debris was spun down at 4000 RPM, 4ºC for 20min and the viral supernatant was purified further using 17, 25, 40, 60% Iodixanol gradients in PBS, 1mM MgCl2, 2.5mM KCl (Optiprep, Sigma). 0.0005% phenol red was added to alternate layers for visualisation. Viral lysate (14-15mL) was added to the top of the gradient in a 36mL Optiseal tube (Beckman 362183) and centrifuged for 90min at 48-50,000rpm (Beckman VTi50 rotor) at 16ºC. The 40% viral fraction was harvested, diluted in PBS and centrifuged for 30min 3500rpm at 4'C (Eppendorf 5810R) in Millipore Amicon 100K columns (UFC910008) for a total of 3 times for removal of iodixanol and concentration of virus. 150-250μL of virus was recuperated after the last spin, aliquoted and stored at -80ºC. The titer was determined by calculating viral genome copies from RT-qPCR using primer pairs for the CMV promoter, and serial dilutions of linearized vector plasmid for the standard curve. Purified AAVs were treated with DNase to remove residual plasmid DNA from purification (15 min, 37'C, DNAseI Roche followed

by 10min inactivation at 95'C) and then incubated with proteinase K (Sigma Aldrich, 37'C, 15 min) before quantification with RT-qPCR.

2.4.3 WT and SiR ΔG rabies virus production

2.4.3.1 Generation of SiR virus production cells

HEK293T cells expressing GFP, rabies G (chimeric oG or N2C), and TEV protease to high levels needed to be selected for efficient viral productions. Lentiviruses encoding these proteins along with a puromycin selection marker were produced to transduce the exogenous cassette into production cell lines. For infection, HEK293T cells were split into T75 flasks at 40% confluency with 5mL lentiviral supernatant and 5mL culture medium. The following morning, viral supernatant was changed to fresh cell culture medium. 24 hours after medium change, successfully transduced cells were selected for with 1ug/ml of puromycin. A negative control of uninfected cells were also incubated with 1ug/ml of puromycin to assess puromycin toxicity. Usually, transduced cells were selected over a period of 2-3 days. Following puromycin selection, cells were additionally sorted for high GFP expression by FACS on a Beckman Coulter MoFlo High Speed Cell sorter.

2.4.3.2 Recovery of SiR virus from plasmids

For each viral production, 4 wells of a 6 multi-well plate were seeded with HEK-TGG (HEK 293T expressing TEV protease, B19 oG or N2C G protein, and GFP) and transfected with PEI at around 85% confluency with:

- the pSAD or pN2C SiR genome (12μg)
- 4 plasmids expressing the viral proteins N, P, L, G for either the N2C or B19 strain (6, 6, 3, 3 μg respectively)
- pT7 polymerase (2.4μg)

The following day, supernatant was replaced with fresh medium with additional gentamicin at 50μg/mL. After medium change, cells were left to recover from transfection and pooled into a T75 flask when reaching 100% confluency (typically 24-48 hours after medium change for Sir B19, 48-72 hours after medium change for SiR N2C). When the re-plated cells approached 80% confluency, the medium was changed to 2%FBS/DMEM solution and incubated at 35°C, 2% CO_2 for 2-3 days. For viruses that could be visualised with a fluorescent marker, typically one would observe viral spreading in >50% of production cells around this time. Once viral infection was visible in >50% of cells, or once cells approached

80% confluency, viral supernatant was collected from cells after incubation for 2-3 days at 35°C, 2% CO_2. Two rounds of cell expansion and supernatant collection were conducted and used for subsequent high titer viral production.

2.4.3.3 High-titer SiR production

HEK-TGG cells were plated on a T75 flask to 80-85% confluency and and transduced with 5-6mL of supernatant collected from cells transfected with plasmids. The following day, the medium was exchanged to 2% FBS/DMEM and cells incubated at 35°C and 3% CO_2. Once confluent, cells were split 1:2 into a T175 plate. When cells approached 80% confluency (and for visible viruses, when 80% of cells were infected with virus), the medium was changed to 2% FBS/DMEM and incubated for 24-48 hours for the first supernatant collection. Subsequently, fresh 2% medium was added to the cells and incubated for another 24-48 hours for the second supernatant collection. The length of supernatant collection was determined by the kinetics of viral production and cellular viability. After collection of viral supernatant, cell debris was removed by centrifugation at 4000 rpm for 10 min followed by filtration with 0.45 mm filter (Millipore, SLHV033RS). The virus was concentrated by ultracentrifugation (4°C, 25000rpm, 4h or longer, Sw32Ti rotor) on a 20% sucrose cushion[189], resuspended in 50-75uL PBS and stored at -80°C.

2.4.3.4 WT ΔG rabies virus production

WT ΔG N2C or B19 viruses were produced very similarly to SiR viruses but with HEK-GG cells (HEK 293T expressing the B19 oG or N2C G protein and GFP). Typically, incubation times for supernatant collection were approximately halved for WT viruses compared to SiR viruses due to increased viral replication and spreading kinetics in cell culture.

2.4.4 Rabies virus titration

2.4.4.1 Functional titration *in vitro*

Functional titration of rabies virus was performed through *in vitro* infections of HEK293T cells with serially diluted viral preparations as described in previous reports[189–191]. HEK293T cells were plated onto 6 well plates and at 70-80% confluency infected with serially diluted virus. 48h p.i. cells positive for virally expressed fluorescent markers were quantified by FACS. The number of infectious units per ml was calculated referring to the dilution giving a low multiplicity of infection (1-10%) by multiplying the number of infected cells by the

dilution factor. This method is referred to as the "traditional" method of titration in this thesis as it is the most commonly used method used for the titration of previous experimental rabies virus variants.

For the titration of SaCas9-gRNA expressing rabies viruses without fluorescent markers, the same protocol was used for the infection of cells with serially diluted virus. However, cells were then subsequently stained for 1 hour at room temperature with a Cy3 conjugated anti-FLAG antibody and imaged on an inverted Leica SP8 confocal microscope. Virally infected cells were then counted manually using the ImageJ image processing program.

2.4.4.2 RT-qPCR

Due to cell tropism and fluorescence intensity biases in the traditional titration method, relative titers between different rabies virus variants (WT vs SiR and B19 vs N2C) was quantified by RT-qPCR. Viral genomic RNA was first obtained either by direct lysis of viral productions using a solution of IGEPAL CA-630+BSA or by RNA purification of viral productions using commercial kits (Directzol RNA MiniPrep, ZymoResearch). Fold changes of viral genome copies between productions was quantified by RT-qPCR using Taqman probes for CRE sequences present in both SiR 1.0 and SiR 2.0 viruses.

2.5 *In vitro* experiments

2.5.1 Single particle rabies virus sequencing

RNA viral genomes were extracted using a Direct-zol RNA Miniprep kit (Zymoresearch R2051). The purified RNA genomes were converted to cDNA using a barcoded primer that annealed to the 3' rabies virus leader sequence and contained a random 8-nucleotide sequence on its 5'end[192] (Superscript IV, Invitrogen). Thus, a unique random index was incorporated onto the 5' end of each cDNA molecule corresponding to each individual viral particle's RNA genome. The cDNA sequence was then amplified from the 5' end through to the first half of the rabies virus P gene using Q5 Hot Start High-Fidelity DNA Polymerase (NEB) giving a PCR amplicon of approximately 2kb length. The amplicons were extracted from an agarose gel using a QIAquick Gel Extraction Kit and inserted into a standard cloning vector pBluescript KS (+) opened with restriction enzymes KpnI and KbaI (NEBuilder® HiFi DNA Assembly). The plasmids were transformed into Stbl3 competent E. coli, and 48 clones were isolated and purified for sequencing. The sequencing procedure is summarised in Fig. 3.3. For each clone, the 8-nucleotide index and 3'-end of the N gene were sequenced. The

sequences were aligned using Snapgene 4.1.9 with the reference sequence based on plasmid pSiR-N2C-iCRE-P2A-mCherryPEST used for viral production. The primer sequences are as follows:

- N2C RT primer with 8nt barcode: TCAGACGATGCGTCATGCNNNNNNNNT-CACGCTTAACAACAAAACCAGAG
- Fw PCR primer: CCACCGCGGTGGCGGCCGCTCTCAGACGATGCGTCATGC
- Rv PCR Primer: CTAAAGGGAACAAAAGCTGGGTACTGAGAGAAAGCAGACG-TTGTCTCC

2.5.2 *In vitro* immune response characterisation

Mouse neuro2A cells were seeded onto polylysine coated glass cover slips in 6 well plates and infected at 80% confluency with WT B19 oG GFP rabies virus at a titer of $5x10^6$ to give approximately 60% infected cells. For the positive control, 100 ng/ml of murine recombinant TNF-α (aa 80-235) was added to the culture medium (410-MT, Bio-Techne Ltd.). For the negative control, equal volumes of PBS were added to the culture medium. For imaging of immune responses at specified time points post-infection, cells were then fixed for 10 minutes with PFA 4% at room temperature and washed in PBS. Cells were incubated with anti-NF-kB (p65) and anti-phospho p38 MAPK antibodies under the conditions specified in Table 2.4 for immunofluorescence analysis. The glass coverslips were mounted in Fluoromount-G (ThermoFisher) with DAPI for image acquisition on an inverted Leica SP8 microscope. For RT-qPCR analysis of immune response genes, RNA was extracted from cells at specified time points using RNeasy Mini kits (Qiagen). RT-qPCR was conducted using the rotor-gene SYBR Green PCR Kit (Qiagen) with the primer pairs listed for various immune response genes in Table2.2 which were previously experimentally verified[193–195]. Data was acquired on a Corbett Rotor-Gene 6000 real-time PCR machine. Fold changes between samples was calculated using the double delta Ct analysis with GAPDH as the housekeeping gene.

2.5.3 T7 endonuclease mismatch cleavage assay

For screening of NeuN gRNA sequences mouse neuro2A cells were transfected with CMV-SaCas9 and U6-gRNA DNA expression plasmids. For tests of rabies virus mediated CRISPR-Cas9 targeting, mouse neuro2A cells were infected with WT ΔG B19 rabies viruses expressing gRNA and SaCas9. 3 days post-transfection and/or infection, genomic DNA was

extracted from infected cells (Qiagen, DNeasy Blood Tissue Kits). The NeuN locus (exon 1) was PCR amplified using the following primers:

- Fw PCR primer GGCAATCAGGACGCCACGG
- Rv PCR primer GCTACACTTCCTACCTCCCAGATGG

PCR amplicons were then denatured-renatured and incubated with the T7 endonuclease I enzyme for cleavage of heteroduplexes. Fragments were then visualised by gel electrophoresis on a 2% agarose gel with SYBR Gold Nucleic Acid Gel Stain (Invitrogen, S11494). Fragmentation analysis was conducted on BioRad Image Lab software. Cleavage efficiency was calculated using % cleavage of PCR amplicons (100 x quantity of the two fragmented bands/total quantity of unfragmented and fragmented bands).

Gene	Fw Primer Sequence (5'-3')	Rv Primer Sequence (5'-3')
IL 6	TGCCTTCTTGGGACTGATGC	GACAGGTCTGTTGGGAGTGGT
RIG I	GCAAGTGCTTCCTCCTGACC	ATGCGGTGAACCGTCTTTCC
MCP1	AGCCAACTCTCACTGAAGCC	TGTGACTCGGACTGTGATGC
TNFa	GCACTCCCCCAAAAGATGGG	TGTGAGGGTCTGGGCCATAG
IFIT1	AGGCTGGAGTGTGCTGAGAT	TCTGGATTTAACCGGACAGC
IFIT2	CACCTTCGGTATGGCAACTT	GCAAGGCCTCAGAATCAGAC
OASL1	TCCTGGAGGACTGGTTTGAC	ACTGAAGACGTGGACCATCC
Mx1	CGGCAGAAGCTTGCCAAATT	AGACCAGTCCTCTAGGCAGG
IFNa2	TCTGTGCTTTCCTCGTGATG	TTGAGCCTTCTGGATCTGCT
IFNa5	CTGCCTGAAGGACAGAAAGG	TCATTGAGCTGCTGATGGAC
GAPDH	GGAGAAGCTGCCAATGGATA	TTACGCTTGCACTTCTGGTG

Table 2.2 Primer sequences for mouse immune response genes used in this thesis

2.5.4 Testing of gRNAs using a double reporter cell system

HEK 293T cells were seeded onto 6 well plates and transfected with polyethylenimine (PEI) at 80% confluency with 2.5μg of the double reporter plasmid and 2.5μg gRNA-SaCas9 expressing plasmid. Medium was changed 24 hour post-transfection and 72 hours post-

transfection cells were trypsinised and prepared for FACS analysis on a Becton Dickinson LSRII SORP Flow cytometer.

2.5.5 Differentiation of neuro2A cell line

For neuro2A differentiation, cells were plated at 20% confluency and the medium was changed to DMEM + 0.1% FBS the following morning. During differentiation, the cell medium was replaced once every two days.

2.5.6 Generation of neuro2A synthetase cell line

Wild type neuro2A cells were transfected with linearised ROSA plasmid using PEI. Positive cells were selected with geneticin for a course of 4 weeks (0.4mg/mL selection concentration, 0.2mg/mL maintenance concentration). Subsequently, we carried out two rounds of fluorescence-activated cell sorting (FACS) for mCHERRY positive cells using a MoFlow sorter to obtain a stable mCHERRY positive cell line of 99% purity.

2.5.7 Cell mixing experiments

For cell mixing experiments, donor cells and acceptor cells were initially cultured in separate 6 wells or 10 cm dishes and transfected with a donor or acceptor plasmid construct respectively using PEI. The following day, cells were trypsinised, pelleted down and incubated in PBS with 2.5mM Mg^{2+} and DNAse1 for 15 minutes. Cells were then mixed at specificed ratios (usually 1:1) and replated. For confocal imaging studies, cells were grown to confluence on L-polylysine coated coverslips, fixed in 4% PFA for 10 minutes at room temperature and washed twice with PBS.

2.5.8 Click reaction for analysis of azide-incorporated proteins by confocal imaging, western blot analysis or mass spectrometry

For confocal imaging of azide-incorporated proteins *in vitro*, the click reaction was performed on neuro2A cells post-fixation using the Click-iTProtein Reaction Buffer Kit (Invitrogen). AF488-alkyne (Sigma) was used as the click detection reagent. For quantification of azide-incorporated proteins by western blot, cells were lysed in RIPA (Radioimmunoprecipitation assay buffer) and purified cellular proteins run on a NuPAGE Bis-Tris Gel (Invitrogen). Proteins were incubated with biotin-alkyne followed by HRP-streptavidin for quantification

with enhanced chemiluminescence western blotting substrate (32106, ThermoScientific). For identification of azide-incorporated cells by mass spectrometry, cells were lysed in UREA buffer and azide labelled proteins enriched from total cell protein lysates using a alkyne agarose protein enrichment kit (C10416, Invitrogen). Following on-resin trypsin digestion samples were run on a mass spectrometer by the LMB mass spectrometry facility.

2.6 *In vivo* surgical procedures

2.6.1 Stereotaxic intracerebral injections

Adult mice aged between 6-12 weeks were used for stereotaxic intracerebral injections. Mice were anaesthetised with isofluorane (induction at 3% in 2L/min of oxygen, maintenance at 1-2% in 2L/min oxygen) and mounted onto a stereotaxic frame (David Kopf Instruments). Rimadyl (2 mg/kg body weight) was administered subcutaneously for anti-inflammatory peri-operative care prior to viral injection. Viruses were injected at the relevant coordinates (Table) relative to bregma. A small 500μm diameter hole was drilled in the skull to allow viral injection with a 35G beveled Hamilton nanofil needle. 400-800nL of virus was injected depending on the experiment (400nL for focal injections, 800nL for more widespread coverage). The needle was left in the brain 5-10min post injection to prevent viral backflow following needle withdrawal. Rabies viruses and lentiviruses were injected at titers of around 3x10^8 infectious units/ml. AAVs were injected at around 5x10^{12} genomic copies/ml.

Target Region	Anterior/Posterior (mm)	Medial/Lateral (mm)	Dorsal/Central (mm)
CA1	-2.3	1.65	1.48
CA3	-2.06	2.65	2.45
NAc	+1.3	1.3	4.75
BLA	-1.65	2.9	4.7

Table 2.3 Stereotaxic coordinates for intracerebral injections Anterior/Posterior and Dorsal/Ventral coordinates were measured from Bregma. Medial/Lateral coordinates were measured from the midline.

2.6.2 Intraocular injections

Adult mice aged between 6-12 weeks were used for intraocular injections. Mice were anaesthetised with isofluorane delivered through a nose cone (induction at 3% in 2L/min of

oxygen, maintenance at 1-2% in 2L/min oxygen). To dilate the pupil, tropicamide was applied to the eye for 2 minutes, followed by pheylephrine hydochloride. Viscotear gel was then applied to the eye before injection with a 65RN Hamilton syringe with a ga34/10mm/pst4 needle. The needle was positioned to enter the eye from behind the ring muscle and pointed towards the posterior side of the eye before injection of 2μl of virus or protein solution. Recombinant AAVs were injected at titers of 8 x 10^{12}. Cholera toxin B subunit-AF488 was injected at 3mg/mL (C22841, Invitrogen). Human mutant A53T alpha synuclein-Dylight594 was kindly prepared by Sophie Morgan from Goedert Lab and injected into the right eye (3μl at 400 μM) in aggregate form. The needle was left in the eye for 1 minute before withdrawal. Viscotears and proxymetacaine were applied to the eye immediately post-procedure and chloramphenicol drops added later during recovery.

2.6.3 Intramuscular injections

C57BL/6J WT or transgenic pups were fostered onto CD-1 foster mothers shortly after birth. At P3, pups were anaesthetized with isofluorane (induction at 3% in 3L/min oxygen, maintenance at 1-2% in 1-2L/min oxygen). A small incision was made in the skin over the quadriceps muscle and 5μL of viral mixture was injected with a picospritzer (Parker Hannifin). The viral mixture consisted of 2.5μL of rabies virus and 2.5μL of AAVs expressing either oG or N2C G protein and a nuclear H2B FLAG epitope for staining. B19 strain rabies viruses were injected at titers of 3x10^9 to enable efficient targeting of motor neurons (titers below 10^9 led to inefficient motor neuron labelling). N2C strain rabies viruses were injected at titers of 3x10^8 for efficient motor neuron targeting. AAVs were injected at 3x10^{12}.

2.6.4 In utero electroporation

CD-1 embryos were electroporated at embryonic day (E) 14.5-15.5. The noon of a day when a vaginal plug was found was designated as E0.5. The mother was anaesthetized with isofluorane in 1L/min oxygen (induction at 3-5%, maintenance at 1.5-2%). Buprex (0.1mL subcutaneous at 0.03mg/mL for 0.1ug/g of body weight) and ritrodine (0.1mL intraperitoneal at 14 mg/mL for 0.05mg/g of body weight) were administered post-induction of anaesthesia. A midline skin incision of the lower abdomen was made followed by a paramidline incision in the muscle sheath. The uterus was then gently pulled out onto a sterile surgical cover. Embryos were gently positioned with blunt ended round forceps and a pulled glass capillary containing DNA (1ug/mL) diluted in fast green dye for visualization was injected into the ventricles using a mouth pipette. The electrodes (CUY650P5, nepagene)

were positioned with the positive electrode covering the somatosensory cortex and the DNA electroporated with transfer pulse settings: pulse number 5, 950ms interval, 50ms pulse length, strength 45V and polarity + (NEPAGENE Super Electroporator NEPA21 Type II). During electroporation, tissues were kept moist with warm NaCl 0.9%. Following the procedure, the uterus was re-inserted into the abdomen and previous surgical incisions sutured. The surgical wound was disinfected with Videne antibiotic solution, and the mother recovered at 37'C for 5-10 minutes. A timer was set from the beginning of anaesthesia induction and the whole procedure limited to 40 minutes to ensure good recovery of the mother and successful pregnancy. Glass capillaries were pulled before procedure to give 0.8-1cm tip lengths (e.g. 1B120F-4 borosilicate glass capillaries pulled on P2000 shutter with sequence: heat 600, FIL 4, velocity 35, del 130, pull 65). The procedure was adapted from Wang and Mei[196] and with much help from Alfredo Llorca Molina.

2.6.5 *In vivo* BioID

Recombinant AAV expressing GFP and biotin ligase (AAV-GFP-BioID2) was injected into the eye as described above in section 2.6.2 of methods. 3 weeks p.i., biotin was administered once intraocularly (0.3mg) and daily intraperitoneally for a period of 7 days (1mg/day). Mice were then perfused for immunohistochemistry analyses (section 2.7.3 of methods).

2.7 Immunohistochemistry

2.7.1 Sample preparation

For immunohistochemistry experiments, mice were terminally anaesthesized with Euthatal (0.2mL) and transcardially perfused with 20mL ice cold PBS followed by 20mL 4% paraformaldehyde (PFA).

For analysis of brain samples, brains were dissected out and post-fixed overnight in 4% PFA at 4°C. Post-fixation, samples were dehydrated in 30% sucrose/PBS overnight for cryoprotection. Subsequently, free-floating coronal sections (35 m thickness) were cut from the brain using a freezing sledge microtome for immunofluorescence.

For spinal cord samples, spinal cords were dissected out and fixed for 4-8 hours at 4°C. Post-fixation, samples were dehydrated in 30% sucrose/PBS overnight for cryoprotection. Spinal cords were subsequently frozen in O.C.T. compound (VWR, Radnor, PA) and sliced into 30 μm sections using a CM1950 cryostat (Leica, Wetzlar, Germany). Sections were

mounted on glass slides immediately after cryosectioning using an anti-roll blade and stored at -80°C.

2.7.2 Antibody staining

Primary and secondary antibodies used in this thesis and the concentrations and incubation lengths are listed in the tables below (Table2.4, 2.5). All incubation steps were conducted at 4'C on a rocking platform. Antibodies were diluted in blocking solution (1% bovine serum albumin and 0.3% Triton X-100 in PBS), unless stated otherwise. Between primary and secondary antibody staining, sample were washed four times with PBS at room temperature.

Primary Antibody	Catalogue number	Floating Sections	Thin Sections
Rabbit RFP	ABIN129578	1:2000, 24h	1:1000, 24h
Goat RFP	AB0040-200	1:100, 72h	1:50, 72h
Chicken GFP	Aves GFP 1020	1:2000, 24h	1:1000, 24h
Rabbit TEVs	PA1-119	1:200, 24h	1:75, 72h
Goat Iba1	NB 100-1028	1:150, 72h	1:100, 72h
Rabbit GFAP	Dako-Z0334	1:500, 24h	1:300, 72h
Rat FLAG	637304	1:500, 24h	1:250, 24h
Rabbit RIG-I	ab45428	1:100, 72h (in 3%BSA, 10% donkey serum, 0.3% Tx)	-
Mouse V5	V8012	1:1000, 24h	-
Mouse CRE	C7988	1:500, 24h	-
Rat OLLAS	MA5 16125	1:300, 24h	-
Rabbit NeuN	ab128886	1:2000, 24h	-
Rabbit FLAG	PA1-984B	1:500, 24h	-
Rat CTIP2	ab18465	1:300, 72h	-
Rabbit NF-kB (p65)	Cell signalling, 8242	1:400, 48h	-
Rabbit phospho p38 MAPK (D3F9)	4511L	1:500, 48h	-

Table 2.4 List of Primary antibodies and working dilutions

Secondary Antibody	Catalogue number	Floating Sections	Thin Sections
Donkey anti chicken AF488	703-545-155	1:1000, 24h	1:500, 24h
Donkey anti RAT AF488	A21208	1:1000, 24h	1:700, 24h
Donkey anti rabbit Cy3	711-165-152	1:1000, 24h	1:500, 24h
Donkey anti goat Cy3	705-165-147	1:1000, 24h	1:500, 24h
Donkey anti goat Cy5	705-175-147	1:1000, 24h	1:500, 24h
Donkey anti rabbit Cy5	711-175-152	1:1000, 24h	1:500, 24h
Donkey anti rat Cy5	712-005-153	1:1000, 24h	1:500, 24h
Goat anti rat Cy5	ab6565	1:1000, 24h	1:500, 24h
Streptavidin-HRP	Thermo Scientific, 21130	1:10000, 24h	-

Table 2.5 List of Secondary antibodies and working dilutions

2.7.3 Visualisation of biotinylated proteins

Biotinylated proteins were visualised either by incubation with Streptavidin-AF555 (3 hours, room temp, 5 μg/mL in blocking solution) or streptavidin-HRP (Table2.5). After incubation with streptavidin-HRP, the biotinylated protein signal was amplified by 10 minutes incubation with the TSA Plus Cyanine 3 system (Perkin Elmer, NEL744001KT). Streptavidin-HRP incubation enables higher sensitivity of biotinylated protein visualisation as the HRP enzyme can subsequently catalyze the the covalent deposition of fluorophores adjacent to the immobilized enzyme resulting in signal amplification.

2.7.4 Image acquisition

For cell counting experiments, spinal cord or brain sections were automatically detected and imaged on a robot assisted Nikon High Content Analysis microscope equipped with a 10X air objective (0.45 NA) operated by NIS-Elements HC software (Nikon, Tokyo, Japan). For higher resolution images of certain immune markers or candidate trans-synaptic proteins, fluorescence was visualised using a Leica SP8 inverted confocal microscope.

2.7.5 Immunofluorescence analysis

For cell survival experiments in the brain, the whole hippocampus was sampled. For cell survival experiments in the spinal cord, the whole region of T5-S5 was sampled. In both

cases, every fourth 35 or 30 m thick section respectively was acquired on a Nikon High Content Analysis microscope and RFP or YFP positive cells manually counted using Nikon HCA software.

For quantification of immune response markers, a total of 10 sections were analyzed for each animal. These sections covered the maximally infected hippocampal area. Positively stained immune cells were then quantified for cell number, cell area, cell circularity, cell elongation and average fluorescence intensity using Nikon General Analysis software.

2.8 RNAseq

2.8.1 RNA extraction and cDNA library preparation

Mice were injected in triplicates with WT B19, WT N2C or SiR N2C rabies viruses. 1 week p.i., mice were culled and the injected hippocampi dissected out for homogenization (Tissuelyser II, QIAGEN) and total RNA extraction with direct-zol RNA Miniprep kit (Zymoresearch R2051). The quality and quantity of purified RNA was analysed using Agilent RNA 6000 Pico chips on a Agilent 2100 Bioanalyzer. Subsequently, cDNA libraries were generated from total RNA using a TruSeq® Stranded Total RNA Library Prep Kit (Illumina, San Diego, CA) using low throughput single adaptors for each individual sample. The samples and pooled library were quantified using the KAPA Library Quantification Kit for Illumina Platforms (KR0405) and sequenced on an Illumina HiSeq 4000 machine (paired end, 150bp read length).

2.8.2 RNAseq data processing and analysis

RNA-seq sequence reads were trimmed using Trim Galore (version 0.4.2) with default parameters to remove the standard Illumina adapter sequences.

For differential expression analysis of host genes, RNA-seq reads were mapped to the GRCm38v95 mouse genome assembly using STAR version 2.5.3amodified167 and reads with MAPQ scores <20 were discarded. Downstream quantitation was conducted using the RNA-seq quantitation pipeline in SeqMonk software (www.bioinformatics.babraham.ac.uk/projects/seqmonk/). Raw counts were used to conduct differential expression analysis using the DESeq2 statistical filter. Genes were considered to be differentially expressed if $p<0.05$ after Benjamini and Hochberg multiple testing correction. Gene ontology (GO) enrichment analysis was conducted using the panther classification system[197] with PBS injected control RNA-seq transcripts used as the background list against which the target list was compared.

Hierarchy diagrams were generated using the Gene Ontology enRIchment anaLysis and visuaLizAtion tool (GOrilla) using a single ranked list of genes[198].

For viral mutation detection, RNA-seq reads were aligned to reference viral genomes using the Burrow-Wheeler Aligner[199] (BWA-0.7.17d). The reference viral genomes were taken from the viral genome starter plasmids used for viral productions. We then used the freeBayes software to detect polymorphisms present within the length of a short RNA-seq read. For quantification of mapped read density, SeqMonk software was used. A running window generator was used to create probes of 30bp length all along the viral genome. Subsequently, read count quantitation was carried out to generate log2 RPM values for each 30bp probe.

2.9 General Analysis

Throughout the thesis, mean values are accompanied by SEM. Statistical analyses and graphical visualisation of data were performed using Prism 8 software (Graphpad). The significance of differences between the means of experimental conditions are depicted as follows: * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$.

Statistical analyses used throughout the theses (unless stated otherwise) are:

- Two-tailed, unpaired Student's T-tests in experiments containing two groups separated by one independent variable
- One-way ANOVA where more than two groups are separated by one independent variable. Tukey correction was performed for multiple comparisons where every group is compared to every other group.

Chapter 3

The Development of Self-Inactivating Rabies Virus 2.0

3.1 Introduction

The computations underlying behaviour are implemented by neuronal circuits. In order to understand how these computations are carried out, one must a) determine the anatomical structure by which the network is connected, b) correlate the concerted activity of neurons within these networks to behaviours or other experimentally tractable aspects of brain function and c) establish causality between circuit activity and brain function.

Until now, much focus has been dedicated to the development of tools that allow us to address challenge a) the anatomical dissection of complex biological network architectures. The importance of developing such tools has been appreciated from the beginnings of modern neuroscience. In 1979, Crick included as one of three suggestions for "invaluable" developments in neuroscience, "a method that would make it possible to inject one neuron with a substance that would then clearly stain all the neurons connected to it, and no others"[200]. In addition, Brain Research through Advancing Innovative Neurotechnologies (BRAIN), one of the largest neuroscience initiatives to date, has as one of its main goals "the aim of deÞning the structural map of all neurons and synaptic connections of entire nervous systems". Serial EM reconstruction is the gold standard of circuit mapping and enables polysynaptic mapping in the absence of prior knowledge of network structure. A full connectomic map of the entire *C. elegans* nervous system was mapped in this way in 1986[201], and with increasing improvements in volume EM imaging and reconstruction hardware and software, larger nervous systems can be traced with this approach. Recently, a complete adult *Drosophila melanogaster* brain was imaged with EM at synaptic resolution[202], and circuit mapping is currently around 30% complete.

However, for larger mammalian brains (e.g. rodents and primates), serial EM reconstruction is not yet a practically feasible method due to the larger volumes that need to be covered to perform synaptic tracing. In such cases, monosynaptic tracing using rabies virus is the only way of labelling the direct presynaptic neurons to a targeted neuronal group in the absence of prior hypotheses within an experimentally tractable time frame[203–207]. In addition, the ability to use the virus in live animals *in vivo* allows not only for network tracing post-mortem, but live network manipulation. This is crucial as anatomical network conÞgurations in isolation cannot explain to us how neural circuits generate behaviour. In order to understand brain function, we also need to observe live network activity and establish behavioural causality in live animals[208,209]. The experimental tractability and circuit speciÞcity of the rabies virus in mammalian behavioural model organisms sets its up nicely to be used not only as a tracing vector but also as a vector to enable circuit-speciÞc recording or

manipulation experiments. However, an inherent challenge of using the rabies virus for such long-term functional studies is its inherent cytotoxicity and immunogenicity[190]. So although it targets circuits very well, it can have detrimental effects on the very circuits that are being studied. Previously in our lab, a self-inactivating rabies (SiR 1.0) was developed whereby the addition of a PEST domain to the C terminus of one of the five viral proteins generated a trans-synaptic circuit-specific virus which could label neurons and then be degraded by the host cell[191]. However, for the short time during which the SiR is active (after initial infection but before viral shut-off) the virus still exerts cytotoxic and immunogenic effects on the host cell. In this chapter, I discuss the development of a new SiR vector, called Self-Inactivating Rabies 2.0 (SiR 2.0) with markedly reduced immunogenicity, higher neurotropism and trans-synaptic transfer efficiency compared to its predecessing rabies vectors. With this, we aim to forward the development of a non-toxic and minimally-immunogenic circuit-specific viral vector that can be used to conduct long term network-to-behaviour experiments. In addition, we expect continued improvements of the rabies virus in such a way could lead to its eventual development as a promising therapeutic vector to target circuit-based neurological disorders.

3.1.1 The historical development of rabies virus as a trans-synaptic tracing tool

The *Rabies lyssavirus* is a member of the *Rhabdoviridae* family and has a single stranded, negative strand RNA genome of approximately 12kb encoding 5 viral proteins: N, P, L, G and M. The viral nucleoprotein (N) encapsulates the RNA genome enabling it to act as a template for transcription and replication by the viral polymerase complex. This polymerase complex consists of the large protein (L) which functions as the polymerase and phosphoprotein (P) which is a non-catalytic subunit. The RNA-N complex together with viral proteins P and L form the ribonucleoprotein (RNP) core of virus particles and forms cytoplasmic active RNP complexes in virus-infected cells[210]. The viral RNP core is bridged to an encapsulating host-derived membrane via the matrix protein (M). Lastly, the rabies glycoprotein (G) is a transmembrane protein that facilitates viral entry into host cells and virus release. Although the receptor for the rabies virus to which the glycoprotein binds is still unknown, the glycoprotein is essential for viral infectivity[211] and trans-synaptic spread[212].

In the wild, the rabies virus is a highly virulent pathogen that is transmitted via the saliva of a rabid animal into a new host through existing scratches in the skin or bites. The virus then enters neurons at peripheral sites and continues a sequence of viral replication and retrograde trans-synaptic spread until it spreads from peripheral organs to the spinal cord and

then to the brain[121]. As such, the rabies virus vector encompasses some qualities which are highly advantageous for circuit neuroscientists: high neurotropism and efficient retrograde trans-synaptic transfer. For widespread use as an experimental tool however, the inherent toxicity and immunogenicity of the virus also needs to be overcome. Thus, there has been a long history of iterative modifications to various rabies virus strains to develop the virus as a safe trans-synaptic tracing tool[190,191,211,213,214].

Initially, viruses isolated from the wild were "fixed": a term used to describe the repeated passaging of viruses isolated from the wild in cell culture or animals resulting in a virus that had a stabilized incubation period and virulence. In 1994, a reverse genetics technique was established that enabled for the first time, the direct production of infectious rabies virus from plasmid cDNA in cell culture. This not only enabled a more consistent method of virus production for experiments, it also enabled the ability to introduce mutations, exchange genes between vaccine and pathogenic strains, or excise certain genes altogether[213]. This reverse genetics technique was first conducted on a genomic sequence of an attenuated rabies virus vaccine strain called SAD-B19. Subsequently, the next notable development was that of the G-deleted rabies virus (ΔG rabies)[211,214]. Here, the viral G protein was deleted from the same SAD-B19 strain, thus modifying it into a mono-synaptically restricted ΔG rabies virus. As the viral genome did not encode G, the virus on its own could no longer spread trans-synaptically. However, if the G was expressed specifically in a starter population *in trans*, the virus would be able to spread to its direct presynaptic partners but no others as no other cells would express the glycoprotein required for further trans-synaptic transfer[203,204]. This monosynaptic restriction was important as the polysynaptic rabies virus was able to jump multiple synapses at different rates, resulting in ambiguity as to the number of synapses crossed in the circuit. In addition, the starting population could be pre-defined to a small subpopulation of neuronal cells by pseudotyping the rabies virus with EnvA (the envelope protein of subgroup A avian sarcoma and leukosis virus (ASLV-A)). The EnvA coat protein binds specifically to its cognate TVA receptor which can be exogenously expressed in starter cells but is not expressed by normal mammalian cells[214]. Thus, the combination of mono-synaptic restriction and EnvA pseudotyping enabled the unambiguous labelling of the direct pre-synaptic inputs into a pre-defined starting population[215].

However, the mono-synaptically restricted WT ΔG SAD-B19 rabies virus, which by most intents and purposes was optimized for tracing experiments, still has limitations for long-term behavioural studies as the virus is cytotoxic and can cause infected neurons to die within a timescale of weeks[191]. As of date, two parallel approaches have been taken in an attempt to reduce the inherent cytotoxicity of this trans-synaptic virus.

In 2016, a ΔG CVS-N2C virus was published where the problem of toxicity was reduced by applying the same monosynaptic restriction and EnvA pseudotyping method as previously described in the SAD-B19 strain to a more virulent strain called CVS-N2C[190]. Increasing virulence of rabies virus strains often results in advantageous characteristics such as increased neurotropism, decreased immunogenicity, and delayed toxicity. A comparison of the WT ΔG CVS-N2C and ΔG SAD-B19 virus revealed that the ΔG CVS-N2C virus showed noticeably delayed toxicity. The survival of infected cells *in vivo* was extended to around 4 weeks for retrogradely infected cortical projection neurons and around 2 weeks for retrogradely infected hippocampal neurons which was longer than for the ΔG SAD-B19 virus[190]. In addition to delayed cytotoxicity, the ΔG CVS-N2C virus exhibited 10- to 20 -fold higher transsynaptic tracing efficiency than the ΔG SAD-B19 virus, enabling a more complete labelling of direct pre-synaptic inputs into starter cells.

Another parallel approach taken by our lab was to abolish, rather than delay, rabies virus-induced cell toxicity by applying a "Self-inactivating Rabies virus" (SiR) technology. Here, a PEST degradation signal was added to the viral nucleoprotein N in the original ΔG SAD-B19 strain. This PEST tag directs the nucleoprotein for degradation by the proteosome, thus reducing its availability to bind to the viral RNA genome which is essential for viral transcription and replication. It was found that this led to the gradual disappearance of the virus from infected cells *in vivo* over approximately 3 weeks and consequently, the once-infected cells were observed to be alive and functionally unimpaired for effectively the whole life-span of the animal[191].

3.1.2 The innate immune response to rabies virus infection

Host organisms have co-evolved with viral pathogens, resulting in a complex interplay between viral and host cell molecular machinery. In general, the host organism needs to quickly identify foreign antigens that are "pathogenic" and trigger appropriate immune responses to the pathogen in question. On the other hand, pathogens have developed many mechanisms to evade recognition by host cells or to misdirect or dampen the immune responses initiated by host cells.

The first step of host antiviral activity is initiated by genetically encoded pattern recognition receptors (PRRs) of the innate immune system. PRRs bind repeating antigenic signatures called pathogen associated molecular patterns (PAMPs) that are representative of different pathogenic challenges. The rabies virus, as a cytosolic ssRNA virus, presents PAMPs such as antigenic viral RNA intermediates that are not usually produced by the host cell as well as for-

eign viral proteins. The PRRs that recognise rabies infection include retinoic acid–inducible protein I (RIG-I) which binds uncapped 5'-triphosphate RNA (3pRNA) presented by viral genomic RNA and Toll-like receptor 3 (TLR-3) which recognizes dsRNA produced during the replication of RNA viruses. They are found to be upregulated in neurons following viral infection *in vitro*[216,217] and by Purkinje cells in the cerebellum in humans following rabies infection[218]. Following activation, these PRRs trigger signaling cascades which lead to IRF3 activation and the transcription of interferon (IFN)-β, as well as IFN-stimulated genes (ISGs). The activation of type-I interferon ((IFN-α/β)) responses establishes antiviral responses in infected cells as well as directing appropriate adaptive immune responses through the release of inflammatory cytokines (Fig. 3.1).

In response to host immune responses, the rabies virus has in turn developed many anti-immune mechanisms (Fig. 3.1). In fact, only minimal inflammation and necrosis is observed in the brains of deceased human rabid patients and experimental mice infected with pathogenic street strains of the rabies virus[220,221]. Conversely, lab-attenuated non-pathogenic rabies strains induce extensive inflammation, cytokine and chemokine expression and apoptosis of infected cells in the CNS[193,220,222,223]. Therefore, it appears that more lethal street strains somehow evade host immune responses to promote long-term survival of the virus and enable their high pathogenicity. In contrast, less pathogenic strains have greatly reduced abilities to cause rabid disease after inoculation into an animal and show a much higher propensity to trigger host immune responses (e.g. the SAD-B19 attenuated vaccine strain). The mechanisms by which this immune evasion is executed needs further study, but appears to be due to a mixture of different factors. For example, street viruses whilst not triggering apoptosis in infected neurons, do trigger apoptosis in inflammatory cells, possibly using this as a mechanism to prevent the release of cytotoxic cytokines following cell lysis[220]. In addition, attenuated strains through the induced activation of CXCL10, IFN-γ and IL17, enhance BBB permeability enabling entry of immune effectors into the main site of viral replication in the CNS whilst pathogenic strains do not[224,225].

Further studies into the differences between pathogenic and non-pathogenic strains have revealed that pathogenic gene variants of the 5 virally encoded proteins increase immune evasion or antagonism. For example, switching of N, P or M genes from a lethal pathogenic strain (Ni) into the genetic background of a non-lethal strain (Ni-CE) results in lethal hybrid strains which show increased immune evasion mechanisms[226]. The lethal N-hybrid strain shows reduced RIG-I activation of the IRF-3 pathway[227] which can be attributed precisely to variations in AA273 and 394[228]. The lethal P-hybrid strain showed decreased sensitivity to host type I IFN responses[226]. Mechanistically, the viral protein P can prevent IRF3 S386

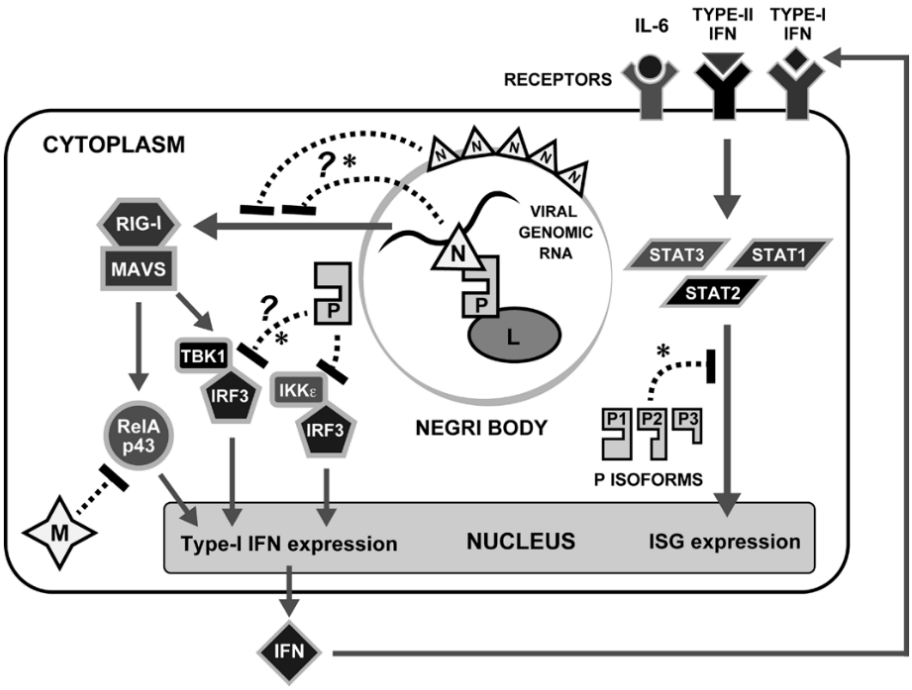


Fig. 3.1 Interferon activation and antagonism by the rabies virus
Rabies virus infection is detected by viral RNA sensor RIG-I which is a helicase with a N-terminal CARD domain that recruits the adaptor Cardif/VISA/MAVS/IPS-1 which acts as a scaffold to recruit signalling components that activate either the IRF3 pathway or NF-kB pathways involving RelAp43. This results in the assembly of an enhanceasome complex on the IFN-b promotor thus activating the expression of type-I IFNs and proinflammatory cytokines. Type I IFNs are released from cells and bind to cell receptors (IFNAR) to activate JAK/STAT signalling pathways that induce the expression of antiviral ISGs. Rabies viral proteins can antagonise this innate viral response at many steps. The N protein can reduce RIG-I activation possibly by inhibiting PAMP-PRR recognition due to its roles in encapsidating the RNA genome and/or forming cage-like structures around virally active Negri bodies. The P protein inhibits the JAK/STAT pathway at multiple steps whilst the M protein inhibits NF-kB pathway activation by inhibiting RelAp43. Host cellular molecules are indicated with white letters, rabies viral proteins are indicated with black letters. Mechanisms directly implicated in pathogenicity are indicated by an asterisk. Figure taken from Ito et al. 2016[219]

phosphorylation[229], inhibit nuclear accumulation of IFN activated STATs[230,231] and inhibit DNA binding of STAT1[232]. The M protein has also been found to target and inhibit RelAp43, a member of the NF-kB pathway[233]. The exchange of G from lethal street viruses into a lab attenuated strain also results in decreased chemokine expression, immune cell infiltration and BBB permeability enhancement[234]. In addition more pathogenic strains show a lower level of G expression and incorporation into virions which contributes to increased immune evasion[235,236].

3.1.3 Aims

Given the clear advantage of pathogenic viral strains for long-term circuit targeting (i.e. increased neurotropism, reduced cytotoxicity and reduced immunogenicity), we aimed to further improve the first-generation SiR vector by applying the same self-inactivation technology to a more virulent CVS-N2C strain. Although the SiR 1.0 modification essentially abolishes the toxicity of the ΔG SAD-B19 virus by shortening its transcriptionally active life-span, the SAD-B19 strain from which it is derived is an attenuated vaccine strain with high immunogenicity and very fast transcriptional kinetics[194]. Therefore, during the short transcriptionally active phase of the SiR 1.0 before viral shut off, the vector is notably immunogenic and can have short-term detrimental effects on labelled circuits. In contrast, the recently developed WT ΔG CVS-N2C shows significantly delayed cytotoxicity, reduced immunogenicity and increased neurotropism compared to the SAD-B19 strain, but still causes cell death as it does not have a switch-off mechanism like the SiR[190].

Therefore, we aimed to combine the inherent reduced immunogenicity and increased neuroinvasiveness of the CVS-N2C strain with the SiR technology to produce a second generation SiR virus (from now on called SiR N2C, or SiR 2.0) that self-inactivates itself very much like the first generation SiR virus (SiR B19, or SiR 1.0) but also enables:

(A) More efficient peripheral targeting of neural circuits

(B) Increased trans-synaptic transfer efficiency

(C) Decreased immune response to SiR infection during the brief period of SiR transcriptional activity after initial infection but before viral silencing.

These developments are particularly important for the long-term targeting of more sensitive neuron types such as spinal motor neurons and with the view of developing the rabies virus as a promising therapeutic vector for circuit-dysfunctions.

3.2 Results

3.2.1 The design and production of SiR 2.0

The SiR 2.0 was designed by the addition of the same TEVs-PEST modification used in the production of SiR 1.0 to the C terminus of the N protein of the ΔG CVS-N2C rabies virus (Fig. 3.2). As mentioned previously, the PEST sequence is a signal peptide that directs the viral N protein for degradation. This leads to the gradual loss of viral proteins necessary for viral translation and replication, leading to self-inactivation of the rabies virus. The PEST signal peptide is attached to the C terminus of the viral nucleoprotein via a linking sequence which includes the Tobacco Etch Virus cleavage site (TEVs). This enables us to cleave off the PEST modification from the N-TEVs-PEST protein during viral production by expressing a highly sequence specific Tobacco Etch Virus Protease (TEVp) within the host cell. By expressing TEVp within production cells, we release the attenuation of the viral translation an replication to facilitate efficient viral recovery from plasmids and high titer viral production *in vitro*.

As the sequence of the 5 genes of the rabies virus are all different in the CVS-N2C strain compared to that of the SAD-B19 strain from which the SiR 1.0 is derived, the production of SiR 2.0 required multiple adjustments. First, we needed to clone the TEVs-PEST modification into the correct location of the CVS-N2C rabies genome plasmid to produce a pSiR N2C plasmid. We also needed to produce 4 plasmids expressing the viral proteins N, P, L and G of the CVS-N2C strain for co-transfection during the recovery of SiR N2C virus from plasmids. We also needed to produce new viral production cell lines expressing TEV protease, GFP and the G protein of the CVS-N2C strain (HEK-TGG). The detailed protocol for SiR 2.0 production is listed in the methods section.

3.2.1.1 SiR 2.0 can be produced at high titer and volume using HEK cells

In order to enable efficient use of SiR 2.0 as an experimental or therapeutic vector, the SiR 2.0 must be able to be produced within a reasonable time frame and be scalable for high titer, high volume productions. The WT ΔG CVS-N2C rabies virus is reported to take 2-3 times longer to produce than the WT ΔG SAD-B19 due to its slower kinetics and production in mouse neuro2A cells as opposed to BHK-21 cells[190]. The SiR B19 also takes approximately two times longer to produce than the WT ΔG SAD-B19 (Table 3.1). As such, producing the SiR N2C in neuro2A cells would in theory take over 2 months to produce from DNA transfection.

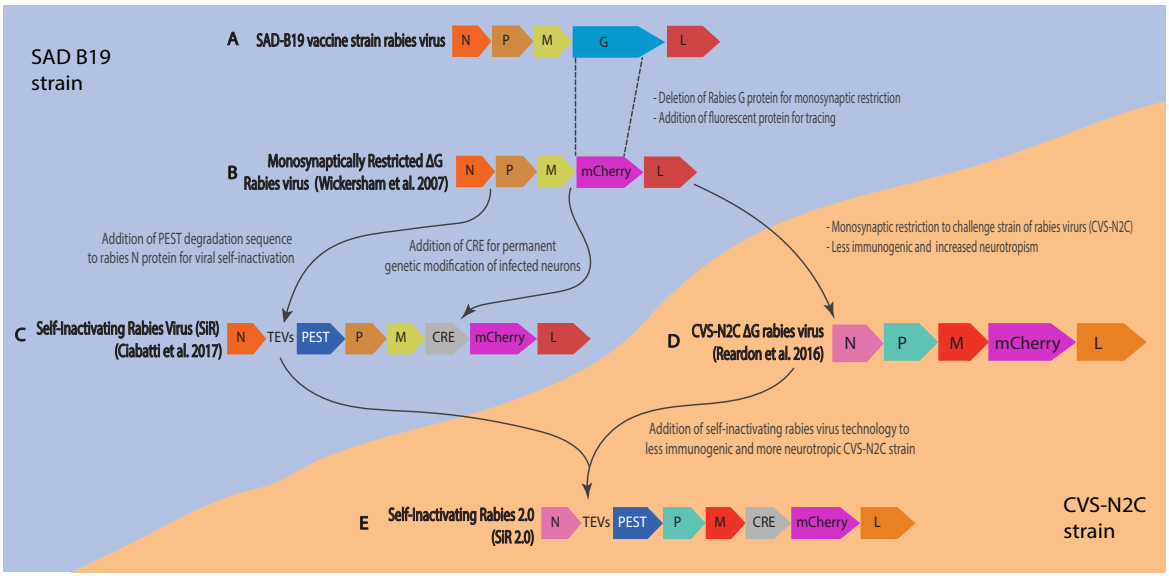


Fig. 3.2 Evolution of SiR 2.0 **(A)** WT SAD-B19 vaccine strain rabies virus which is polysynaptic and highly immunogenic. **(B)** Monosynaptically restricted ΔG SAD-B19 rabies virus. **(C)** SiR 1.0. The monosynaptically restricted SAD-B19 strain is modified further to enable viral self-inactivation, abolishing viral toxicity in host cells. **(D)** Monosynaptically restricted ΔG CVS-N2C rabies virus. This virus originates from a more virulent rabies virus strain and as such displays delayed cytotoxicity, decreased immunogenicity and increased neuroinvasiveness compared to its SAD-B19 counterpart. **(E)** The SiR 2.0. This virus effectively abolishes the delayed cytoxicity of the CVS-N2C strain whilst incorporating the decreased immunogenicity and increased neuroinvasiveness of the CVS-N2C strain for monosynaptic circuit-targeting.

	Cell type used	Length of viral production (weeks)
WT B19	HEK 293T / BHK-21	2-3[189]
WT N2C	Neuro2A	6-8[190]
WT N2C	HEK 293T	2-3
SiR B19	HEK 293T	5-6[191]
SiR N2C	HEK 293T	5-6

Table 3.1 Comparison of production times of WT and SiR viruses of B19 and N2C strains

The slow metabolism of neuro2A cells necessitates long viral supernatant collection periods of up to 28 days for the WT ΔG N2C rabies. The rationale of using murine neuro2A cells in the original WT ΔG N2C production protocol was explained to be due to the high likelihood that the continued maintenance of the virus exclusively in murine neural cells would be important to maintain the high neurotropism of the CVS-N2C strain. To support this hypothesis, it has previously been shown that the passaging of the highly neurotropic CVS-24 strain (from which the CVS-N2C is derived) through BHK-21 cells can select for substitutions in the viral G protein that significantly reduce viral neurotropism[237]. However, in the original paper, the less neurotropic CVS-24 subpopulation that is selected for in BHK-21 cells is referred to as a conserved minor subpopulation already existent in the CVS-24 preparation. In our case, viral productions begin from cDNA and so the starting viral population is homogeneous. In addition, variations in the G protein are recognised as the main driver for the selection of different subpopulations[237]. In our case, as the G protein is expressed from cells and is not encoded in the rabies genome, we did not expect to see a significant reduction in viral neurotropism by using a different cell type. We therefore attempted to produce the SiR N2C using HEK293T cells which have a faster metabolism. However, we needed to verify that the WT and SiR N2C viruses produced in this way still exhibited high neurotropism. In addition, we also wanted to ensure the genetic stability of SiR 2.0 viruses produced from HEK 293T was not compromised as it has been reported that different host cell environments can alter RNA virus mutation rates[238].

We find that we can produce both ΔG WT N2C virus and SiR N2C viruses at high titers using HEK 293T cells in a much shorter time scale then previously reported using neuro2A cells. The WT N2C viruses produced in HEK 293T cells within 2-3 weeks show titers similar to those published previously using neuro2A cells using the same titration protocol (Fig. 3.6C). We also find that SiR 2.0 viral particles produced using this method are genetically stable (Fig. A.2) and that their high neurotropism is also retained *in vivo* enabling efficient central (Fig. 3.8) and peripheral (Fig. 3.9) targeting of neural networks.

3.2.2 Genomic stability of SiR viral particles *in vitro*

All RNA viruses, including rabies viruses, are known to be subject to a high rate of mutations[238,239]. As such, we systematically investigated the genomic stability of SiR. Most importantly, we wanted to investigate the likelihood of the virus to obtain mutations in the crucial C-terminal TEVs-PEST modification of the SiR which could result in the loss of

self-inactivating behaviour. Such WT reverting mutations had been reported recently in two SiR productions[192].

3.2.2.1 The crucial C-terminal TEVs-PEST modification is not selectively mutated during SiR production

First, we sequenced 50 individual SiR viral particles each from 10 independent productions of SiR viruses to investigate the accumulation of mutations in SiR during viral production in cell culture. A barcoding RT-PCR protocol was used to prepare genomic libraries for sanger sequencing of individual viral particles as described in Matsuyama et al. 2019[192] (Fig. 3.3A). This molecular method enabled the unique indexing of each single starting genomic RNA strand by the incorporation of a random octamer barcode in the reverse-transcriptase (RT) primer binding to the 3' leader sequence of the viral genome. In the total of 494 individual viral particles sequenced, we did not observe a single SiR viral particle that had obtained revertant mutations (Appendix A) demonstrating that the occurrence of revertant mutations is rare.

We did observe mutations elsewhere in the rabies virus genome. This was in line with the previous literature regarding the high mutation rates of RNA viruses and confirmed the validity of our sequencing method. A comparison of mutation frequency in different genome regions showed also that there was no positive selection or "hot spot" of mutation accumulation at the attenuating TEVs-PEST coding sequence of SiR viruses (Fig. 3.3D). Rather, we identified a polyA stretch in the intergenic region between the N and P coding sequences which showed a statistically increased mutation rate for insertion mutations.

In summary, these results show that revertant mutations are rare and do not accumulate when the virus is produced in appropriate conditions in cell culture as previously described[191].

3.2.2.2 SiR 2.0 shows decreased mutagenicity compared to SiR 1.0 during viral production

As SiR 1.0 and 2.0 are derived from different rabies viral strains with different replication kinetics and toxicity[190], we suspected that they may also have inherently different mutagenic rates. As we were producing these two SiR variants in the same host cell conditions, we deduced that a comparison of their mutation rates would reflect inherent virus-derived differences in their propensity for mutation accumulation. Analysis revealed that the SiR 2.0 had a significantly lower inherent mutation rate (Fig. 3.3C). This lower mutagenicity may in turn be related to the WT N2C strain's decreased neurotoxicity and slower replication

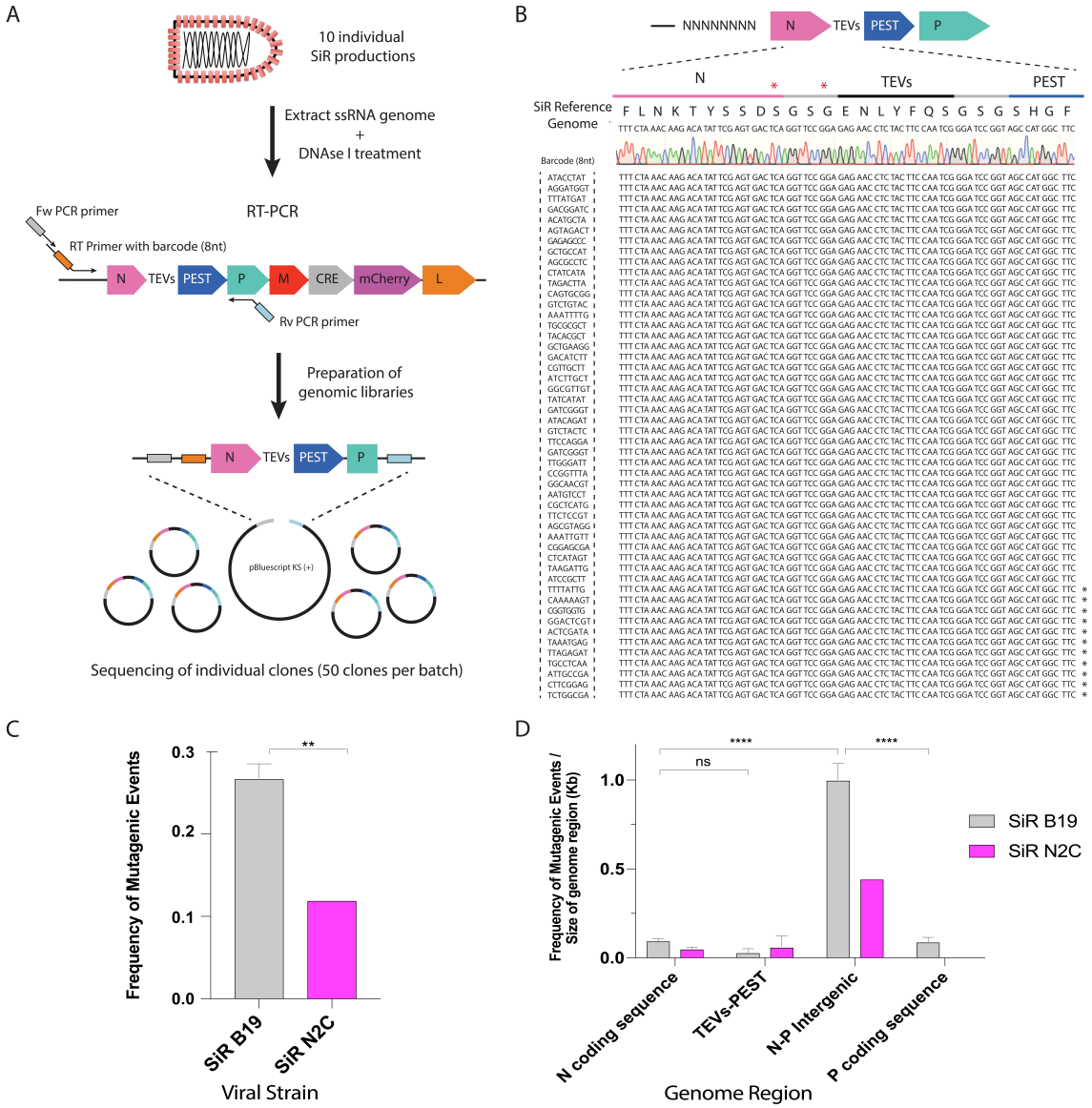


Fig. 3.3 SiR Viral particles do not selectively accumulate nonsense mutations at the TEVs-PEST site

(A) Scheme of procedure to sequence individual viral particles from 10 independent SiR preparations: ssRNA genomes from a viral production are extracted and treated with DNAse I. The RNA is retrotranscribed and the entire N-TEVs-PEST sequence and portion of the P gene is PCR amplified. The amplicon is then inserted into a pBluescriptKS(+) sequencing vector to create genomic libraries. 10 independent productions were prepared in this way and approximately 50 individual clones are sequenced by the Sanger method for each production, giving a total of 500 individual clones. **(B)** Example of one viral production showing no mutations in the region upstream of the PEST domain (clones with mutations in other regions of the analysed amplicon are marked with a star). The position of the nonsense mutations reported in Matsuyama et al. are marked with red stars above the reference sequence. The comprehensive analysis of mutations for all batches is in Appendix A. **(C)** SiR N2C productions show a lower inherent rate of mutagenic variation compared to SiR B19 productions. Two-tailed Student's T-tests. (D) A comparison of mutagenic events in different genomic regions reveals that the TEVs-PEST region of SiR viruses is not a hot-spot for mutation accumulation. Rather, a polyA stretch in the N-P intergenic region shows increased propensity for mutation accumulation compared to other regions of the genome. One-way ANOVA with Tukey correction for multiple comparisons.

kinetics[240,241]. In the context of high volume and titer production of SiRs in cell culture, this is an advantageous quality of SiR 2.0 as it is an inherently more stable virus and reduces the mutagenic variation of RNA virus populations used for *in vivo* experiments.

3.2.3 Genomic Stability of SiR viral particles *in vivo*

In addition to looking at the stability of rabies virus genomes during production *in vitro*, we also investigated the mutagenicity of the rabies viruses *in vivo* following infection of neurons in the central nervous system. This was to investigate a number of different things.

First, to investigate the robustness of the SiR N terminal TEVs-PEST addition *in vivo*. The *in vivo* system is very different to the non-neuronal *in vitro* system from which the viruses are produced, and this is likely to impact viral replication biology. Most importantly, unlike in cell culture, infected neurons do not express the TEV protease enzyme. The TEV protease is required to cleave and consequently "switch off" the PEST domain induced viral attenuation during high titer viral production *in vitro*. As no TEV protease is expressed in neurons *in vivo*, if a mutant WT viral particle were to arise it would have a significant selective advantage over the SiR, leading to amplification of the revertant viral population. Thus, we wanted to directly experimentally verify the stability of the TEVs-PEST modification in SiR viral particles *in vivo* both in primarily infected neurons, as well as trans-synaptically labelled neurons.

Second, we also compared the differing mutagenicities of the WT B19 and N2C strains *in vivo* to further verify our findings that the N2C strain showed decreased mutagenicity from the barcoding Sanger sequencing method *in vitro*.

3.2.3.1 Next generation sequencing of SiR N2C RNA extracted from primarily infected neurons *in vivo* shows absence of mutation accumulation in TEVs-PEST modification

In order to investigate viral mutagenicity *in vivo*, we injected the CA1 of adult mice with PBS or SiR N2C virus and sacrificed 1 week post injection (p.i.) for RNA extraction from the hippocampus. Ribosomal RNA was extracted from the preparation and the remaining total RNA was prepared for next generation sequencing (NGS) using a single-indexed library. Following sequence acquisition, the sequencing reads from each injected brain were aligned to the SiR N2C reference viral genome using Burrows-Wheeler Aligner software[199] (Fig. 3.4A). Successful alignment was then confirmed by visualization of sequence alignments using the Seqmonk program (https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/).

Visualization of read alignments from SiR N2C injected brains onto the SiR N2C viral genome showed, as expected,very high numbers of successfully mapped reads. In contrast, there were a very small number of chance alignments from the PBS injected samples (Fig. 3.4B).

Following successful alignment, we first investigated whether there were changes in the aligned read density over the TEVs-PEST coding sequence compared to other regions of the viral genome. This was to assess for possible deletions of the TEVs-PEST coding sequence from the viral genome. Visual observation and quantification of read density did not reveal any appreciable difference between the number of reads aligned to this part of the genome compared to other coding regions (Fig. 3.4B, D). However, a decrease in read density was observable in the N/P intergenic region as would be expected due to the lack of mRNA transcription from intergenic sequences (Fig. 3.4C, D).

Subsequently, the sequencing reads that were aligned to the viral genome were screened for small polymorphisms using a genetic variant detector called freeBayes (version 1.3.0)[242]. FreeBayes uses bayesian inference to find small polymorphisms that are smaller than the length of a short-read sequencing alignment. Specifically, it can detect: SNPs (single-nucleotide polymorphisms), indels (insertions and deletions), MNPs (multi-nucleotide polymorphisms), and complex events (composite insertion and substitution events).

First, to verify this pipeline for the detection of viral mutations, we conducted mutation analysis for virally aligned reads against a rabies reference genome sequence that had purposefully introduced SNPs, indels, MNPs and complex mutations at known locations. As expected, we found that BWA alignment followed by freeBayes polymorphism detection very robustly identified these apparent "mutations" in the virally aligned sequence reads (Table 3.2). All introduced mutations were detected in 3 out of 3 brain samples with a high read depth of >2,000.

Having validated our method, we then screened the virally aligned reads from SiR N2C infected hippocampal samples against the unedited, original SiR N2C genome to identify mutations that would have accumulated during the total time from high titer viral production *in vitro* to 1 week post-injection *in vivo*. From the 3 samples obtained from 3 injected mice, we identified a total of 89 mutagenic events with varying read depths across the viral genome (Fig. 3.4E). As expected, the frequency of detection of these spontaneous mutations was much lower than that identified in the pipeline verification (3-200 vs. 2,000-40,000). The individual positions and nature of mutations identified are detailed in table B.1 in Appendix B. Notably, none of the 89 identified mutagenic events identified from the 3 injected samples fell into the TEVs-PEST coding region (Fig. 3.4F). Taken together, this suggests that the

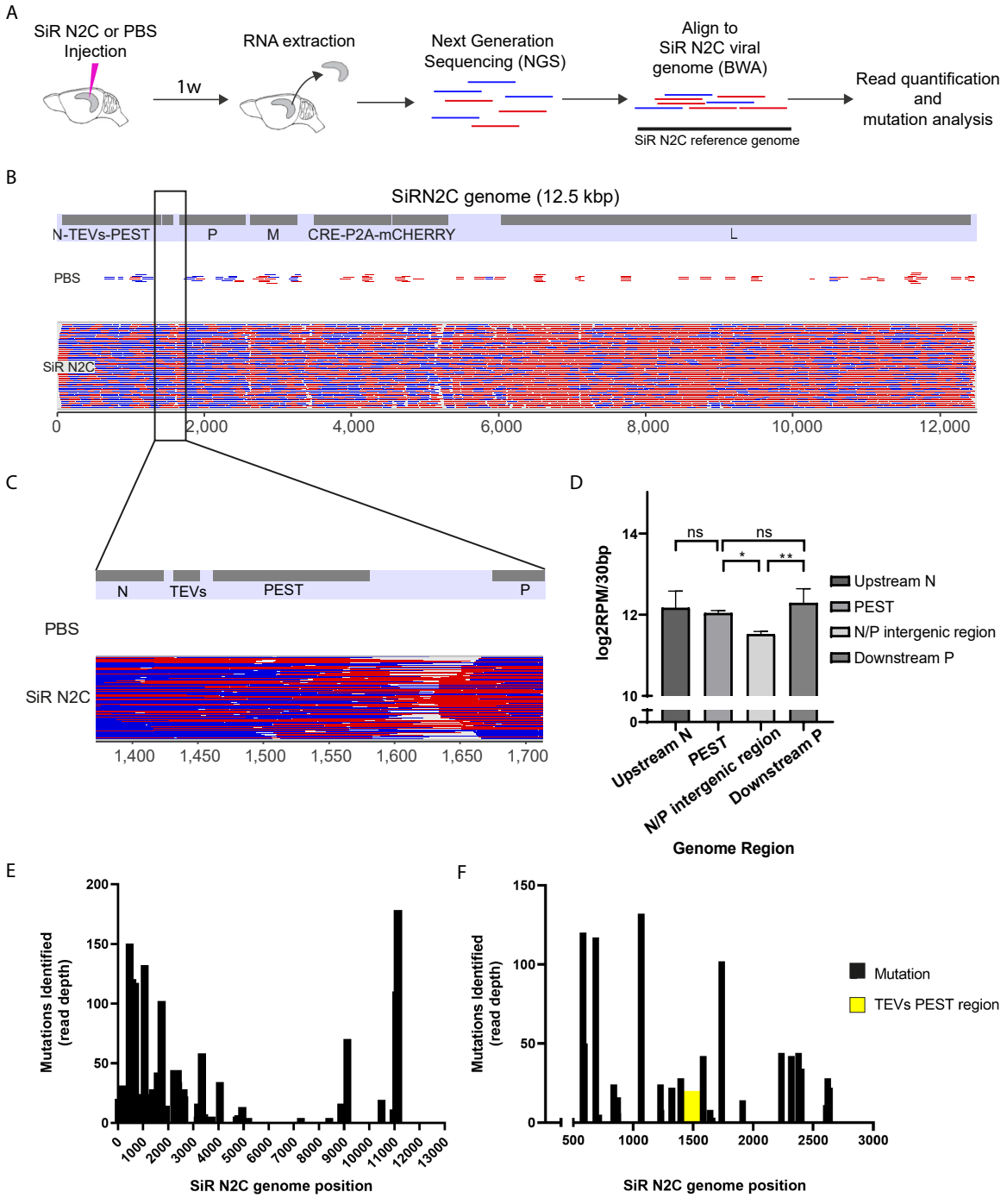


Fig. 3.4 RNA-sequencing of SiR N2C infected hippocampi shows stability of TEVs PEST modification *in vivo* (A) The CA1 of mice were injected with SiR N2C virus or PBS control. 1 week p.i. the hippocampi were dissected for total RNA-seq sample preparation. Sequenced reads were aligned to the SiR N2C viral genome using Burrows-Wheeler Aligner[199] (BWA). Aligned reads were then quantified and screened for mutations using freebayes variant detector[242]. **(B)** Visualisation of high density of reads from SiR N2C injected samples aligned to the SiR N2C genome using Seqmonk software. Samples from PBS injected samples show only very few chance alignments in comparison. **(C)** Zoom in to TEVs-PEST region of the genome to show high read density at N, TEVs-PEST and P coding sequences, but reduced read density at the N/P intergenic region. **(D)** Quantification of read density to viral genome regions. N, PEST and P coding regions show no significant differences in read density (n=3, one-way ANOVA with Tukey correction for multiple comparisons). **(E)** Visualisation of mutagenic events identified across SiR N2C genome. (**F**) Zoom in to the TEVs PEST coding region (yellow) shows zero mutations identified in this region from a total of 3 injected samples.

Introduced Mutation			Freebayes Detection		
Position	Reference	Mutation	Mutation Type	Frequency (out of 3 samples)	Total Read Depth
1066	G	A	snp	3	38422
9119	G	A	snp	3	3112
9338	T	C	snp	3	2887
9497	C	T	snp	3	2627
10533	G	A	snp	3	2137
10708	GT	AC	mnp	3	2162
10714	T	C	snp	3	2259
11801	TCCACCG	CCCCCCA	complex	3	2327

Table 3.2 Verification of mutation detection pipeline from next generation sequencing data

SiR N2C virus does not particularly accumulate mutations on the criticial TEVs-PEST modification either during production *in vitro* or following 1 week of infection of neurons *in vivo*.

3.2.3.2 NGS reveals reduced mutagenicity of N2C strain compared to the B19 strain *in vivo*

We also applied the same mutation analysis pipeline to NGS reads obtained from hippocampal samples infected with WT ΔG B19 and N2C viruses produced from HEK cells as described in the methods section. In line with our findings from Sanger sequencing of viral particles *in vitro*, we found that the B19 strain showed a higher accumulation of mutations then the N2C strain. We identified 32 different polymorphic events in WT B19 injected brains as opposed to only 2 in WT N2C injected brains (Table 3.3 and 3.4). In line with our findings from *in vitro* Sanger sequencing of individual viral particles, there was a predilection for insertion mutations to accumulate in polyN sequences in the viral genome (specifically polyA and polyT sequences in both intergenic and coding regions). Again this supported the inherently increased genome stability of the N2C strain, and thus its increased suitability for repeatable production of more genomically homogenous populations of virus for experimental and therapeutic use.

3.2.3.3 SiR retains the TEVs-PEST domain *in vivo* upon transsynaptic transfer

In addition to characterizing SiR genome stability *in vitro* and *in vivo* through RNA sequencing, we also examined the stability of the TEVs-PEST site at the protein level *in vivo*. We

	Sequence Type	Position	Reference	Mutation	Mutation Type	Frequency (out of 3 samples)	Total Read Depth
Leader Sequence	leader		no mutations identified				
N	gene	736	A	G	snp	1	5
		757	T	C	snp	1	5
		763	T	C	snp	1	5
		778	T	C	snp	1	3
		793	A	G	snp	1	3
		799	T	A	snp	1	3
		811	A	G	snp	1	3
		1036	C	T	snp	1	7
N/P	intergenic	1475	GAAAAAAACTAA	GAAAAAAAACTAA	ins	1	11
P	gene		no mutations identified				
P/M	intergenic		no mutations identified				
M	gene		no mutations identified				
M/iCREGFP	intergenic		no mutations identified				
iCREGFP	gene		no mutations identified				
iCREGFP /L	intergenic	5402	GAAAAAAACATT	GAAAAAAAACATT	ins	1	5
L	gene	6574	T	C	snp	1	7
		6636	TAAAAAAATGAT	TAAAAAAAATGAT	ins	1	201
		8567	CTTTTTTGGGA	CTTTTTTTGGGA	ins	1	161
		9073	A	G	snp	1	8
		10315	G	A	snp	1	6
		10605	C	T	snp	1	4
		10624	GCC	TCT	complex	1	5
		10644	TCTTTTA	CCTATTG	complex	1	6
		10662	C	T	snp	1	6
		10710	G	T	snp	1	3
		10717	A	G	snp	1	3
		10722	TATCG	CATCA	complex	1	3
		10743	TC	CT	complex	1	3
		11787	G	A	snp	1	3
		11799	A	C	snp	1	3
		11809	A	C	snp	1	3
		11832	C	T	snp	1	4
Trailer Sequence	trailer	11867	C	T	snp	1	6
		11880	GAC	AAT	complex	1	6
		11889	G	A	snp	1	6
		11895	T	C	snp	1	6
		11916	CTA	TTG	complex	1	4

Table 3.3 WT ΔG B19 mutation accumulation post-viral production and 1 week injection *in vivo*

	Sequence Type	Position	Reference	Mutation	Mutation Type	Frequency (out of 3 samples)	Total Read Depth
Leader Sequence	leader		no mutation identified				
N	gene		no mutation identified				
N/P	intergenic		no mutation identified				
P	gene		no mutation identified				
P/M	intergenic		no mutation identified				
M	gene		no mutation identified				
M/ iCREmCHERRY	intergenic	3192	CAAAAAAACACA	CAAAAAAAACACA	ins	1	49
iCREmCHERRY/ L			no mutation identified				
L	gene	7295	A	G	snp	1	116
Trailer Sequence	trailer		no mutation identified				

Table 3.4 WT ΔG N2C mutation accumulation post-viral production and 1 week injection *in vivo*

did this by infecting neurons *in vivo* with the SiR virus, and then testing for TEVs-PEST expression in infected neurons by using an antibody that binds specifically to the TEVs linker protein.

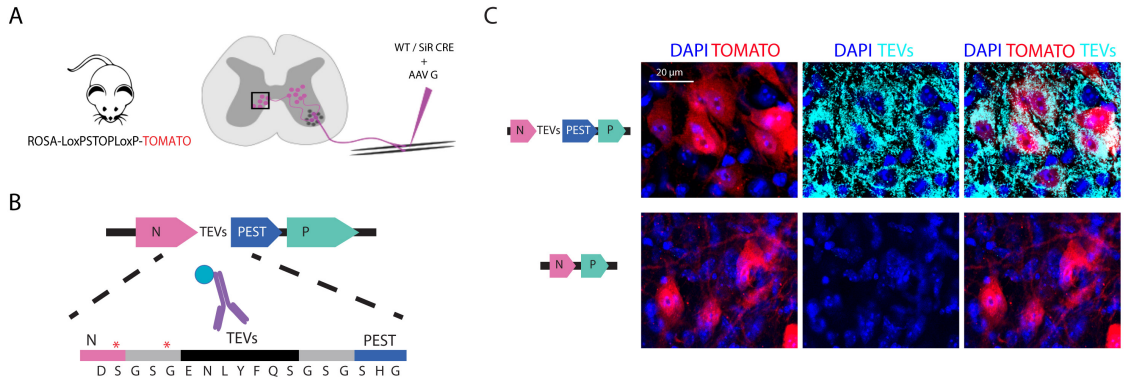


Fig. 3.5 SiR Viral particles do not selectively accumulate nonsense mutations at the TEVs-PEST site
(A) Intramsucular injection of p3 pups with an SiR or WT N2C rabies virus expressing CRE and an AAV expressing G required for pre-synaptic transfer of viral particle results in trans-synaptic tracing of premotor neurons from primarily infected motor neurons. **(B)** TEVs specific antibody recognizes the expression of the TEVs domain added to the C terminus of N in SiR viruses. **(C)** Premotor neurons trans-synaptically infected by SiR (top panel) but not WT G Rabies (bottom panel) show TEVs staining.

The use of antibody immunofluorescence also enabled us to spatially distinguish the stability of SiR genomes in both primarily infected neurons, and spatially distinct trans-synaptically infected secondary neurons. Using this method, we aimed to directly experimentally investigate the recent suggestion that non-revertant SiR virus particles could not propagate trans-synaptically *in vivo*[192]. The assumption here was that the replicated viral particles *in vivo* enabling trans-synaptic viral transfer must be WT revertants, as the SiR virus should be unable to replicate *in vivo* in the absence of TEV protease. However, published findings showing that SiR transcription attenuation takes a period of 2-3 weeks to fully inactivate the virus following *in vivo* injection suggested to us that revertant mutant accumulation may not necessarily be the only explanation for the ability of SiR to jump trans-synaptically *in vivo*.

In order to directly investigate this hypothesis, we first trans-synaptically traced premotor neurons in the mouse spinal cord by co-infecting spinal motor neurons with SiR-CRE and AAV-G by intramuscular injection (Fig. 3.5A). We then visualized whether trans-synaptically propagated SiR viral particles in premotor neurons had developed nonsense mutations at the PEST site by staining pre-synaptic neurons with an antibody that recognizes the TEVs domain downstream to the reported nonsense mutations (Fig. 3.5B). SiR-traced premotor neurons (but not WT-traced premotor neurons) were positively stained, indicating the absence of premature stops and the presence of an intact TEVs-PEST tail at the C-terminal of the N protein of trans-synaptically spread SiR viral particles (Fig. 3.5C). As such, these results show that SiR particles with intact TEVs-PEST modifications can propagate trans-synaptically *in vivo* and supports the genomic stability of SiR *in vivo*.

3.2.4 Development of alternative titration methods for SiR viruses

The established method of rabies virus titration is by fluorescence activated cell sorting (FACS) of virally expressed fluorescent proteins in human HEK 293T cells infected with serially diluted virus[190,191,243]. However, we found that for the case of SiR 2.0, this method was not suitable due to its lower tropism for non-neural HEK 293T cells, low transcriptional levels and self-inactivating nature, all of which significantly decreased the fluorescence levels of infected HEK 293T cells. Without additional staining, we were in fact unable to visualise SiR 2.0 driven fluorescence in HEK 293 T cells at all (3.6D). Compared to the very low apparent titers *in vitro*, we observed high numbers of virally infected neurons *in vivo* for the same viral productions of SiR 2.0 highlighting the need for alternative titration methods.

We proposed that these apparent differences in titer could be due to several possible reasons. First, the N2C strain has lower transcriptional levels then the B19 strain mak-

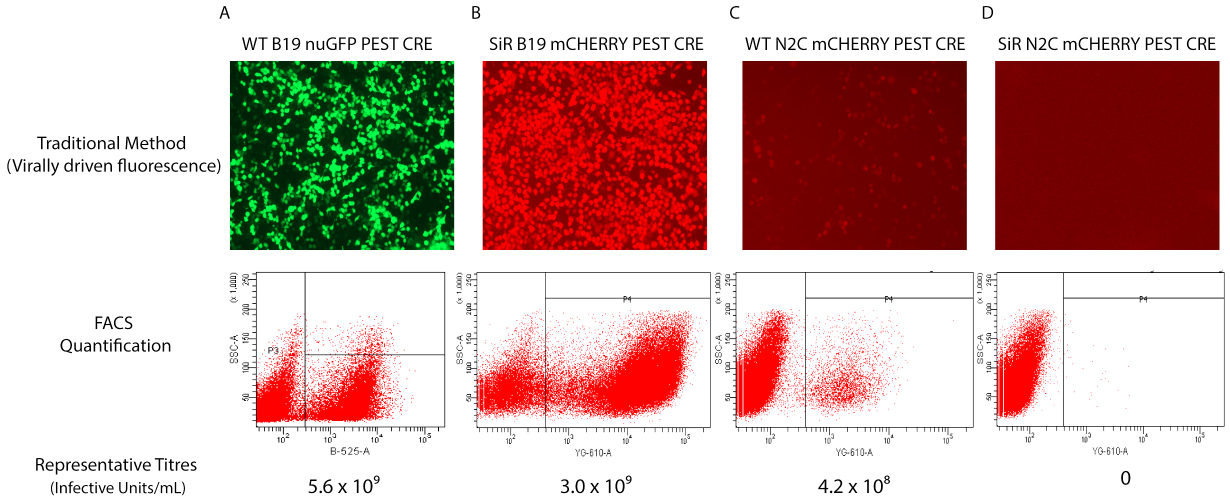


Fig. 3.6 Traditional viral titration by quantifying virally driven fluorescence signals Demonstration of the titration of **(A)** WT B19, **(B)** SiR B19, **(C)** WT N2C, **(D)** SiR N2C viruses by quantifying virally driven fluorescence of infected HEK 293 T cells. Light microscopy images of highly infected cells (first row) show that the viral expression of fluorescent proteins is lower in SiR viruses then in WT viruses (C vs. D), and also lower in the N2C strain compared to the B19 strain (B vs. D). This results in a reduced efficiency (C) or lack (D) of detection of virally infected cells by FACS quantification (second row) and thus unrepresentative titer calculations (bottom row). FACS quantification dot-plots represent each detected cell as a single dot. The axes are logarithmic and represent the level of side scatter (SSC) on the y-axes and the level of GFP or RFP fluorescence intensity on the X-axes (using B-525-A or YG-610-A filters respectively).

ing infected cells harder to visualise and so leading to a decrease in apparent titers using fluorescent markers (compare Fig. 3.6B and D). Second, the addition of the TEVs-PEST degradation domain to the viral N protein in SiR viruses reduces transcriptional efficiency and thus reduces fluorescent protein expression levels further (Fig. 3.6C and D). This shifts the apparent SiR titers to lower levels than for WT viruses. Third, the CVS-N2C strain is known to have higher neurotropism than the SAD-B19 strain. As such, the N2C may be less effective in infecting non-neural human HEK 293T cells than the less neurotropic B19 strain, leading to decreased cell infectivity and thus decreased apparent titers.

In order to fairly compare these different viral strains, we had to develop new methods of viral titration that would enable us to accurately quantify and compare the amount of virus present in preparations spanning WT vs SiR viruses, and B19 vs N2C strains. Initially we produced a HEK 293T stable cell line expressing constitutive BFP and floxed mCHERRY thus enabling the detection of CRE activity driven by the different viruses. We hypothesized that this would give a more binary read out of viral quantity and that detection would not be as dependent on transcriptional levels of the virus. This development confirmed that there was indeed a much higher titer of SiR CRE viruses than had been measured using the traditional method (Fig. 3.7). This confirmed that weaker fluorescent levels of SiR viruses had previously led to under-estimations of SiR titers compared to WT viruses.

However, we were still unable to quantify SiR 2.0 viral particles using this CRE reporter method suggesting that the increased neurotropism of the N2C virus was decreasing the efficiency of HEK 293T cell infection *in vitro*. This discrepancy was also observed in a more neurogenic neuro2A cell line (N2A). Therefore, we decided to develop a more direct molecular method of viral quantification. Following the direct lysis of viral productions using a solution of IGEPAL CA-630+BSA or following RNA purification of viral productions using commercial kits, fold changes of viral RNA between productions was quantified by RT qPCR using Taqman probes for sequences present in both SiR 1.0 and SiR 2.0 viruses. This revealed that compared to the apparent "zero" titer of SiR 2.0 preps observed with other methods, SiR 2.0 productions were only around 3 fold lower in titer than SiR 1.0 productions (Fig. 3.7). In addition, the relative titers of WT N2C viruses also increased significantly. Typically, cell based titration methods quantify WT N2C preparations as being 100 fold lower than WT B19 preparations (Reardon et al. 2016[190], Fig. 3.6). Using RT-qPCR methods, we found that the fold difference between WT B19 and N2C viruses was much smaller and more on the order of 2-5 (Fig. 3.7). Such information was crucial to inform our further *in vivo* work.

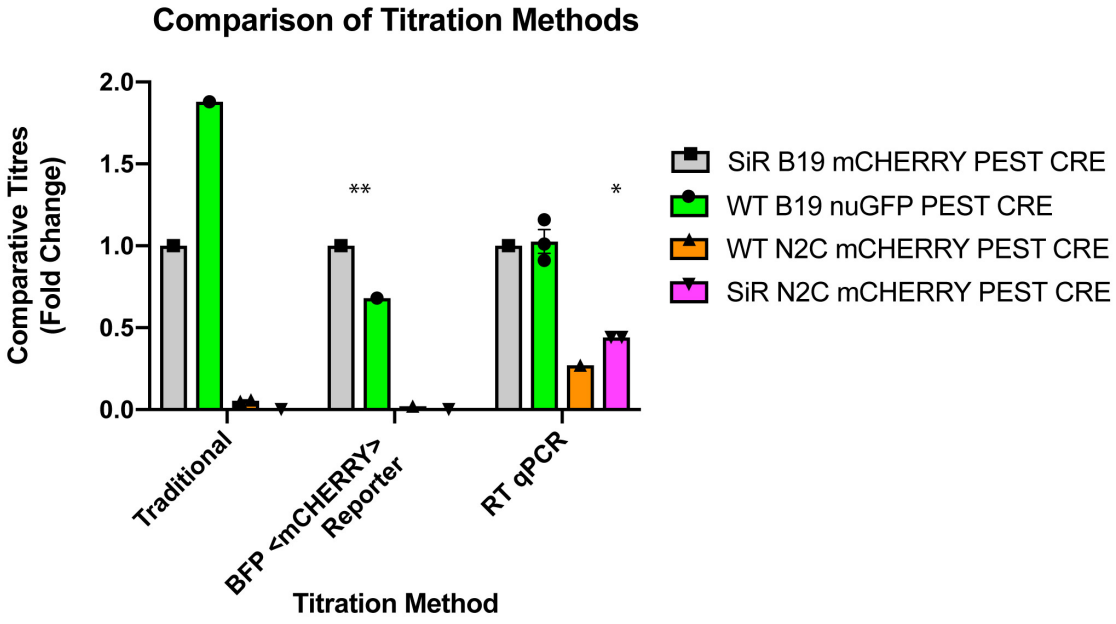


Fig. 3.7 Comparison of new titration methods 4 productions of WT B19, SiR B19, WT N2C and SiR N2c viruses were titered using three methods. Results show that only the RTqPCR method was able to detect SiR N2C viral particles for quantification (*). The use of a CRE-reporter cell line, rather then traditionally used virally driven fluorescence, to quantify CRE-expressing SiR and WT B19 viruses resulted in a higher proportional increase in the titers of CRE-expressing SiR B19 viruses compared to WT B19 viruses (**)

3.2.5 Confirmation of SiR 2.0 self-inactivation *in vivo*

Following high titer productions of SiR 2.0, we injected the virus both intracerebrally and intramuscularly to confirm that the SiR 2.0 could infect neuronal populations *in vivo* and exhibit the properties of viral switch-off. This would enable long-term survival of infected neurons as would have been predicted by the effects of the C terminal TEVs-PEST design.

3.2.5.1 SiR 2.0 self-inactivation kinetics are slower then for SiR 1.0

We monitored viral self-inactivation *in vivo* by injecting SiR 2.0 expressing mCHERRY fluorescent protein and CRE recombinase into a YFP CRE reporter mouse. In this way all virally infected cells were permanently labelled with a YFP fluorescent marker, whilst active viral transcription could be monitored by mCHERRY fluorescence. We found that following injection of the SiR 2.0 (but not the WT N2C control virus expressing the same transcriptional cassette) there was a gradual increase in the proportion of mCHERRY-/YFP+ neurons, representing increasing viral self-inactivation over time (Fig. 3.8B). Interestingly, comparison of the self-inactivation kinetics of the new SiR 2.0 with SiR 1.0 revealed that viral self-inactivation was slower in the SiR 2.0, most probably due to different viral transcription and replication mechanisms and kinetics in the N2C strain.

3.2.5.2 SiR 2.0 shows no observable cytotoxicity in hippocampal and motor neurons

To confirm long-term survival of virally infected cells, we monitored the total number of virally infected neurons over time *in vivo*. As we knew that the hippocampal and spinal motor circuits are especially susceptible to viral cytotoxicity, we longitudinally monitored virally infected cells in both these neuronal populations. We found no significant decrease of YFP+ virally infected cells for up to two months in the hippocampus following intracerebral stereotaxic injection of virus into CA1 of YFP CRE reporter mice (Fig. 3.8C). We also monitored the number of SiR 2.0 infected neurons in the spinal motor-premotor circuit over time and again found that there was no decrease in the total number of SiR 2.0 CRE infected neurons following intramuscular injections into pups of a TOMATO CRE reporter line. This was in contrast to the SiR B19 and WT N2C virus which did exhibit cytotoxicity and loss of virally infected neurons over a period of 1-4 weeks (Fig. 3.9D). This is the first assessment of motor and premotor neuron survival following intramuscular SiR infection *in vivo* and presents the SiR 2.0 as the preferred vector of choice for long term targeting of spinal motor circuits.

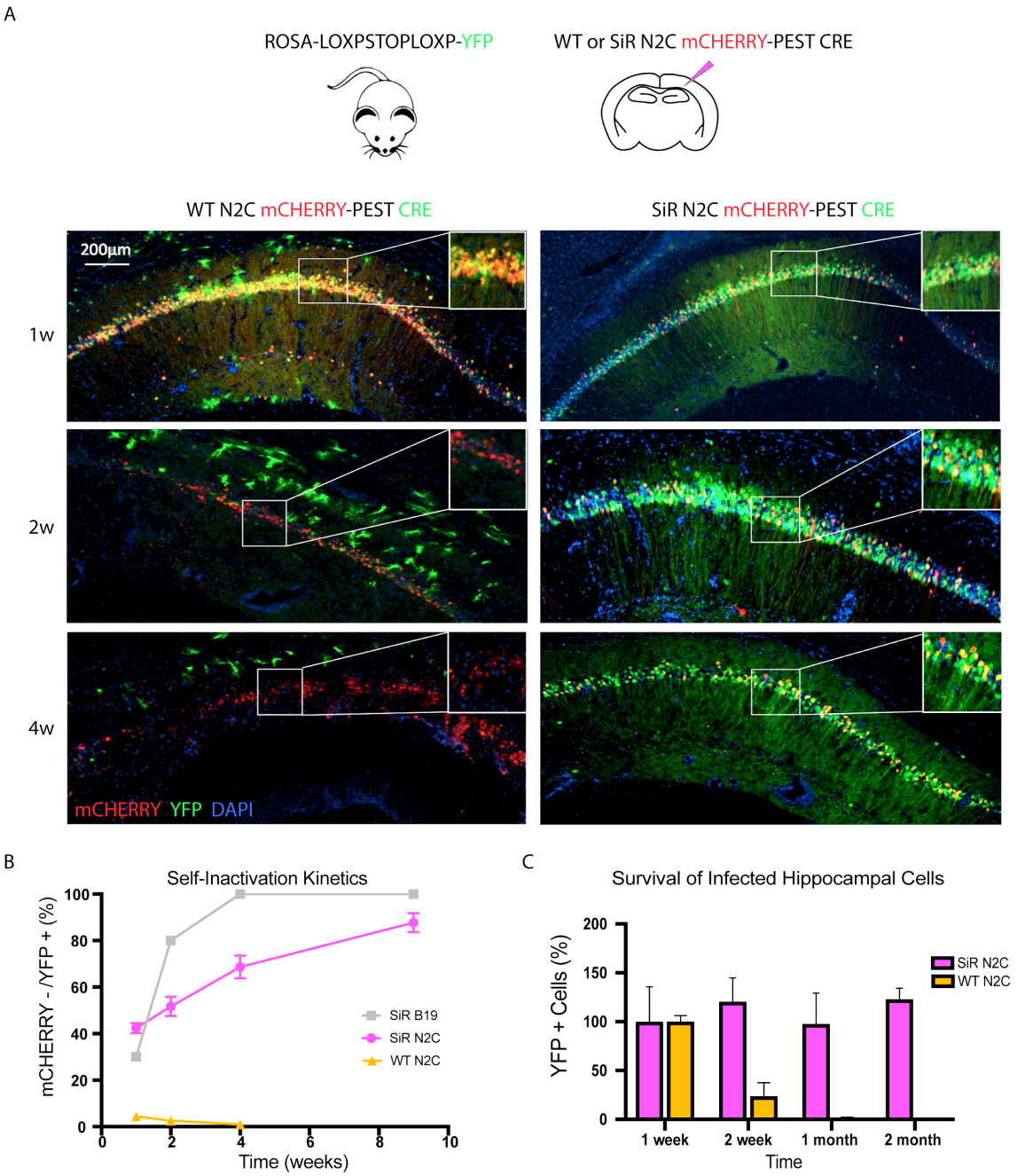


Fig. 3.8 SiR 2.0 self-inactivation kinetics and reduced cytotoxicity in CA1 (A) WT or SiR N2C rabies viruses expressing a fluorescent mCHERRY protein and CRE recombinase were injected into the CA1 of YFP CRE reporter mice, followed by longitudinal monitoring of infected cell numbers and fluorescence. **(B)** Proportion of mCHERRY-/YFP+ neurons at 1week, 2 weeks, 1 month and 2 months p.i. The gradual increase in the proportion of mCHERRY-/YFP+ cells over time following injection with SiR viruses represents viral self-inactivation. No such self-inactivation is observed with the WT N2C virus (mean ± SEM, n = 3 animals per time point). **(C)** Number of YFP+ neurons at 1week, 2 weeks, 1 month and 2 months p.i. normalized to 1 week time point (mean ± SEM, n = 3 animals per time point).

3.2.5.3 Addition of CRE is not sufficient to reduce toxicity of WT virus

It was previously suggested that an alternative mechanism by which the SiR CRE viruses exhibited decreased cytotoxicity could be due to the expression of a CRE recombinase which could reduce viral cytotoxicity[192]. Here, we find that a comparison of WT N2C and SiR N2C viruses expressing exactly the same cassette including a CRE recombinase show clear differences in self-inactivation (Fig. 3.8B) and survival of infected neurons (Fig. 3.8C, Fig. 3.9D). This suggests that the addition of a CRE sequence is not sufficient to reduce toxicity of the WT N2C virus. Rather, as these viruses differ only by the addition of a TEVs-PEST addition to the C terminus of the N, the reduced cytotoxicity and self-inactivation of the SiR virus compared to the WT virus is much more likely due to the effects of our appended N protein modification.

3.2.6 SiR 2.0 shows enhanced neurotropism and trans-synaptic transfer compared to oG coated B19 rabies virus in motor circuits

In addition to confirming SiR 2.0 viral self-inactivation and reduced cytotoxicity, we also confirmed that the SiR 2.0 exhibited increased neuroinvasive properties as would be expected from the newly modified N2C strain. When compared to the WT ΔG SAD-B19, the WT ΔG CVS-N2C strain has been reported to show 10- to 20- fold better trans-synaptic tracing efficiency in corticostriatal and spinal motor circuits[190].

The trans-synaptic transfer of the WT ΔG SAD-B19 strain has also recently been improved through the engineering of an optimized glycoprotein (oG) which is composed of a mouse codon-optimized chimeric G consisting of the transmembrane and cytoplasmic domain of the original B19G of the SAD-B19 strain with the extracellular domain of the rabies Pasteur virus strain G. Use of the oG to complement the WT ΔG SAD-B19 strain instead of the B19G has also been shown to improve monosynaptic tracing efficiency by around 10- to 20- in thalamocortical circuits[244].

Therefore when compared to the B19G complemented WT ΔG SAD-B19, both the oG coated SAD-B19 and the CVS-N2C strain show 10- to 20- fold increased trans-synaptic transfer in different intracerebral and spinal circuits. However, as far as we are aware, there has been no direct comparison of the trans-synaptic labelling efficiency of the two improved viruses in the same experimental set up. In the following experiments I directly compare the neurotropism and trans-synaptic labelling efficiency of the oG coated SAD-B19 against the CVS-N2C strain in motor-premotor circuits in the spinal cord following peripheral intramuscular injection.

3.2.6.1 SiR 2.0 shows increased peripheral neurotropism for motor neuron terminals in muscles

Here, we conducted a comparison of the total number of primarily infected motor neurons following injection of either the SiR N2C virus or the oG coated SiR B19 virus into the quadriceps of P3 pups (Fig. 3.9A). Equal volumes and comparable titers of either virus were injected. In general, we would inject around 3 fold higher titers of oG coated SiR B19 purely due to the fact that injections with titers below 5 x 10^8 (as set with traditional titration methods) resulted in unsuccessful labelling of motor neurons post-procedure. Quantification of infected motor neurons revealed that the N2C strain was significantly more efficient at infecting peripheral motor neuron terminals, giving an average of around 10- fold higher total number of infected motor neurons despite being injected at 3-fold lower titers (Fig. 3.9B).

3.2.6.2 SiR 2.0 shows increased trans-synaptic transfer efficiency

We also compared the trans-synaptic transfer efficiency of the SiR N2C compared to SiR B19 by calculating the ratio of pre-synaptically labelled premotor neurons per primarily infected motor neuron. Motor-premotor targeting was achieved by intramuscular co-injection of SiR viruses with AAVs expressing the oG or the N2C G protein and a nuclear FLAG tag (H2B-FLAGx3) for visualisation (Fig. 3.9A). Both the SiR B19 and SiR N2C encoded the same mCHERRYPEST CRE cassette, and the design of the G complementing AAVs were also identical apart from the coding sequence of the G protein. We found that SiR 2.0 shows both faster and more complete targeting of pre-synaptic inputs. At 3 days p.i. the SiR 2.0 already exhibited a premotor-motor neuron ratio of 6, whilst the SiR 1.0 exhibited only a ratio of 0.33 (Fig. 3.9C, representative images in Fig. 3.9E, G). This shows that despite the lower transcriptional levels and slower replication of the N2C strain, the virus can transport faster and more efficiently into secondary trans-synaptic neurons. At 1 week p.i. the premotor-motor ratios increased further to an average of 110 labelled premotor neurons/motor neuron for the SiR N2C as opposed to 40 for the SiR B19 (Fig. 3.9C, representative images in Fig. 3.9F, H).

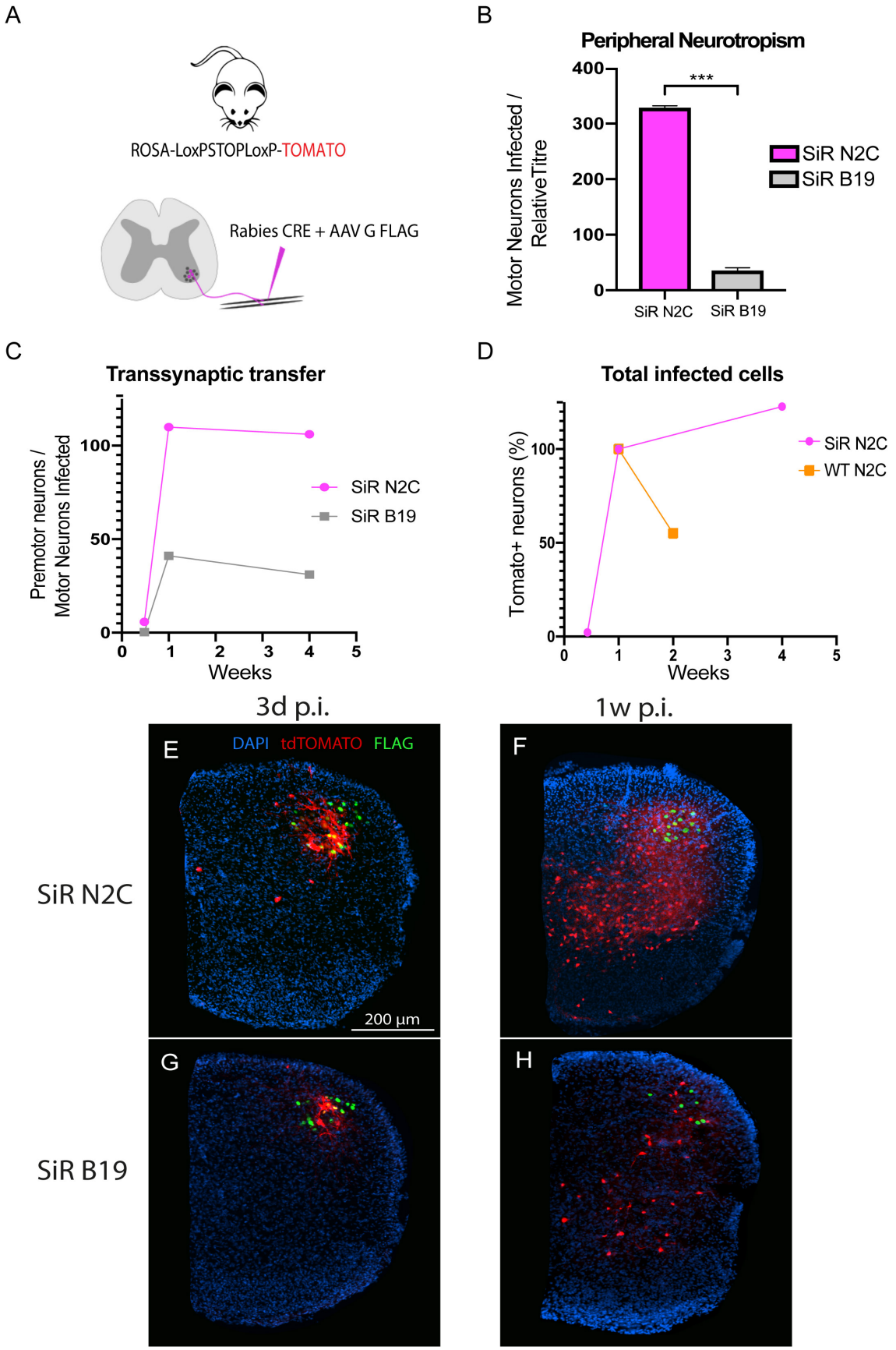


Fig. 3.9 SiR 2.0 shows enhanced neurotropism and trans-synaptic transfer efficiency in motor-premotor spinal circuits *Legend on following page...*

Fig. 3.9 (A) TOMATO CRE reporter P3 pups were co-injected with a SiR CRE virus (B19 or N2C) and AAV-H2BFLAGx3-P2A-G (oG or N2CG) in the quadriceps muscle for mono-synaptic tracing from motor neuron terminals. **(B)** Quantification of total motor neurons infected post-injection, normalised by relative titer of the two viruses as determined by RT-qPCR. (mean ± SEM, n = 3, Two-tailed, unpaired Student's T-test) **(C)** Quantification of premotor neurons trans-synaptically labelled per infected motor neurons infected at 3d, 1w and 4w. The SiR N2C shows both faster and more complete mono-synaptic tracing compared to the SiR B19. **(D)** Quantification of total virally infected cells at 3d, 1w and 4w. **(E-H)** Representative images of SiR N2C or B19 infected neurons (tdTOMATO, red) and AAV infected neurons (FLAG+, green) at 3d and 1w p.i. Double positive (tdTOMATO+,FLAG+) neurons are starter neurons from which premotor neurons are trans-synaptically traced.

3.2.7 SiR 2.0 shows markedly reduced immunogenicity compared to other rabies tracing vectors

Having confirmed the self-inactivation and higher neuroinvasiveness of the SiR 2.0, we also investigated the immunogenicity of the new SiR 2.0 vector compared to the other existing rabies vectors for circuit tracing (WT ΔG B19, SiR B19, and WT ΔG N2C viruses).

3.2.7.1 Innate immune response to rabies virus infection in mouse neuroblastoma cells *in vitro*

First, we began by investigating whether the secondary neuronal cell line (neuro2A) could activate an innate immune response to rabies virus infection. Mouse neuroblastoma cells have been shown to detect and respond to viral infections *in vitro*. For example, infection of neuro2A cells with japanese encephalitis virus, a positive strand single-stranded RNA virus, has been shown to be recognized by the pattern recognition receptor RIG-I PRR leading to downstream activation of MAPK and NF-kB signalling[245]. RIG-I has also been shown to play a significant role in the activation of innate immune responses to rabies virus infection in bone marrow derived dendritic cells[246]. As such, we decided to investigate the neuronal response to rabies virus infection by infecting mouse neuro2A cells with an oG coated WT ΔG B19 GFP virus. For the positive control, we added recombinant mouse TNFα to the culture medium. The two cognate TNFα receptors p55 (TNF-RI) and p75 (TNF-RII) are expressed on neurons and are known to upregulate the NF-kB signaling pathway on activation[247].

24-48 hours post infection, cells were fixed and immunostained for NF-kB and p38 MAPK (Fig. 3.10A). We found that there was a clear activation of NF-kB within neuro2A cells following infection with the virus that was maintained for up to 48 hours. The MAPK pathway was also activated within neuro2A cells but the response was more transient and only

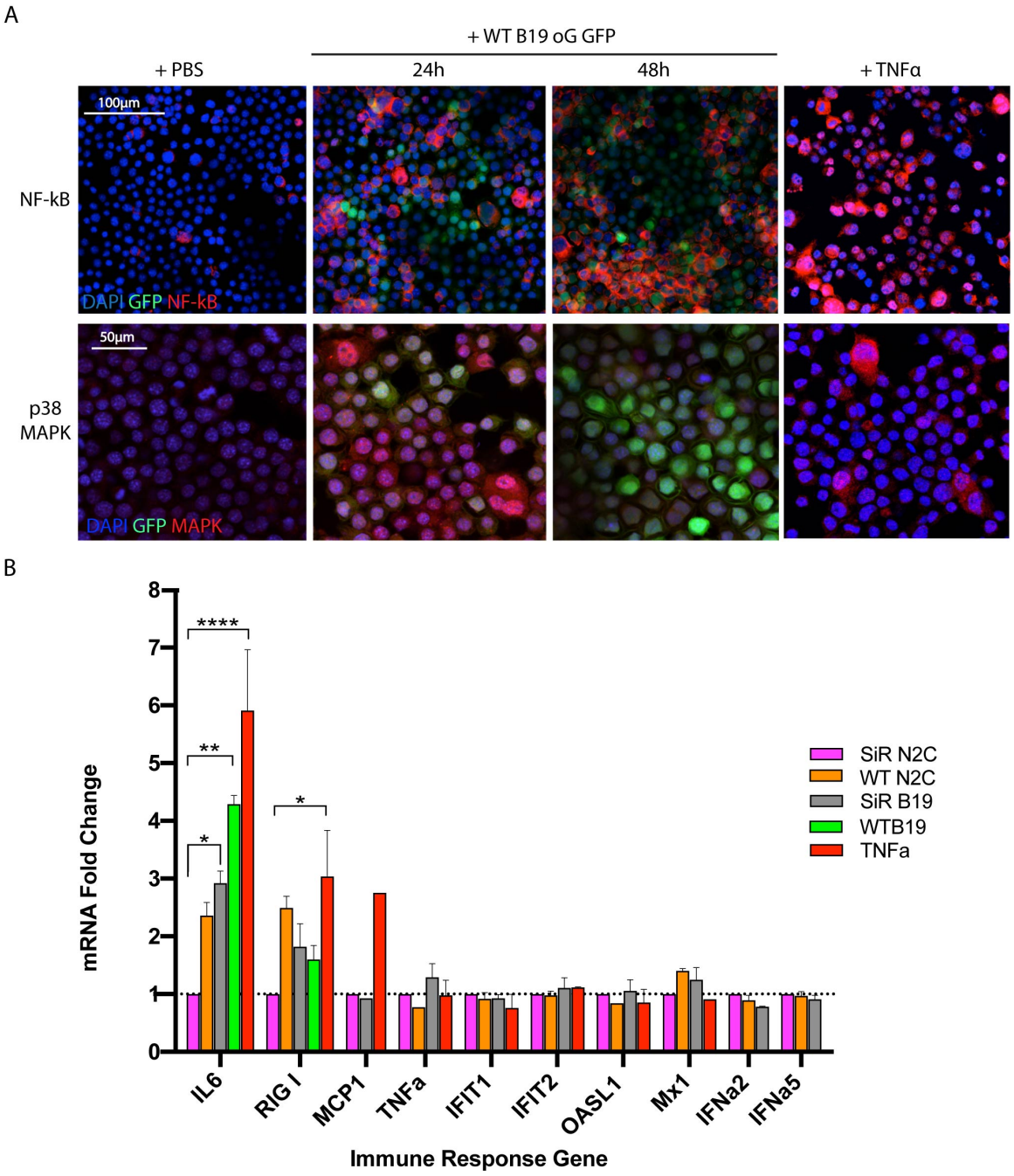


Fig. 3.10 Rabies viral infection can induce innate immune responses *in vitro* **(A)** Rabies infection of mouse neuro2A cells induces both NF-kB and p38 MAPK activation. oG coated WT B19 rabies virus expressing GFP was used to infect mouse neuroblastoma cells in culture and fixed 24 and 48 hours post infection for immunostaining for NF-kB (upper row) or p38 MAPK (bottom row). Equal volumes of PBS was added to cells as a negative control, and 100ng TNFα added as positive control for induction of innate interferon responses). **(B)** RTqPCR fold changes of selected immune response genes following rabies virus infection *in vitro*. Fold changes normalised to SiR N2C induced immune response gene transcript levels.

visible at the 24 hour time point but not the 48 hour timepoint. Having confirmed that mouse neuro2A cells could activate immune responses to WT B19 rabies virus infection, we wanted to compare the degree and extent of this immune activation by neuro2A cells following infection by each of the WT B19, SiR B19, WT N2C and SiR N2C. We also wanted to investigate any potential differences in the molecular responses to infection between each of the four vectors. Therefore we performed RTqPCR analyses on infected neuro2A cells with primers specific for various interferon stimulated genes and cytokines downstream to NF-kB and MAPK activation. We also investigated mRNA fold changes in the potential upstream PRR RIG-I which could potentially have triggered the observed NF-kB and MAPK response. We found that the IL-6 response showed significant variation between the 4 viruses, with WT B19 causing the strongest IL6 activation followed by SiR B19, WT N2C and lastly the SiR N2C. This indicated that the degree of immune response activation following neuro2A infection was lesser in SiR compared to WT viruses, and lesser in the N2C strain compared to the B19 strain. However, none of the other immune responses genes investigated showed significant differences in transcript levels between the four vectors.

This experiment indicated to us that the rabies virus did induce immune responses *in vitro*, and that there was potentially a difference in the degree of immune activation caused by the different viral vectors. This led us to further characterise the degree and molecular nature of immune activation following infection by the four rabies virus vectors in a more physiologically relevant setting *in vivo*. Although the *in vitro* system did reveal some patterns, we wanted to investigate further using a more physiological setting where not only undifferentiated neuroblastoma cells, but mature neurons, astrocytes, microglia and other immune related components could all act in concert to react to rabies virus infection.

3.2.7.2 SiR 2.0 downregulates the activation of innate viral sensor RIG-I *in vivo*

To investigate the early initiation stage of host immune responses to rabies viral infection *in vivo*, we studied the induction of RIG-I expression. RIG-I is an innate pattern recognition receptor that binds 5 triphosphate-ended RNA and activates potent type I IFNs in response to rabies virus infection[216]. We investigated this RIG-I response 1 week p.i. as this is within the short transcriptionally active time window of SiR viruses after initial infection but before viral shut off. CRE recombinase expressing rabies vectors were injected into the CA1 of YFP CRE reporter mice and sacrificed 1w p.i. for both studies at the transcriptional level (RT-qPCR from total extracted hippocampal RNA, Fig. 3.11A) and protein level (immunofluorescence, Fig. 3.11C).

RT-qPCR analysis of transcriptional levels of RIG-I and two downstream interferon stimulated genes IFIT1 (Interferon-induced protein with tetratricopeptide repeats 1) and Mx1 (Interferon-induced GTP-binding protein Mx1) confirmed that infection with the B19 strain led to an induction of RIG-I driven interferon responses when compared to PBS injected controls (Fig. 3.11B). Comparison of RIG-I, IFIT1 and Mx1 transcript levels between WT B19 and WT N2C injected hippocampi also revealed that the B19 strain induced higher levels of RIG-I activation than WT N2C (Fig. 3.11B). The same pattern of RIG-I activation was apparent from immunofluorescence studies. WT B19 infected hippocampal neurons trigger a more widespread area of positive RIG-I staining with very intensely stained RIG-I+ neurons and glial cells. In contrast, WT N2C infected neurons trigger significantly lower levels of RIG-I staining that is not visible at low zoom images (Fig. 3.11D, left inset). In addition, we also compared the RIG-I response in WT vs SiR viruses in both strains and found that the attenuation of the N protein in the SiR viruses resulted in further reduction of RIG-I responses in comparison to their parent WT strains. Reduced RIG-I activation was evident by a reduced intensity of RIG-I staining in SiR B19 transduced areas compared to the WT B19 (Fig. 3.11D, left inset) as well as by decreased density of RIG-I stained cytoplasmic antiviral stress granules (avSGs)[248–250] surrounding SiR N2C infected cells compared to the WT N2C (Fig. 3.11D, right inset). Taken together, we found that the attenuated SiR of the less immunogenic N2C strain resulted in a viral vector that showed minimal RIG-I activation compared to all other vectors.

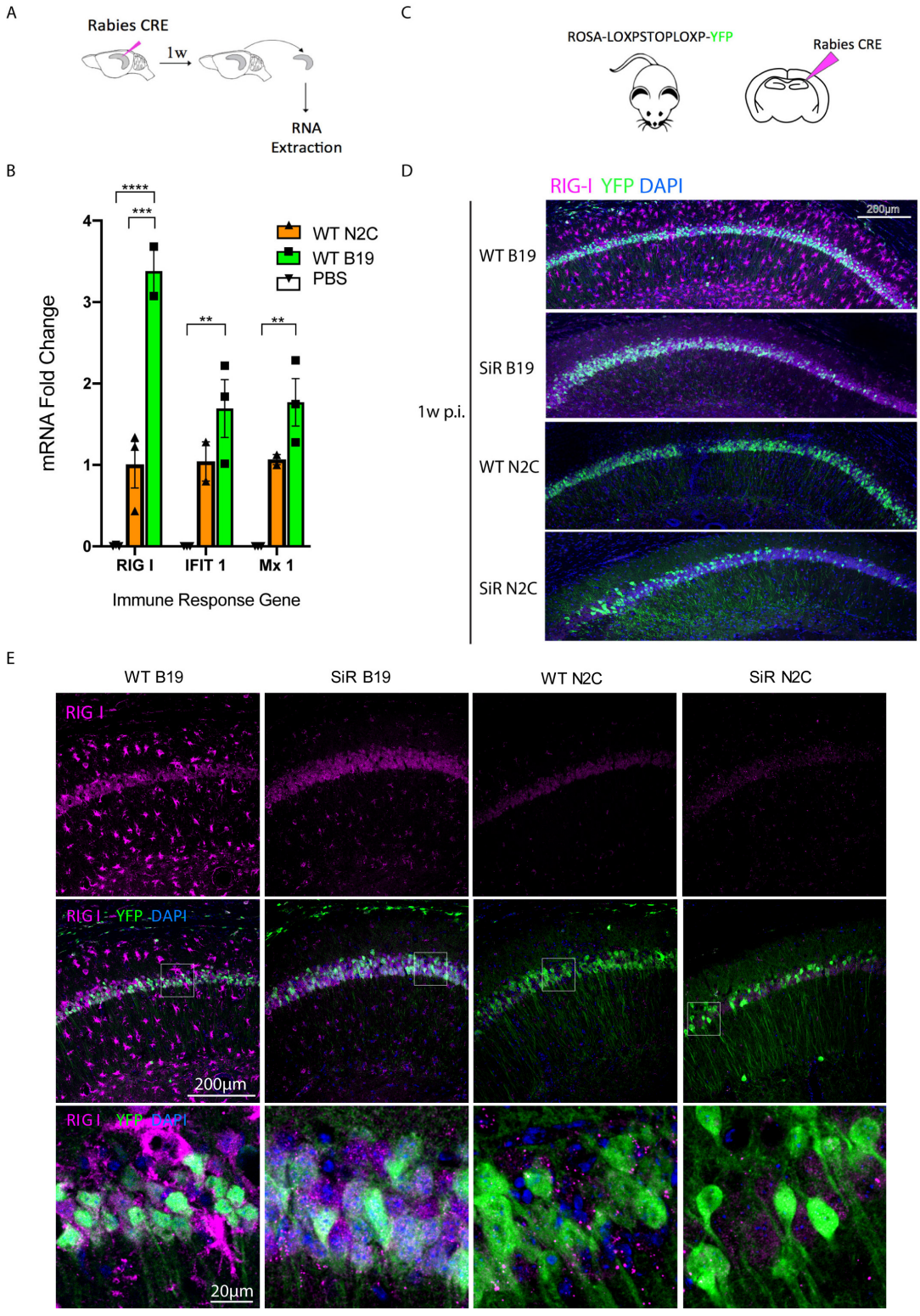


Fig. 3.11 SiR N2C attenuates activation of innate viral sensor RIG-I *in vivo* **(A)** Mouse CA1 was injected with PBS, WT B19 or WT N2C rabies virus. 1 week p.i. total RNA was extracted from hippocampi for RT-qPCR. **(B)** mRNA fold change of immune response genes (RIG-I, IFIT1, Mx1) normalised to WT N2C injected transcript levels. (mean±SEM, n=3, two-tailed unpaired student's T-tests) **(C)** YFP CRE reporter mice were injected in CA1 with rabies viruses expressing CRE recombinase (WT B19, SiR B19, WT N2C or SiR N2C) for immunofluorescence studies 1 week p.i. **(D)** Hippocampal slices were stained for RIG-I (magenta), YFP (green, virally infected cells) and DAPI (blue, nuclei). **(E)** Higher resolution images to show the "speckle" like distribution of RIG-I positive cytoplasmic antiviral stress granules.

3.2.7.3 SiR 2.0 triggers significantly reduced microglial responses following *in vivo* injection

Microglia are resident macrophages of the central nervous system and form a critical role in innate immunity. In the post-natal brain, microglia exist in a "resting" ramified form acting as sensitive surveyors of brain parenchyma. On detection of microbial infections or brain damage, they transform into an "active" ameboid form which can rapidly migrate into areas of brain injury to perform pro-inflammatory macrophagic functions[251]. They can also themselves be infected by viruses. For example, the rabies virus can infect microglia directly[252].

Although the underlying signalling mechanisms underlying this switch from "rest" to "active" state are not well understood, studies have shown that this microglial response is heterogeneous, with microglia responding to different inflammatory signals in different ways depending on anatomical location, species and age[253,254]. The role of microglia in inflammation can also be varied depending on their polarisation to the typically "damaging" pro-inflammatory phenotype (M1) characterized by TNF-α and IL-1β expression or the "repairing" anti-inflammatory phenotype (M2) characterized by IL-10 and TGF-β expression.

We investigated the degree of activated microglial recruitment to rabies virus infected neurons in the hippocampus (Fig. 3.12) and the spinal cord (Fig. 3.13). We found a pattern similar to that observed for RIG-I activation in both anatomically distinct circuits. The B19 strain induced stronger microglial responses than the N2C strain, and again the SiR viruses induced weaker microglial reponses compared to their WT counterparts (Fig. 3.12A, Fig. 3.13A-D). In combination, the attenuated SiR virus of the less immunogenic N2C strain led to very minimal microglial activation and recruitment compared to all other vectors. Quantification of Iba1+ cells showed that the SiR 2.0 virus recruited markedly smaller numbers of activated microglial cells then all other vectors in the hippocampus (Fig. 3.12B) and in the spinal cord (Fig. 3.13E). Quantification of other cellular parameters of Iba1+ cells, such as cell area (Fig. 3.12C and Fig. 3.13F), circularity (Fig. 3.12D) and Iba1+ staining intensity (Fig. 3.12E) showed that the few Iba1+ cells that were detected around SiR 2.0 infected cells did not show as prominent features of microglial activation as for the other vectors.

Longitudinal studies into Iba1+ activation between WT and SiR N2C infected neurons showed that the peak of microglial responses in the hippocampus are at 2 weeks p.i., and that the WT N2C induces stronger microglial responses then the SiR N2C at all time points post-infection (Fig. 3.12F). Longitudinal comparisons of microglial responses between SiR B19 and SiR N2C infected motor and pre-motor neurons in the spinal cord also revealed

that SiR B19 induces stronger microglial responses then the SiR N2C at all time-points investigated (Fig. 3.13E).

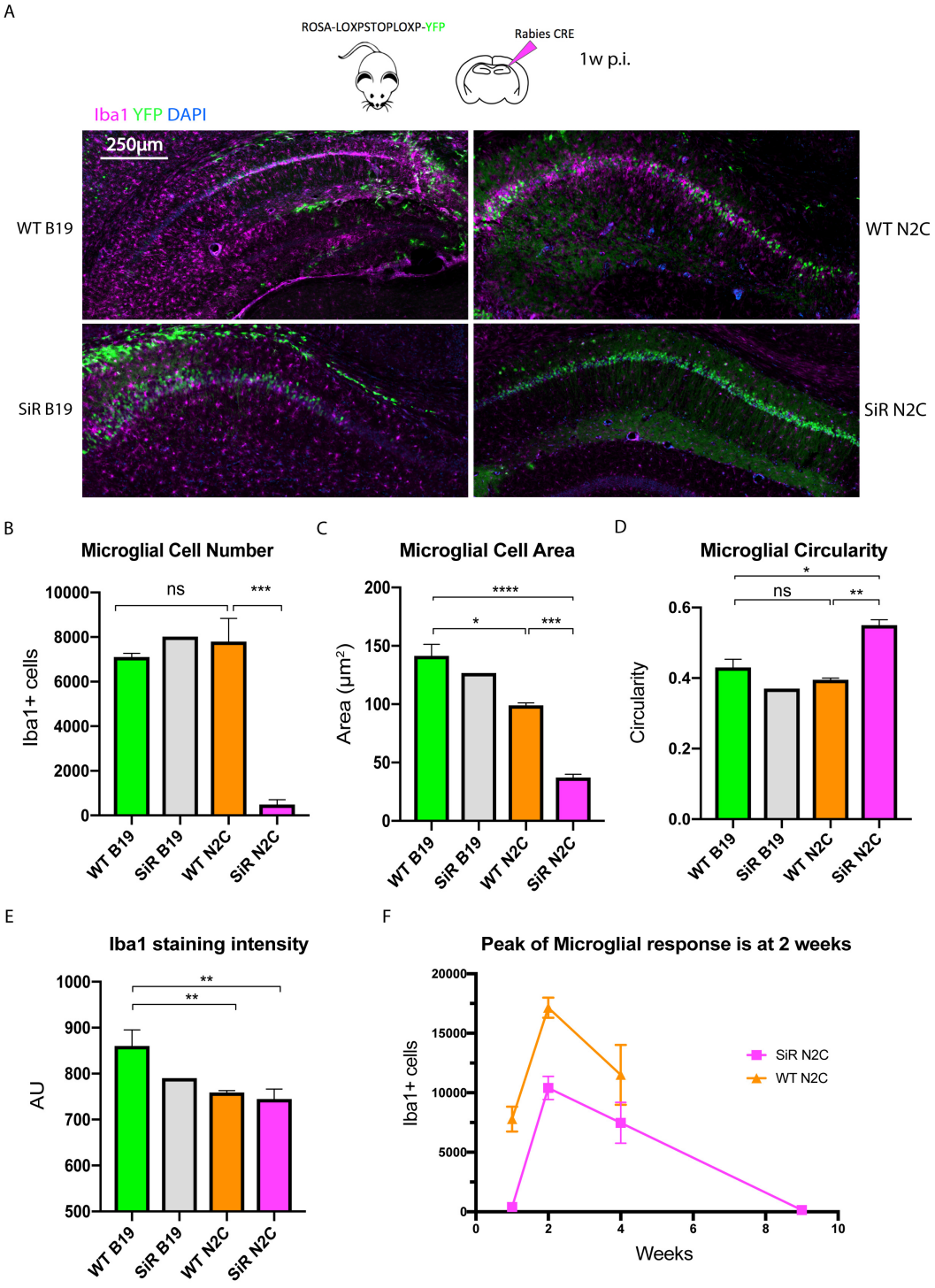


Fig. 3.12 SiR 2.0 shows minimal microglial activation in the hippocampus compared to alternative rabies virus vectors *in vivo* **(A)** YFP CRE reporter mice were injected in CA1 with rabies viruses expressing CRE recombinase (WT B19, SiR B19, WT N2C or SiR N2C) and sacrificed 1 week p.i. for immunofluorescence studies. Hippocampal slices were stained for microglial marker Iba1 (magenta, microglia), YFP (green, virally infected cells) and DAPI (blue, nuclei). **(B-E)** Quantification of Iba1+ cell number, area, circularity and staining intensity staining (mean±SEM, n=3, one-way ANOVA with tukey correction for multiple comparisons). **(F)** Longitudinal quantification of Iba1+ cell numbers in infected hippocampi at 1, 2, 4 and 9 weeks post-injections with WT or SiR N2C viruses (mean±SEM, n=3). WT N2C 9w time point not investigated as WT N2C infected neurons die 4w post-injection.

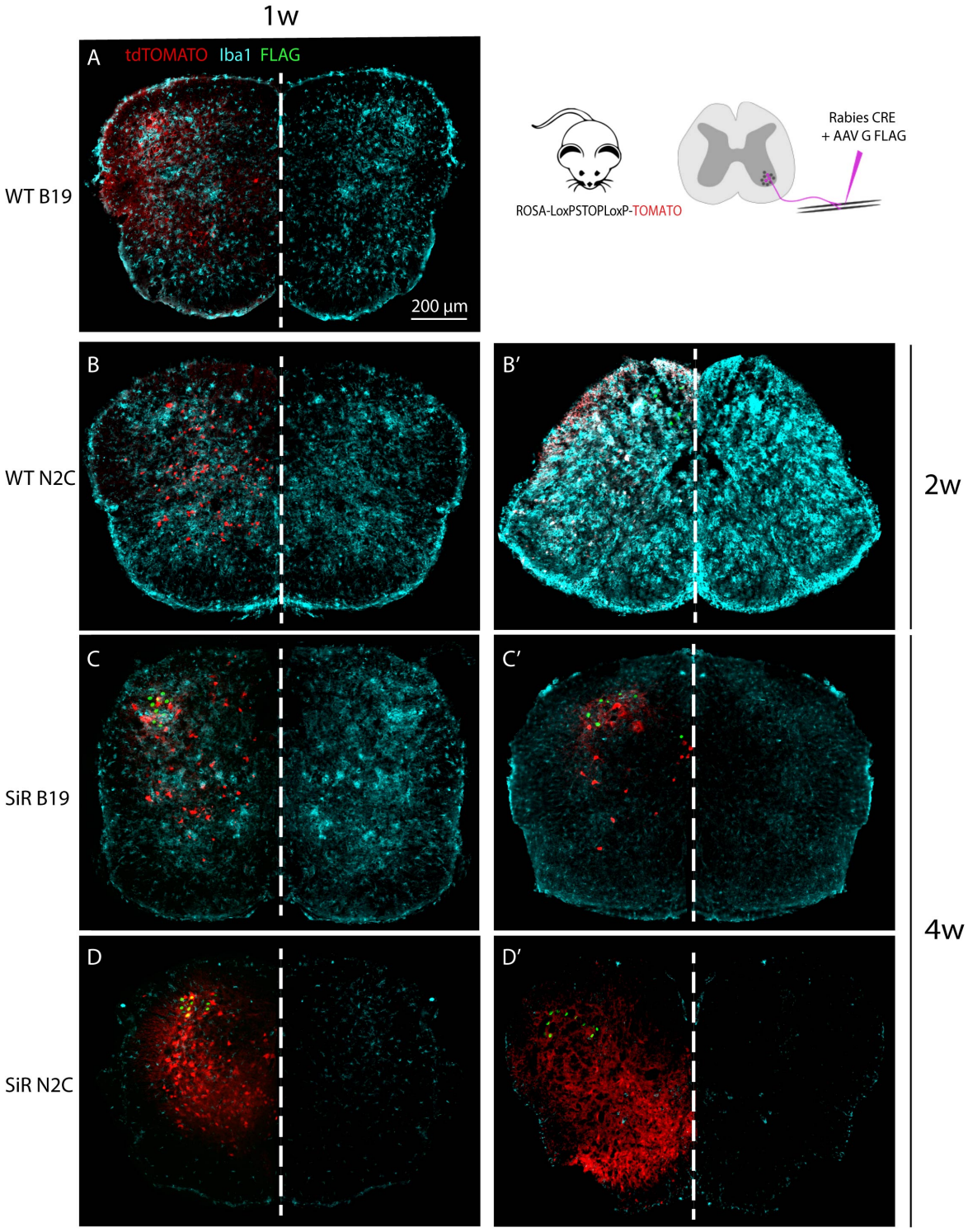


Fig. 3.13 SiR 2.0 shows minimal microglial activation in the spinal cord compared to alternative rabies virus vectors TOMATO CRE reporter pups were co-injected at P3 with a rabies virus expressing CRE recombinase (WT B19, SiR B19, WT N2C or SiR N2C) and an AAV expressing nuclear FLAG and the rabies G of the appropriate strain (N2C G or oG). Spinal cord sections were stained for Iba1 (cyan, microglia), TOMATO (red, virally infected cells) and FLAG (green). **(A-D)** representative images of spinal cord sections 1 week p.i. SiR N2C shows markedly reduced Iba1+ staining compared to all other rabies vectors. **(B')** WT N2C infected spinal cord section 2w p.i. during period of infected cell loss shows intense Iba1+ staining. **(C' and D')** SiR B19 and SiR N2C infected spinal cord section 4w p.i. respectively. Right side of white-dotted line shows reflected mirror image of injected side with only Iba1+ staining. *Legend continued on following page.*

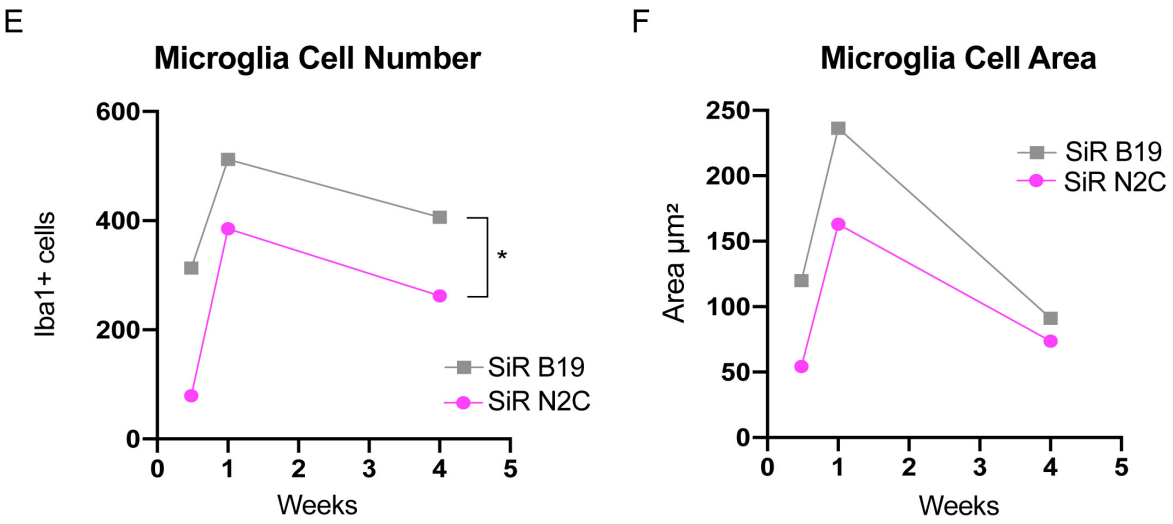


Fig. 3.13 SiR 2.0 shows minimal microglial activation in the spinal cord compared to alternative rabies virus vectors (E) Number of Iba1+ cells at 3d, 1w, 4w p.i. with SiR B19 or N2C viruses. **(F)** Area of Iba1+ cells at 3d, 1w, 4w p.i. with SiR B19 or N2C viruses.

3.2.7.4 SiR 2.0 triggers significantly reduced astrocyte responses following *in vivo* injection

Astrocytes are the most abundant glial cells in the central nervous system and play multiple roles including regulation of the CNS micro environment, synaptic transmission, neuronal metabolic and trophic support, maintenance of the BBB and in innate immunity. Following immune activation, astrocytes undergo reactive astrocytosis or astrogliosis. These reactive astrocytes are hypertrophic, have longer and thicker cytoplasmic processes and show enhanced expression of various proteins such as the glial fibrillary acidic protein (GFAP)[255]. The mechanisms underlying astrogliosis are not well understood, but it is known that astrocytes can be activated by many different triggers (e.g. direct infection by virus, neuronal damage or death etc.). Microglial release of cytokines such as IL-1, IL-6, TNF-α, IFN-γ is widely accepted to play a critical role in the induction of astrogliosis in viral infection[256]. Many of these pro-inflammatory molecules trigger the expression of inducible NO synthase (iNOS) expression in astrocytes, and NO has been shown to play a major role in the induction of GFAP activation[257]. Once activated, astrocytes can increase BBB permability[258], express proinflammatory cytokines[259] (e.g. TNF-a, IL-1, IL-6, macrophage CSF and granulocyte-macrophage CSF) and execute antiviral IFN responses[257].

Studies into the role of astrocytes in rabies virus pathogenesis have shown that rabies virus can directly infect astrocytes *in vitro*[252]. In addition, it has been found that non-pathogenic rabies (e.g. the SAD-L16 vaccine strain[260] or lab-attenuated CVS-B2C strain[261]) triggers

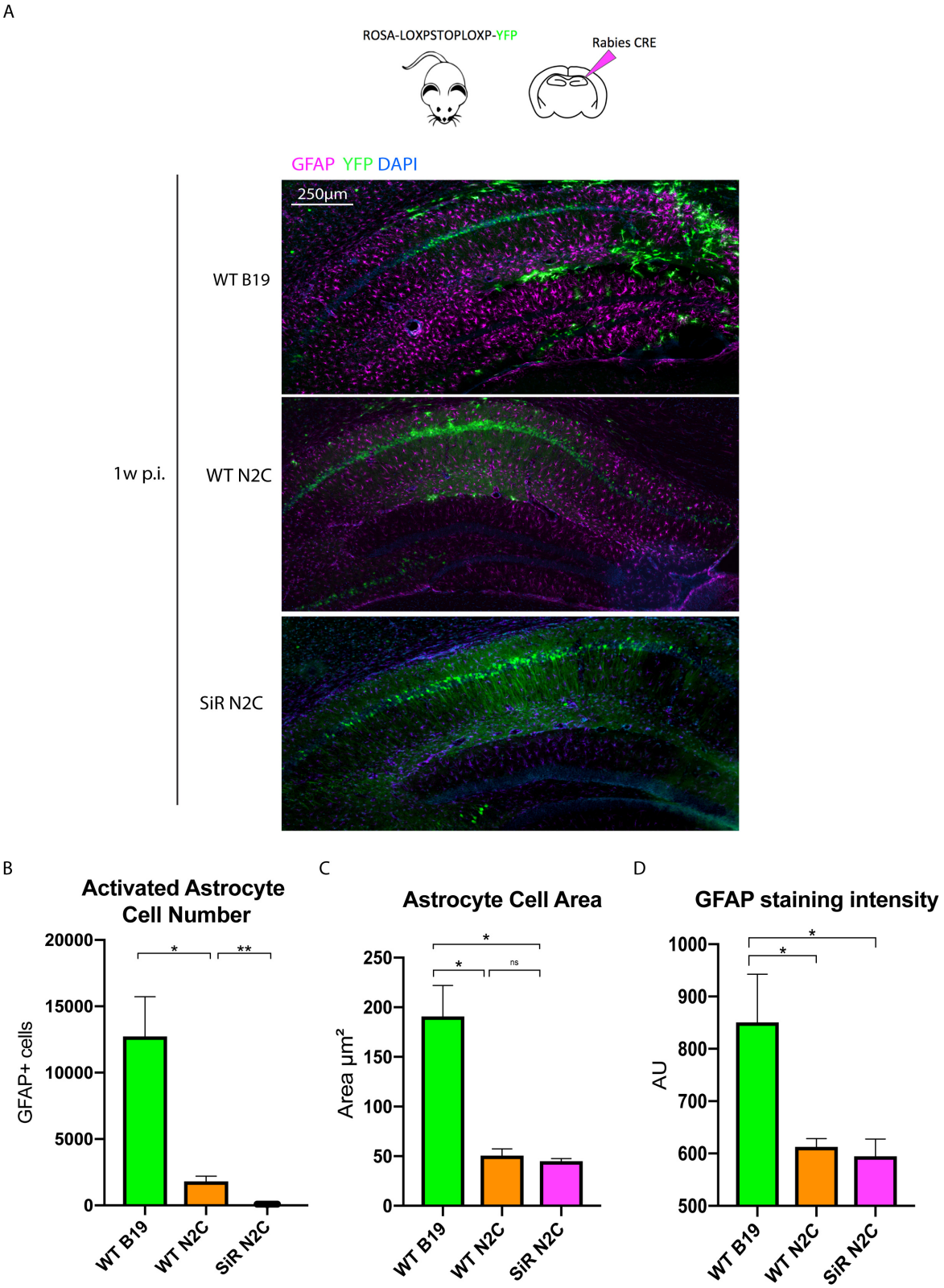


Fig. 3.14 SiR 2.0 shows minimal astrocyte activation in the hippocampus compared to alternative rabies vectors *in vivo* (A) YFP CRE reporter mice were injected in CA1 with rabies viruses expressing CRE recombinase (WT B19, WT N2C or SiR N2C) and sacrificed 1 week p.i. for immunofluorescence studies. Hippocampal slices were stained for astrocyte marker GFAP (magenta, astrocytes), YFP (green, virally infected cells) and DAPI (blue, nuclei). **(B-D)** Quantification of GFAP+ cell number, area, staining intensity staining (mean±SEM, n=3, one-way ANOVA with tukey correction for multiple comparisons).

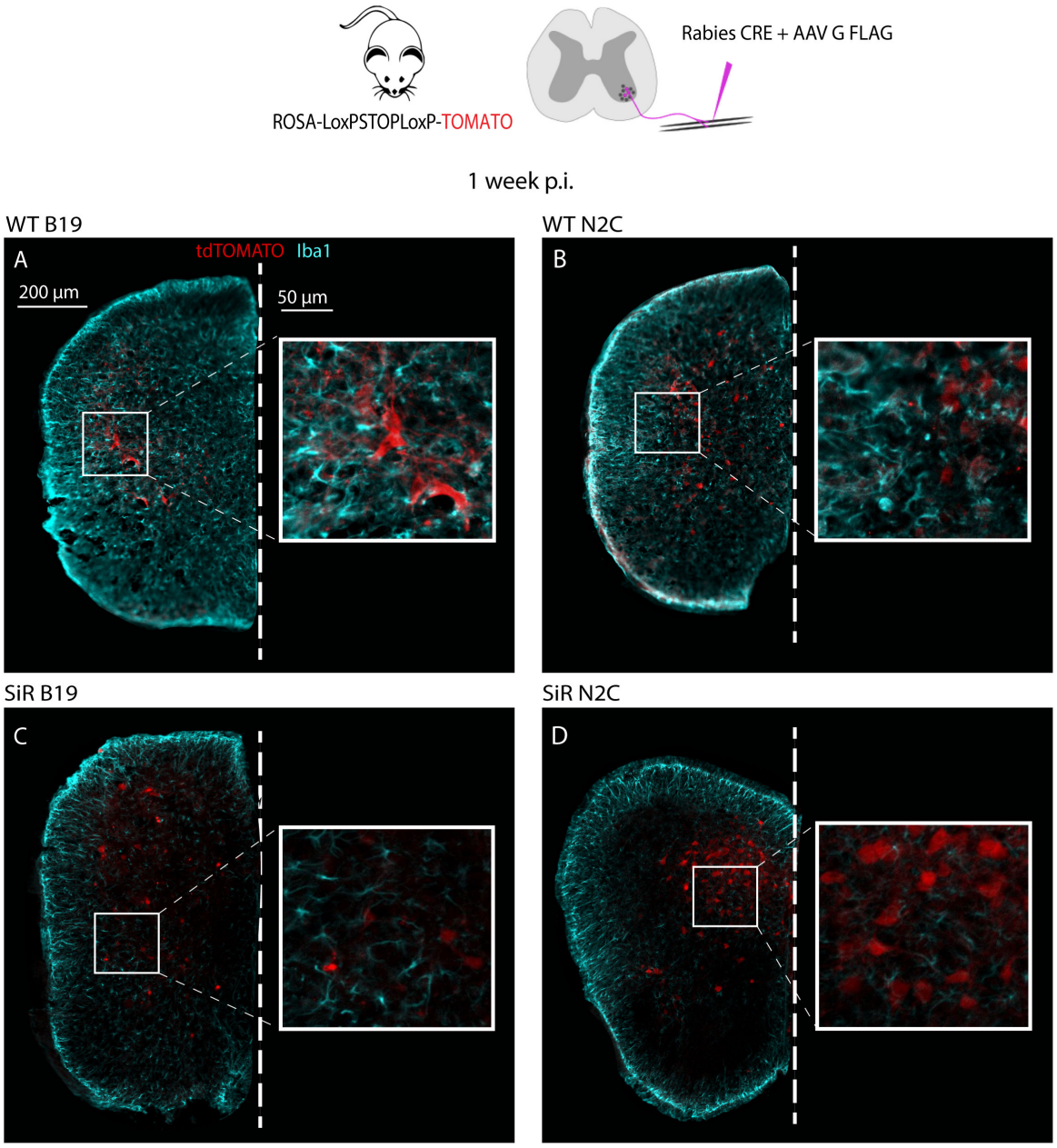


Fig. 3.15 SiR 2.0 shows minimal astrocyte activation in spinal motor circuits compared to alternative rabies vectors *in vivo* TOMATO CRE reporter pups were co-injected at P3 with a rabies virus expressing CRE recombinase (WT B19, SiR B19, WT N2C or SiR N2C) and an AAV expressing nuclear FLAG and the rabies G of the appropriate strain (N2C G or oG). Spinal cord sections were stained for GFAP (cyan, astrocytes), TOMATO (red, virally infected cells) and DAPI (blue, nuclei). **(A-D)** representative images of spinal cord sections 1 week p.i. SiR N2C shows reduced GFAP+ staining compared to all other rabies vectors.

RIG-I induced IFN responses in infected astrocytes which results in abortive infections. In contrast, pathogenic street isolates of rabies virus can maintain persistent astrocyte infections by evading immune recognition.

Here, we investigated the induction of astrocyte activation following both central and peripheral injections of SiR 2.0 compared to other experimental rabies virus vectors (WTB19, SiRB19, WTN2C). Again, we found a gradation of response similar to that observed for RIG-I activation and microglial activation. Namely, the B19 strain induced significantly stronger astrogliosis in the hippocampus compared to the WT N2C strain. Immunofluorescence staining for GFAP revealed a more widespread and intense staining for GFAP surrounding WT B19 injected areas then WT or SiR N2C vectors (Fig. 3.14A, representative images). Quantification of GFAP+ cells revealed significantly reduced numbers of activated astrocytes (Fig. 3.14B), reduced astrocyte hypertophy (Fig. 3.14C) and GFAP staining intensity (Fig. 3.14D) following WT N2C infection compared to WT B19. In addition, as for RIG-I and microglial responses, the SiR attenuated viruses showed decreased GFAP activation than their WT counterparts in both the hippocampus (Fig. 3.14A) and the spinal cord (Fig. 3.15 A vs C for B19, or B vs D for N2C). Quantification of GFAP+ cells revealed a markedly reduced number of GFAP+ activated astrocytes surrounding SiR N2C infected neurons compared to WT N2C in the hippocampus.

3.2.8 Investigating the varied transcriptional responses to rabies virus infections *in vivo* via RNA-seq

Given our findings regarding the innate immune response to rabies virus in neuroblastoma cells *in vitro* and in the mouse hippocampus and spinal cord through immunohistochemistry, we wanted to investigate with higher throughput the molecular players involved in this immune response, and how it may be different between WT B19, WT N2C and SiR N2C viruses. Therefore we decided to conduct whole transcriptome shotgun sequencing (more commonly called RNA-seq) in order to identify differentially regulated transcripts following rabies virus infection. Mice were injected with PBS, WT B19, WT N2C or SiR N2C virus in the CA1 region of the hippocampus (n=3 animals, per virus). 1 week p.i. the hippocampi were dissected out, RNA extracted and prepared for next generation sequencing (NGS). As we were interested to look not only at mRNAs but also non-coding RNAs involved in antiviral responses and also the viral genome RNA itself, we used a total RNA preparation kit that removed only ribosomal RNA from our sample preparation.

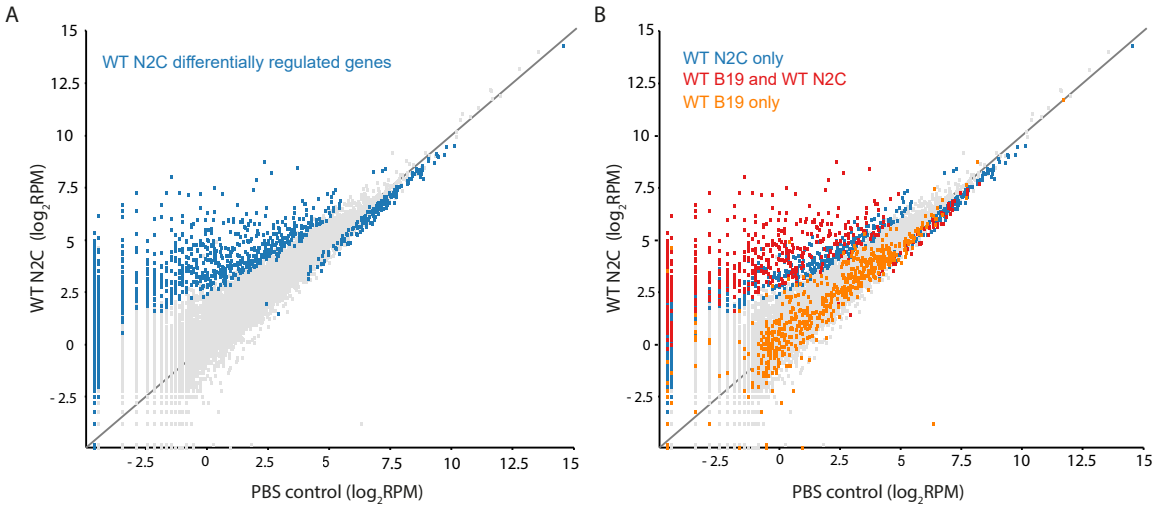
3.2.8.1 Large-scale functional characterisation of transcripts up or downregulated following WT rabies virus infection in the mouse hippocampus

Using the DESeq2 differential gene expression analysis[262] and intensity difference filtering, we identified 1577 differentially regulated transcripts in WT N2C infected hippocampi when compared to PBS injected controls (Fig. 3.16A). For WT B19 infected hippocampi, we also identified 1139 differentially regulated transcripts compared to PBS injected controls. Comparing the differentially regulated transcripts between WT N2C and WT B19 infected hippocampi, we found that around 50% of identified transcripts were overlapping between the two viruses (red, Fig. 3.16B). This suggested to us that the two different strains induce partly different transcriptional responses following injection *in vivo*.

To further understand this large-scale transcriptional response, we conducted gene ontology (GO) enrichment analysis using the panther classification system to identify over-represented gene groups in our list of 1577 differentially regulated genes.

In the up-regulated transcripts, over-represented GO groups revealed many defense immune system processes enriched in WT N2C infected brains (Fig. 3.16C). This was as we were expecting; the parental biological process GO term 0006955 for "immune response" was enriched with a P-value of 9.89E-20. Looking further at the enriched child terms of this GO term we identified more detailed immune system processes such as cytokine production, interluekin signalling, MHC class II antigen presentation, CD4 and CD8+ T cell differentiation, B cell activation, neutrophil activation and monocyte extravasation. These findings were all in line with previous work describing immune responses to rabies virus, including interferon signalling, increased BBB permeability, and induction of humoral immune responses resulting in the production of virus neutralising antibodies (VNAs). Interestingly, when looking at enriched GO terms for the regulation of immune responses, we found that there was an overall pattern of positive regulation of innate immune responses, but a negative regulation of adaptive immune response (Fig. 3.16C and 3.17). This may in part be a result of viral immune evasion mechanisms, as previous research has shown that more virulent rabies virus strains are better at evading increased BBB permeability and neuroinflammation in the CNS post-infection.

For the down-regulated transcripts, we found over-represented GO groups relating to synapse regulation, signalling and neurogenesis (Fig. 3.16D). This suggested to us that WT N2C infection of neurons in the hippocampus at 1 week p.i. was impairing usual synapse functions, as may have been expected from our knowledge that WT N2C infection is eventually toxic and leads to death of infected neurons by 4 weeks p.i.. More specifically, when looking at "cellular component" GO groups we particularly identified the downregulated



C

Gene ontology of upregulated transcripts following rabies injection

GO group	GO term ID	Description	P-value
Biological Process	GO:0006955	immune response	9.89E-20
	GO:0009615	response to virus	1.01E-07
	GO:0045071	negative regulation of viral genome replication	2.59E-06
	GO:0045089	positive regulation of innate immune response	2.74E-05
	GO:0001819	positive regulation of cytokine production	4.1E-05
	GO:0035723	interleukin-15-mediated signaling pathway	4.66E-04
	GO:0002443	leukocyte mediated immunity	6.12E-22
	GO:0019886	antigen processing and presentation of exogenous peptide antigen via MHC class II	2.21E-08
	GO:0043369	CD4-positive or CD8-positive, alpha-beta T cell lineage commitment	8.88E-08
	GO:0042113	B cell activation	8.66E-14
	GO:0035696	monocyte extravasation	4.66E-04
	GO:0042119	neutrophil activation	6.14E-10
	GO:0050868	negative regulation of T cell activation	5.06E-05
	GO:1903038	negative regulation of leukocyte cell-cell adhesion	2.51E-04
	GO:0002820	negative regulation of adaptive immune response	6.66E-04

GO Group	GO term ID	Description	P-value
Reactome	R-MMU-1280215	Cytokine Signaling in Immune system	2.75E-26
	R-MMU-1280218	Adaptive Immune System	4.60E-23
	R-MMU-6798695	Neutrophil degranulation	3.16E-21
	R-MMU-109582	Hemostasis	3.36E-19
	R-MMU-198933	Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell	9.30E-19
	R-MMU-449147	Signaling by Interleukins	7.46E-18
	R-MMU-76002	Platelet activation, signaling and aggregation	2.12E-13
	R-MMU-451927	Interleukin-2 family signaling	2.70E-11
	R-MMU-202733	Cell surface interactions at the vascular wall	7.11E-11
	R-MMU-162582	Signal Transduction	9.52E-11
	R-MMU-1236975	Antigen processing-Cross presentation	1.39E-10
	R-MMU-168898	Toll-Like Receptors Cascades	2.97E-10
	R-MMU-913531	Interferon Signaling	2.91E-09

GO group	GO term ID	Description	P-value
Cellular Component	GO:0009897	external side of plasma membrane	5.10E-16
	GO:0009898	cytoplasmic side of plasma membrane	1.75E-05

D

Gene ontology of downregulated RNAs following rabies injection

GO group	GO term ID	Description	P-value
Biological Process	GO:0050803	regulation of synapse structure or activity	1.98E-11
	GO:0099537	trans-synaptic signaling	1.97E-10
	GO:0022008	neurogenesis	3.03E-10

GO group	GO term ID	Description	P-value
Cellular Component	GO:0043005	neuron projection	3.29E-27
	GO:0097060	synaptic membrane	2.97E-21
	GO:0150034	distal axon	4.08E-11
	GO:0098685	Schaffer collateral - CA1 synapse	2.77E-08
	GO:0098686	hippocampal mossy fiber to CA3 synapse	3.58E-04

Fig. 3.16 Transcriptional response to WT rabies virus infection **(A)** Scatter plot depicting differentially regulated transcripts in WT N2C injected vs. PBS injected control mice (blue). **(B)** Scatter plot depicting genes that are differentially regulated following WT N2C injection (blue) as in subfigure (A) but also depicting genes that are commonly differentially regulated following both WT B19 and WT N2C injection (red), and genes that differentially regulated following WT B19 injection only (orange). **(C)** Gene ontology (GO) terms enriched in upregulated transcripts following WT N2C injection compared to PBS injected controls. **(D)** Gene ontology (GO) terms enriched in downregulated transcripts following WT N2C injection compared to PBS injected controls. n=3 for all experiments

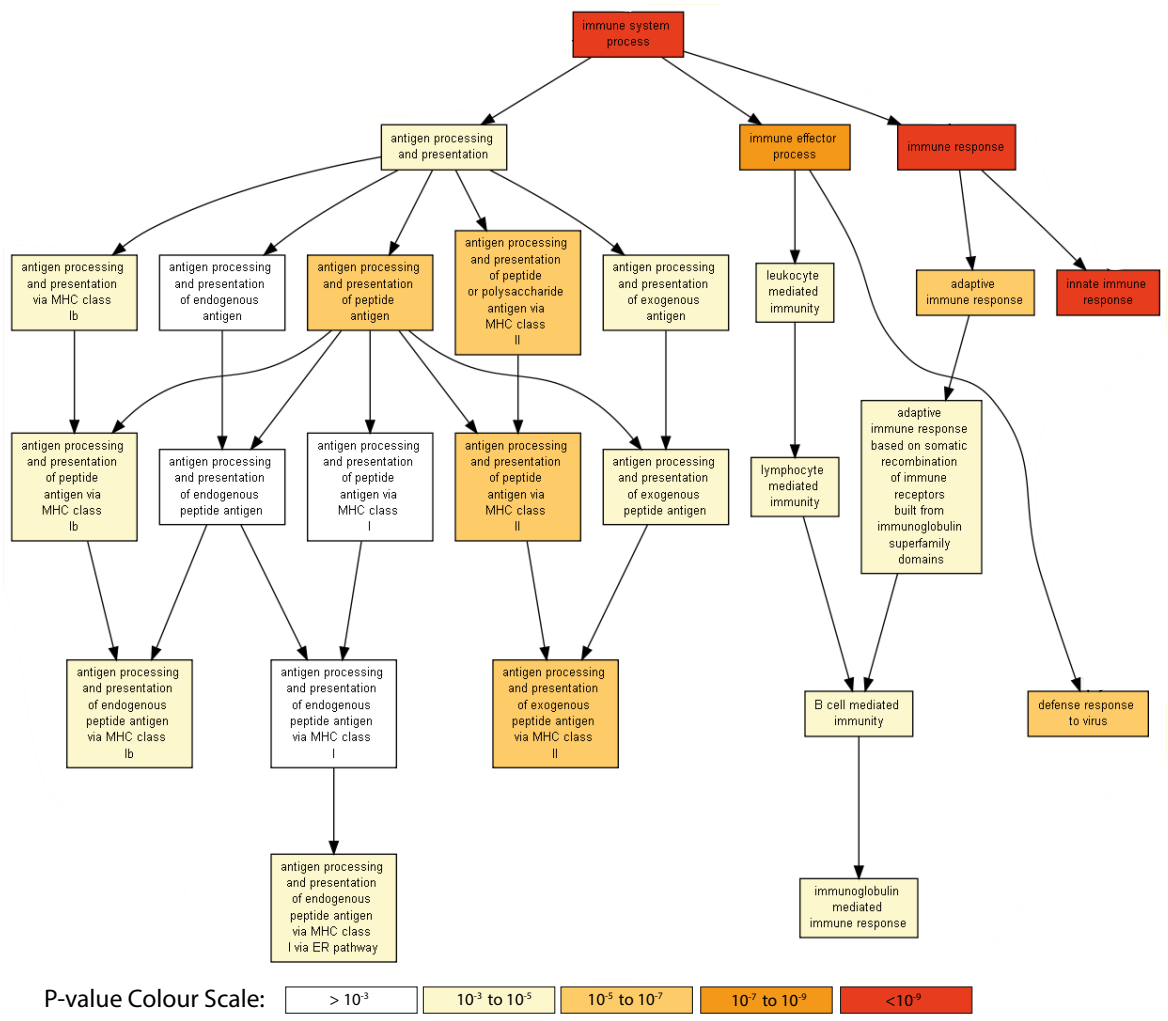


Fig. 3.17 Hierarchy of GO terms of "immune system process" enriched in upregulated RNA transcripts from rabies virus infected hippocampi

GO terms 0098685 and 0098686 signifying gene groups related to "Schaffer collateral - CA1 synapse" and "hippocampal mossy fiber to CA3 synapse" respectively. This was notable as these were the very neurons being targeted for viral infection by our injections. Therefore, downregulation of these GO terms revealed a clear specific impairment of function of neurons infected with WT N2C rabies virus in CA1. In converse, these GO terms were not identified for the 4 identified downregulated genes following SiR N2C infection, showing that at 1 week the SiR N2C does not impair synapse function like the WT N2C virus does (Fig. 3.18A). The 4 transcripts downregulated by SiR N2C are listed in Fig. 3.18B.

3.2.8.2 SiR N2C induces a significantly reduced transcriptional response following infection of hippocampal neurons

When applying the DESeq2 statistical filter to find differentially expressed transcripts in SiR N2C infected hippocampi vs. PBS infected controls, we found that the SiR virus upregulated only 50 genes post-infection compared to the 895 genes upregulated for the WT N2C virus (Fig. 3.18A). The SiR N2C upregulated transcripts are listed in Appendix3 (TableC.1). In a similar line, the SiR N2C virus downregulated only 4 transcripts compared to the 173 transcripts downregulated following WT N2C injection. The SiR N2C did not induce downregulation of synaptic or Schaffer collateral - CA1 proteins as seen with the WT N2C virus (Fig. 3.18A,B), strongly supporting the reduced synaptic toxicity of the SiR N2C virus. The scatter plot in Fig. 3.18C shows all WT N2C differentially regulated genes (blue, purple) as identified by the DESeq2 filter, and highlights in red the transcripts that are also commonly upregulated following SiR N2C infection. Notably one can see that the immune markers that we visualised using immunofluorescence methods (the microglial marker (Iba1), astrocytosis marker (GFAP) and the viral RNA sensor (RIG-I)) were found to be upregulated only following WT N2C infection but not following SiR N2C infection. This was very much in line with the increased staining intensity for these markers observed in WT N2C infected hippocampi compared to SiR N2C hippocampi.

3.2.8.3 Immune related transcripts reveal details of the decreased immunogenicity of SiR N2C

Having observed enriched GO terms related to immune responses, we wanted to study the immune response to WT N2C and SiR N2C infections in more molecular detail. In order to do this we first identified genes upregulated following WT N2C injection that fell under various receptor families, known signalling pathways and antiviral immune processes. The

A

	WT N2C	SiR N2C
Upregulated Transcripts	**895**	**50**
immune system process (GO:0002376)	390	37
Downregulated Transcripts	**173**	**4**
synapse (GO:0045202)	72	0
Schaffer collateral - CA1 synapse (GO:0098685)	12	0

B

Gene Name	Gene Summary	P-value
Jarid2	Regulator of histone methyltransferase complexes. Transcriptional repressor. Plays an essential role during embryonic development.	5.27E-07
Filip1l	Filamin A binding protein. By acting through the filamin-A/F-actin axis, regulates cortical neuron migration and dendritic spine morphology.	8.05E-07
Kpna1	Member of the importing alpha family, involved in nuclear protein import. Involved in viral processes (GO:0016032).	1.07E-05
CT010467.1	18s rRNA, related sequence 5.	1.24E-04

C

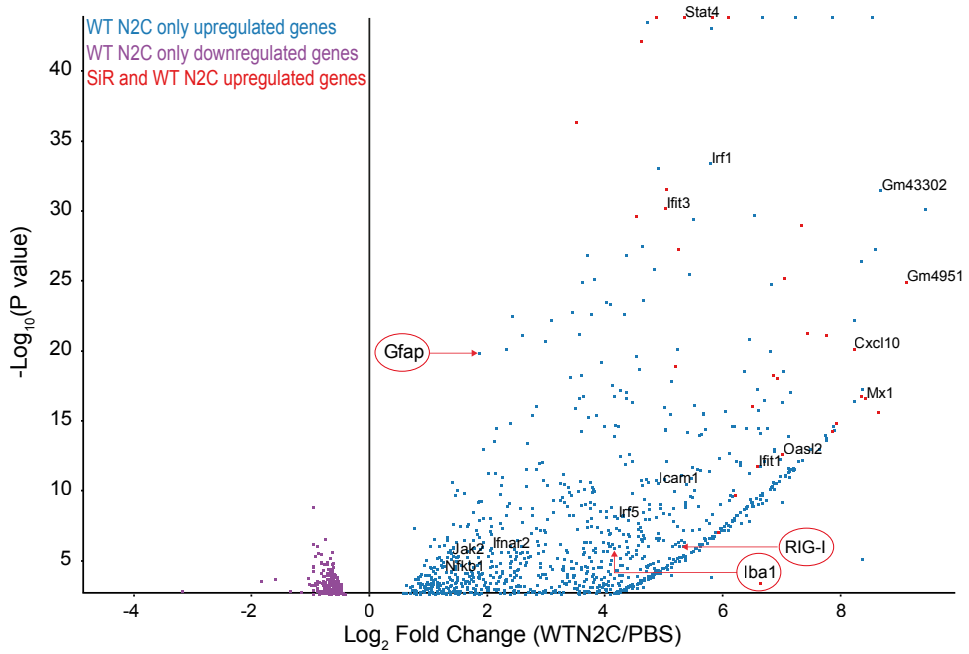


Fig. 3.18 SiR N2C upregulates less transcripts following injection compared to WT N2C (A) Table showing the number of up- or downregulated transcripts following WT or SiR N2C infection as identified by DESeq2 analysis. Identified genes falling under selected GO-terms depicts the lesser immunogenic response and lesser perturbation of synapse biology following SiR N2C injection compared to WT N2C injection. **(B)** Table listing the 4 downregulated transcripts following SiR N2C injection. Gene summaries taken from UniProtKB/Swiss-Prot. **(C)** Scatter plot of genes found to be differentially regulated following WT N2C injection vs. PBS control using the DESeq2 statistical filter (blue, purple). WT N2C upregulated genes that are also upregulated following SiR N2C injection are shown in red. The names of some notable examples of upregulated genes are noted on the upper, right side of the transcript. Examples of only WT N2C upregulated transcripts include GFAP, Iba1 and RIG-I which are in line with histological immunofluorescence findings from previous section and are circled in red.

fold changes in expression level of these genes in WT N2C infected samples compared to PBS injected controls was noted. We then also investigated whether these same genes were upregulated following SiR N2C infection. We found that in most cases the SiR N2C did not upregulate these immune response transcripts, and for the cases where there was co-upregulation, the SiR N2C caused a lesser fold change increase.

Starting from viral detection, WT N2C upregulates pattern recognition receptors (PRRs) of the RIG-I like receptor (RLR), toll-like receptor (TLR) and NOD-like receptor (NLR) family (Fig. 3.19A). Pro-inflammatory signalling pathways downstream of activated PRRs result in the activation of NF-kB and IRF transcription factors which can bind ifna and ifnb promoters to induce the production of type I interferons (Fig. 3.19D). Indeed, we also found many members of the NF-kB and IRF family to be upregulated following WT N2C injection (IRF1, 2, 4, 5, 7 and NFkB1, 2). Type I interferons upregulated by NF-kB and IRF transcription factors in initially infected cells bind IFNA receptors in other cells to activate JAK/STAT (Janus kinase/Signal Transducers and Activators of Transcription) signalling cascades. Activated STATs then induce the activation of interferon stimulated genes (ISGs) that have antiviral functions (Fig. 3.19E). As expected, we identified in our list of WT N2C upregulated transcripts, many genes involved in the antiviral interferon response. WT N2C upregulated IFN receptors, members of the JAK/STAT signalling pathway and downstream ISGs such as those of the 2'-5'-oligoadenylate synthase (Oas) and Oas-like family (Fig. 3.19B).

As for the SiR N2C virus, no PRRs were found to be significantly upregulated following injection. Of the downstream IRF/NF-kB pathway members, only IRF7 was found to be upregulated. In antiviral immune responses, IRF7 forms a positive regulatory feedback loop with type I interferons resulting in enhanced expression of IRF7. Following initial "priming" activation by phosphorylation downstream to activated PRRs, IRF7 binds ifna and ifnb promoters to induce expression of type I interferons. Type I interferons bind IFNA receptors to activate JAK/STAT signalling cascade in other cells. Activated STAT1/2 bind to IRF9 to form IFN-stimulated gene factor 3 (ISGF3) which induces the synthesis of more IRF7[263]. This newly synthesized IRF7 then induces more IFN production leading to a positive feedback loop resulting in a high level of IRF7 and IFN expression following viral infection. In line with this, the SiR N2C was also found to upregulate STAT2 and 4 and the expression of some ISGs (e.g. Mx1, IFIT1,3 etc.). These findings show that the SiR N2C induces a lower grade antiviral immune response following infection compared to the WT N2C virus.

We then looked at WT N2C upregulated transcripts downstream to interferon activation. We observed that there was an upregulation of many IFN induced cytokines associated with increased blood brain barrier permeability and leukocyte infiltration into the CNS (CXCL10, CXCL9, CCL5). It has previously been shown that the expression of inflammatory chemokines is closely associated with leukocyte infiltration and BBB permeability increases following infection by apathogenic lab strains of rabies virus, whilst pathogenic strains do not cause such an upregulation of chemokine expression and do not recruit immune cells into the CNS for rabies virus clearance[264]. The chemokines CXCL10, CXCL9 and CCL5 are three structurally and functionally related IFNg induced proteins that bind the CXCR3+ receptor[265]. They are most commonly attributed to leukocyte trafficking into the CNS, principally acting on activated CXCR3+ CD4+ Th1 cells, CD8+ T cells and NK cells. In addition it is hypothesised that infiltrating Th1 cells following rabies infection differentiate into IL-17 producing Th17 cells which initiates the alteration of tight junctions and increased BBB permeability[266,267]. Intercellular adhesion molecules such as ICAM1 known to facilitate leukocyte endothelial transmigration were also found to be upregulated in WT N2C infected brains.

Lastly, we also found that 3 out of the top 5 upregulated genes following WT N2C infection were major histocompatibility complex (MHC) class I and II proteins involved in the presentation of either cytosolic or extracellular peptide antigens onto the cell surface respectively. MHC class I is expressed in all nucleated cells, whilst MHC II expression is generally restricted to antigen presenting cells (APCs) such as macrophages, dendritic cells and B cells. Both MHC class I and II expression can be upregulated by IFNg signalling. The two NOD-like receptor (NLRs) family members (NLRC5 and CIITA) known to regulate MHC class I and II transcription respectively[268], were also found to be upregulated in WT N2C infections. These findings show that increased antigen presentation is a component of the antiviral immune response to rabies virus and that again, this immune response is decreased in SiR N2C infected hippocampi.

3.2.8.4 Non-coding RNA transcriptional response to rabies infection

Given increasing studies highlighting the role of long non coding RNAs (lncRNAs) in antiviral immune responses, and one recent study that identified 140 lncRNAs differentially regulated following CVS11 rabies virus infection, we also looked at the differential expression of lncRNAs in WT N2C and SiR N2C infected brains. The field of virus-induced lncRNA responses is still fairly young, with the first virus-induced transcript being reported following Japanese encephalitis infection in the mouse brain in 2006[269]. However, there are

A

Viral Detection

	Probe	Log2 Fold Change	P-value	upregulated in SiR N2C (log2 Fold Change)	
Pattern Recognition Receptors (PRRs)	Rig-I	5.25	2.90E-05	no	
	Tlr9	4.56	3.47E-06	no	
	Tlr2	4.48	6.66E-07	no	
	Tlr8	4.05	1.30E-03	no	
	Tlr13	3.91	5.20E-05	no	
	Tlr4	3.84	1.33E-05	no	
	Tlr1	3.79	2.21E-04	no	
	Tlr7	2.84	2.09E-04	no	
	Tlr3	2.28	3.73E-04	no	
Downstream Signalling Pathways/Transcription Factors	Irf7	6.88	1.44E-24	yes	(5.25)
	Irf1	6.17	2.17E-09	no	
	Irf4	5.01	2.95E-05	no	
	Nfkb2	4.17	7.17E-06	no	
	Irf5	4.15	2.78E-08	no	
	Irf2	1.34	4.42E-07	no	
	Nfkb1	1.26	7.32E-05	no	

B

Antiviral Interferon Response

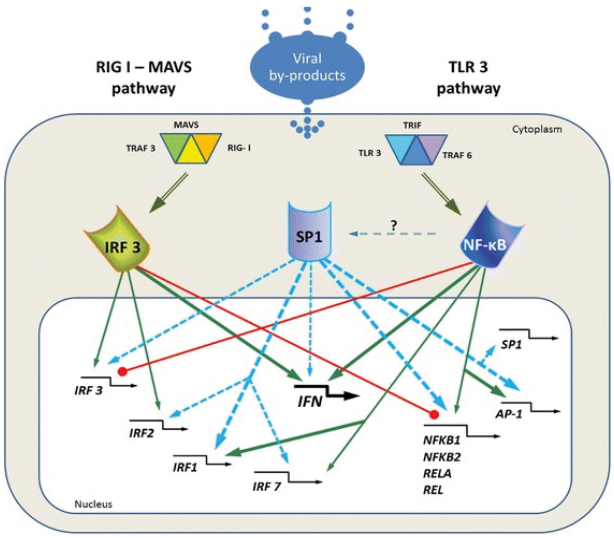
	Probe	Log2 Fold Change	P-value	upregulated in SiR N2C (log2 Fold Change)	
Pro-inflammatory Receptors	Ifngr1	2.50	7.74E-06	no	
	Il6ra	2.25	2.09E-04	no	
	Ifnar2	2.03	2.10E-06	no	
JAK/STAT signalling pathway	Stat4	5.39	1.15E-20	yes	(4.32)
	Stat2	3.78	9.16E-26	yes	(2.49)
	Jak3	3.51	7.79E-05	no	
	Stat3	2.39	4.08E-15	no	
	Jak2	1.43	3.86E-05	no	
Interferon Stimulated Genes (ISGs)	Mx1	8.27	9.81E-17	yes	(6.07)
	Oas3	7.32	4.32E-13	no	
	Oas2	6.94	6.07E-14	yes	(4.64)
	Oasl2	6.94	7.78E-14	yes	(5.45)
	Oasl1	6.74	9.80E-11	no	
	Isg15	6.67	1.76E-10	no	
	Ifit1	6.42	7.11E-27	yes	(4.92)
	Ifitm3	5.27	3.92E-18	no	
	Ifit3	5.13	1.10E-27	yes	(3.50)

C

Immune Cell Responses

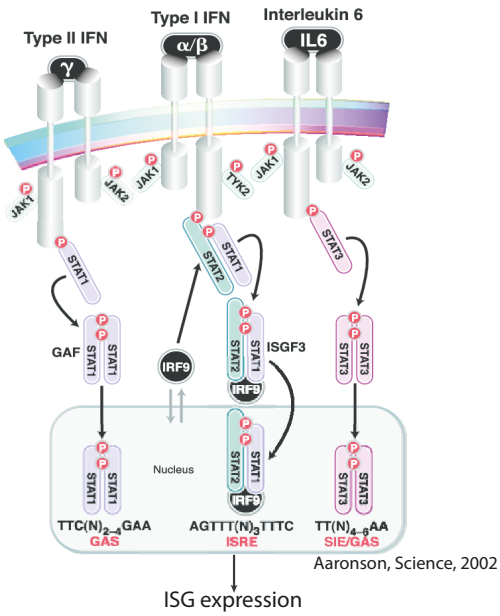
	Probe	Log2 Fold Change	P-value	upregulated in SiR N2C (log2 Fold Change)	
Cytokines	Cxcl10	8.06	5.37E-20	yes	(5.84)
	Cxcl9	7.72	2.52E-14	no	
	Ccl5	6.99	1.16E-11	no	
	Il10ra	4.97	1.86E-14	yes	(3.85)
	Il15	2.67	7.70E-05	no	
	Il17ra	1.87	1.77E-03	no	
Adhesion Molecules	Icam1	4.90	7.41E-11	no	
MHC proteins and regulators	H2-Aa	8.78	7.92E-26	no	
	H2-Q4	8.56	1.21E-18	no	
	H2-Ab1	8.52	8.37E-25	no	
	Nlrc5	7.88	1.81E-41	no	
	Ciita	7.21	5.67E-18	yes	(5.13)

D



Iwanaszko and Kimmel, 2015

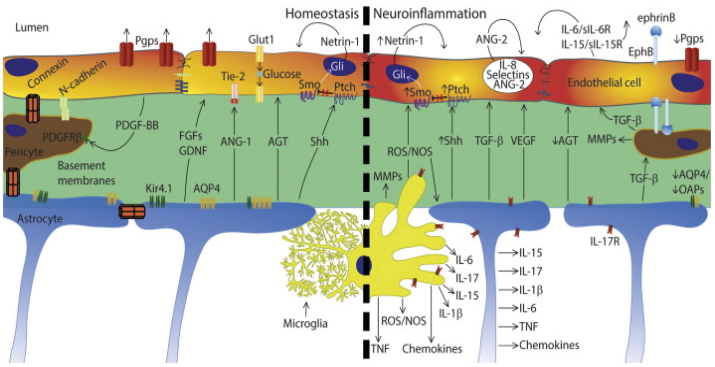
E



Aaronson, Science, 2002

ISG expression

F



Lecuyer, BBA, 2016

Fig. 3.19 Key viral immune response genes upregulated following WT N2C injection but not following SiR N2C injection Legend on following page...

Fig. 3.19 Key viral immune response genes upregulated following WT N2C injection but not following SiR N2C injection Fold change and P-values of WT N2C upregulated genes involved in viral detection **(A)**), antiviral interferon response **(B)** and immune cell responses **(C)**. Most genes are not found to be upregulated following SiR N2C injection, and those that are show lower fold change values. **(D)** Diagram showing the IRF and NF-kB transcription factor responses downstream of PRR activation (copied from Iwanszko and Kimmel. **(E)** Diagram showing pro-inflammatory receptor activation, downstream JAK/STAT pathways resulting in the expression of antiviral interferon stimulated genes (ISGs) (copied from Aaronson et al.). **(F)** Diagram showing the cytokines involved in increasing blood brain barrier permeability in neuroinflammation (copied from Lecuyer).

now many examples showing that lncRNAs can play key roles in host immune repsonses following infection by other RNA viruses. For example, the lncRNA NRAV can inhibit HIV-1 replication and regulate the expression of antiviral genes during influenza virus infections[270].

We identified many upregulated lncRNA transcripts following WT N2C infection. The top 25 upregulated lncRNAs and their fold changes are depicted in Fig. 3.20. Many of these upregulated lncRNAs overlap with previously identified lncRNAs involved in the host immune response to rabies infection and show fold changes of up to 300 fold. Although the roles of these lncRNAs in host immune responses to rabies virus are unclear, our findings are consistent with previously observed dynamic changes in lncRNA expression following rabies virus injection and suggest that they play significant biological roles in the host response. Of the top 25 upregulated lncRNA genes, only 4 were also found to be significantly upregulated following SiR N2C infection using the same DEseq2 statistical analysis (Gm1966, Gm2682, AU020206 and 4930533I22Rik). Again suggesting that the lncRNA host immune response to viral infection is attenuated in SiR infections.

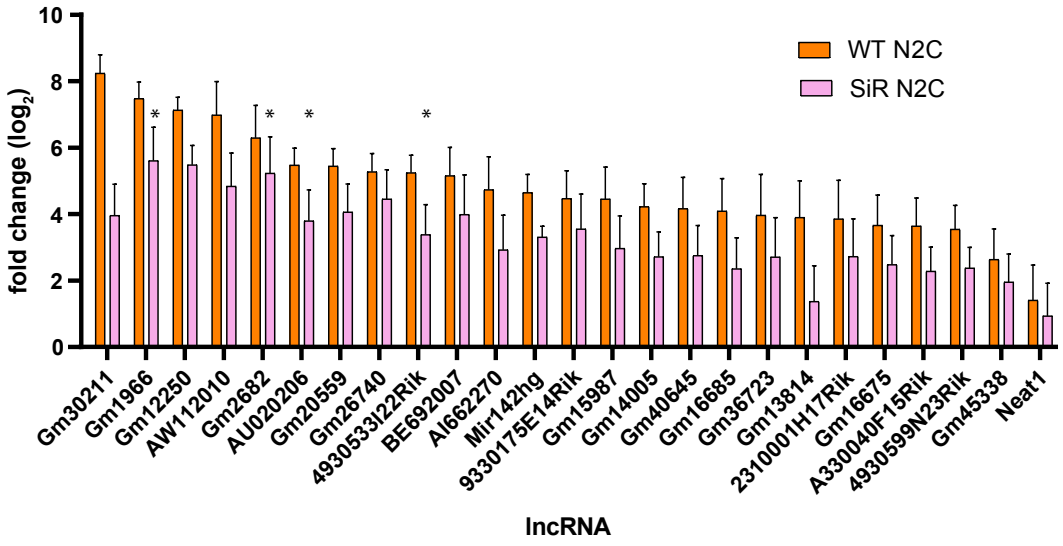


Fig. 3.20 Non-coding RNA transcripts upregulated following WT N2C infection, but not following SiR N2C infection Fold change of the top 25 upregulated non-coding RNA transcripts following viral injection (WT N2C or SiR N2C) compared to PBS injected controls. The 4 transcripts also significantly upregulated following SiR N2C injection (Gm1966, Gm2682, AU020206 and 4930533I22Rik) are labelled with asterisks.

3.3 Summary

In this chapter, I have described the development of a novel self-inactivating rabies vector which we call SiR 2.0. We show here that this vector design, derived from a combination of the viral N protein attenuation technology of SiR 1.0 and the use of a less immunogenic parent strain (CVS-N2C), can be produced *in vitro* successfully to high titers for *in vivo* use.

Given concerns regarding the genomic stability of self-inactivating rabies viruses, especially with regards to "revertant" mutation of the N-TEVs-PEST modification crucial for viral self-inactivation, we have investigated the genomic stability of the TEVs-PEST modification of the SiR at both the sequence and protein level both during viral production *in vitro* and post-infection and trans-synaptic transfer *in vivo*. These results confirm that the N-TEVs-PEST mutation is not susceptible to revertant mutation accumulation in any of the contexts studied.

We then go on to characterize further various aspects of SiR 2.0 vector function. We show that the CVS-N2C strain from which the SiR 2.0 is derived shows greater genome stability then the SAD-B19 strain. This is advantageous in the context of ensuring the quality control of high batch viral productions. In addition, the delayed transcriptional kinetics of the CVS-N2C strain results in a slower self-inactivation of the SiR 2.0 virus compared to

SiR 1.0. This could in theory be useful in certain experimental cases, where more prolonged exogenous transgene expression from the viral genome may be desired. SiR 2.0 also has a greater neuroaffinity for peripheral motor neuron terminals post muscle injection *in vivo* and a greater trans-synaptic transfer efficiency in motor-premotor circuits (enabling better circuit tracing coverage).

Further characterisation of the induction of immune responses to rabies virus infection by the different rabies virus vector variants by immunofluorescence and NGS has shown that the combination of a less immunogenic rabies strain with a TEVs-PEST modification resulting in transcriptional attenuation significantly reduces the immunogenicity of the SiR 2.0 vector compared to its predecessing rabies virus vectors. Our findings suggest that SiR 2.0 is the least cytotoxic and therefore most preferable rabies vector for long-term peripheral targeting of motor neurons.

Chapter 4

Genome-editing of Neural Circuits with Rabies Virus

4.1 Introduction

In the previous section, I have discussed the design and characterization of a novel circuit tracing tool (SiR 2.0) that allows us to genetically modify trans-synaptically connected neurons with increased neurotropic efficiency and reduced perturbations of normal circuit physiology. In these previous examples, genetic modification was conducted using the virally driven expression of CRE recombinase to activate the life-long expression of fluorescent markers in transgenic mice. I have also mentioned in the previous section, the implications of using life-long visualisation of once synaptically traced neurons for further research in synaptic longevity and plasticity. However, as also mentioned previously, anatomical visualisation of network architecture is on its own insufficient for us to decipher exactly how the brain processes information and generates behaviour. Ideally, what neuroscientists need is a vector that can allow us not only to label circuits specifically but also manipulate them in order to visualize dynamic activity and establish clear causality between circuit function and behaviour. In this instance, the rabies virus has been used as such a tool for investigations in mammals. In the case of WT ΔG rabies viruses (SAD B19 or CVS-N2C), virally driven channelrhodopsin2 (ChR2) or GCaMP6 expression has been used within the time frame before the virus causes notable toxicity (1-4 weeks) to monitor or manipulate neuronal activity[190]. For longer time scales, SiR B19 CRE infected cells were found to show effective ChR2 and GCaMP6 activity for up to 4 months p.i.[191,271].

In this chapter, we aim to combine the SiR's unique ability to trace circuits without prolonged toxicity or immunogenicity with another method of manipulating circuits - which is that of genome editing, using the CRISPR/CAS9 technology.

4.1.1 The CRISPR/CAS9 revolution

Gene therapy has been termed as "the holy grail of modern medicine"[272]. Many diseases are caused by defective genes, and genome editing techniques can in theory be used to add, replace or inactivate genes for therapeutic purposes. In order to do this however, one needs a precise method of introducing DNA strand breaks at pre-determined sites out of 3 billion possible cut sites in a patient's genome. Until recently, targeted DNA cleavage in order to achieve such goals relied on the design of specific DNA binding sites into nuclease proteins (e.g.zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs)[273]). However, enzyme dependent systems are not versatile, and re-direction of nuclease activity to different genomic sites requires time-consuming adjustments of protein-DNA binding sites. Other gene silencing methods such as RNA interference are, in contrast,

more easily directed to target genes by alteration of the RNA base sequence but often results in incomplete knock down and requires repeated delivery in order to obtain long-term gene suppression.

Recently, our ability to target specific genomic locations in eukaryotes has seen a huge advancement with the development of CRISPR-Cas9 based systems for genome editing. Clustered regularly interspaced short palindromic repeats (CRISPR) were first observed in bacterial and archaeal genomes in 1993[274,275]. Over the course of two decades, these initially mysterious sequences were revealed to be components of the bacterial adaptive immune system[276,277]. CRISPR sequences are transcribed into RNA molecules called CRISPR RNA (crRNA) that contain complementary sequences to incoming plasmid or phage DNA[278] and direct CRISPR associated proteins (Cas) to create double-stranded breaks 3 nucleotides upstream of the proto-adjacent motif (PAM) in targeted DNA[279]. In type II CRISPR systems, the only protein required for cleavage is the Cas9 endonuclease[280]. This system was shown to transfer immune resistance between different bacteria[281],be modifiable to target genomic loci of choice by changing the crRNA sequence[282], and shown to be functional in cleaving endogenous genome sequences in eukaryotic cells[7,283,284]. A single guide RNA (gRNA), which is a hybrid of the two RNAs used in natural type II systems (a crRNA and trRNA) was also designed and shown to efficiently target Cas9 endonuclease to target genomic loci[285]. The gRNA is approximately 100 bases long, has a hairpin secondary structure and contains a 21bp sequence in its 5' terminus which is complementary to the target genomic site. These developments have output a versatile CRISPR-Cas9 system that enables, by the co-expression of one CAS9 endonuclease enzyme and one gRNA, a means of targeting endonuclease activity to predetermined specific sites of the genome.

Modifications of CRISPR-Cas9 systems for gene engineering has enabled an array of different gene modifications in mammalian cells. For example, Cas9 nickase mutants cause only single-stranded nicks in targeted sequences as opposed to DSBs induced by WT enzymes[282,285,286]. Depending on the type of DNA damage resulting by Cas9 activity (single-strand nicks or double strand breaks (DSBs)) and on the type of repair system used by the host cell (homologous or non-homologous) the CRISPR-Cas9 system can be used to insert novel sequences using pre-designed DNA donor fragments for homologous recombination, or disrupt genes by causing indel mutations as a result of non-homologous end joining of double strand breaks. In addition, the fusion of catalytically inactivated Cas9 to a cytidine deaminase enzyme enables other genome modification methods such as direct targeted base editing (C->T, or G->A on the opposite strand) without the induction of dsDNA breaks[287].

Ultimately, CRISPR-Cas9 systems enable direct permanent editing of eukaryotic genomes after a short period of expression in a targeted cell or tissue, and can easily be designed to target a specific genomic sequence complementary to the 20bp targeting sequence in gRNA and adjacent to a protospacer adjacent motif (PAM). In the past two years, there have been over 325 articles published on the modification and application of such CRISPR-Cas9 systems for basic research and potential therapeutic purposes.

4.1.2 Suitability of SiR viruses for neurological CRISPR-Cas9 gene therapy

Successful gene therapy requires not only a efficacious molecular system to conduct the gene editing, but also a means to deliver these gene editing tools efficiently into target cells that are implicated in disease pathogenesis. Another main barrier to successful gene therapy development is finding a safe and efficient vector system to transfer gene modification tools into cells, tissues, or in our case, neural circuits, of choice.

For the treatment of neuropsychiatric and neurodegenerative conditions, the advantages of using a circuit specific vector to conduct gene modifications includes both the ability to increase the payload of gene therapy to the widespread circuits that are affected in pathogenesis whilst also reducing off-target side effects resulting from transduction of untargeted cells not involved in disease pathogenesis. As discussed in the introduction, this is extremely difficult to achieve using non-circuit specific vectors. Non circuit-specific vectors (e.g. AAVs or lentiviruses) can only target diseased neural circuits either by systemic or widespread CNS delivery which increases off-target anatomical side effects, or more anatomically specific invasive surgical techniques which do not have coverage of widespread networks implicated in disease. In contrast, the rabies virus can, following peripheral injection, trans-synaptically retrogradely transport itself to anatomically distant but connected neurons in appropriate areas of the central nervous system with high efficiency.

In addition to anatomical specificity, one also needs to consider gene-editing specificity. Despite its unprecedented specificity and versatility, CRISPR-Cas9 systems can still have off-target effects[288], and the side-effects of long term expression of endonucleases in cells are unclear. Short-term expression of CRISPR-Cas9 can cleave targeted sites with high efficiency whilst reducing the accumulation of off-target gene modifications. This is a consideration for DNA vectors such as AAVs or lentiviruses which can permanently integrate into the host genome, increasing the risk of insertional mutagenesis but also resulting in more off-target modifications due to long-term expression of gRNAs and endonucleases. Therefore, ideally

the CRISPR-Cas9 should be delivered to target cells for a short period long enough to achieve efficient targeting of the desired loci, but minimal off-target effects. And in this instance, the SiR, which expresses transgenes for a short period of 2-3 weeks before self-inactivation may overcome the problems associated with long term Cas9 gRNA expression.

4.1.3 Aims

In this chapter of the thesis, my aim is to functionally incorporate into the rabies genome the two components of the CRISPR/Cas9 system (the CAS9 endonuclease, and 100bp gRNA). This combines the rabies virus' unique ability to trans-synaptically infect neurons with our unprecedented ability to precisely edit the genome of mammalian cells using CRISPR-Cas9 systems. Such a tool would finally allow researchers to probe the role of specific genes within circuits in the computation of behaviours in live animals, and pave the way for the development of effective circuit-targeted gene therapy for neuropsychiatric and neurodegenerative disorders.

Here, we describe the design, screening and production of rabies viruses expressing both components of CRISPR-Cas9 system to induce targeted DSBs in mouse genomes both *in vitro* and *in vivo*.

4.2 Results

4.2.1 Designing gRNA to target the mouse NeuN gene

In order to test for rabies virus mediated genome editing, we decided to target the NeuN locus. NeuN (also known as Fox-3 or Rbfox3) is a RNA binding protein which is commonly used as a nuclear marker for neurons in vertebrates as it is expressed at a high level in almost all neuronal populations (exceptions include purkinje cells and photoreceptors)[289]. We designed 3 sgRNAs against the first exon of the NeuN gene using the CRISPOR (http://crispor.org)[290] and GPP sgRNA Designer webtool (https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design)[291]. We selected for gRNA sequences with high predicted efficiency and specificity (i.e. high on-target but low off-target activity). The target sequences of these 3 gRNAs with their adjacent PAM sequences are shown in Fig. 4.1A.

We experimentally validated the targeting of the endogenous mouse NeuN locus *in vitro* by these gRNA sequences using a mismatch cleavage assay (Fig. 4.1A). In order to do this, we transfected mouse neuro2a cells with DNA plasmids encoding the Cas9 endonuclease enzyme under a type II CMV promoter and the single gRNA sequence using a type III U6 promoter. 3 days post-transfection, we extracted genomic DNA from infected cells for amplification of the first exon of the NeuN gene (Fig. 4.1B). Any indel mutations introduced by NHEJ repair of gRNA targetted DSBs would also be amplified in the PCR products. Following PCR amplification, PCR products were denatured-renatured, resulting in the annealing of mismatched DNA fragments which can be detected by "resolvase" enzymes, such as the T7 endonuclease enzyme which cleaves DNA heteroduplexes[292]. The resulting reaction mixture was then run on an agarose gel to identify the site and degree of fragmentation. We found that gRNA NeuN3 showed the highest cleavage activity (Fig. 4.1C) and so we used this sequence for further expression from the rabies virus genome.

4.2.2 Preliminary design of a WT ΔG gRNA NeuN Cas9 rabies virus: the molecular challenges of expressing gRNA and Cas9 endonuclease from the rabies genome

Having designed and experimentally validated a gRNA sequence targeting the endogenous NeuN gene (now called gRNA NeuN), we wanted to engineer the rabies virus genome to functionally express both the Cas9 endonuclease and the gRNA NeuN sequence. In order to do this, we had to consider and adjust for various molecular challenges.

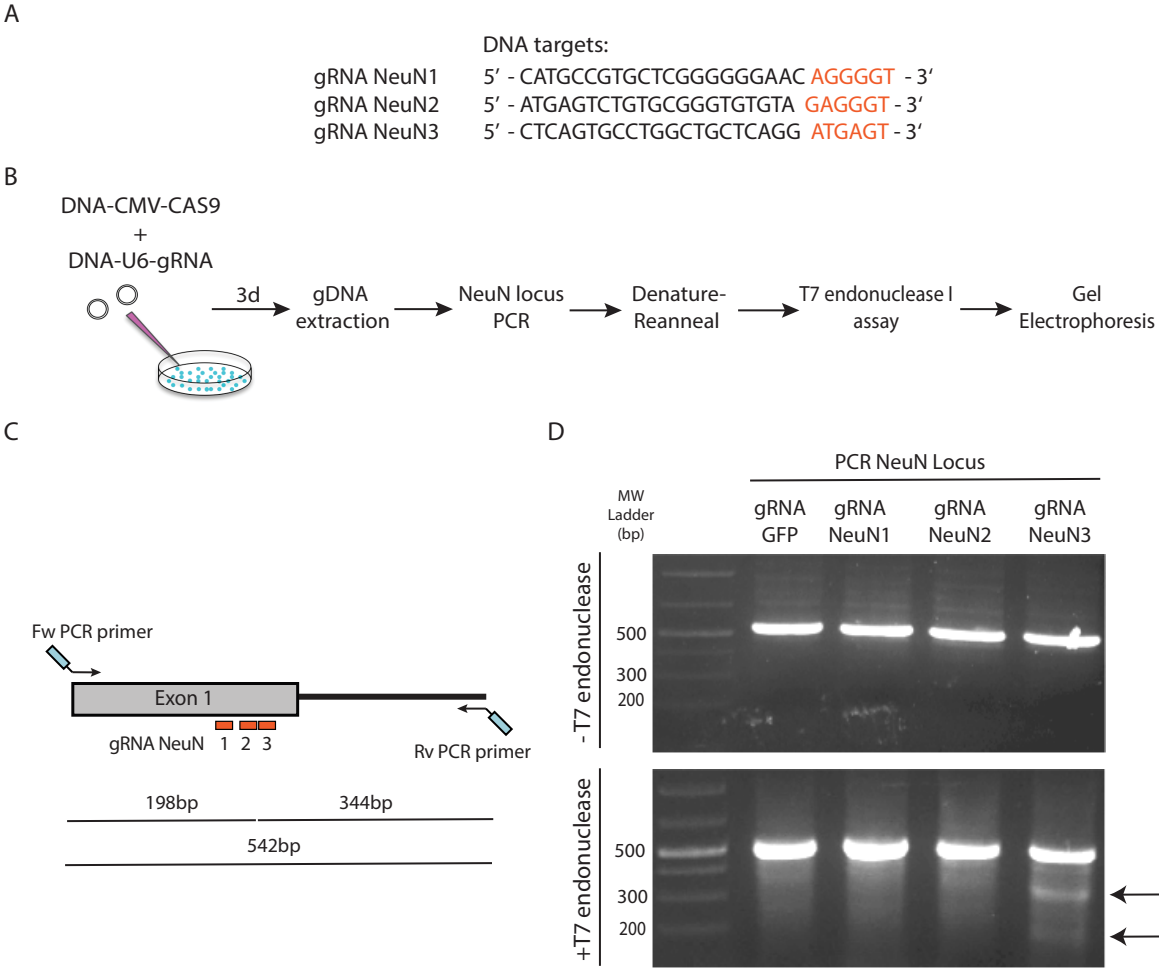


Fig. 4.1 Design and screening of gRNA to target the NeuN locus using the T7 endonuclease I assay (A) 3 gRNAs were designed to target the mouse NeuN locus. Their target sequences (black) and adjacent PAM motifs (red) are shown. **(B)** Mouse neuro2A cells were transfected with DNA plasmids to drive CMV-Cas9 and U6-gRNA NeuN expression. 3 days post-transfection genomic DNA was extracted from cells, the exon 1 of the NeuN locus amplified by PCR, and mutations resulting from effective Cas9 targeting were detected using the T7 endonuclease I assay. **(C)** PCR primers were designed to amplify a region of 542bp length spanning exon 1 and a portion of the subsequent 3' intergenic region. As gRNAs were designed to target near the 3' end of the exon, indel detection would result in fragemntation of the PCR amplicon to give two fragments of approximately 200 and 340bp in size. Exact numbers are given for cleavage caused by gRNA NeuN3 targeting. **(D)** Gel electrophoresis of NeuN PCR amplicon fragmentation by T7 endonuclease I resulting from mismatch cleavage. One can see only gRNA NeuN3 causes siginficant targeting of Cas9 to the NeuN locus.

Our first consideration was the limited coding capacity of viral vectors (Fig. 4.2). The native WT SAD-B19 genome is 11.9Kb[293]. We have previously replaced the G locus (1.5kb) with a CRE-GFP transgene (2kb) without affecting viral recovery from plasmids or high-titer production. In addition, recombinant ΔG rabies viruses have been recovered from plasmids that are up to 13.6kb in size (SADΔG-GFP-ERT2CreERT2) which is 1.7kb larger than the native SAD-B19 genome[294]. Here, we wanted to add to the viral genome sequences for the expression of both the Cas9 endonuclease and gRNA for genome editing. The most commonly used Cas9 endonuclease (that from *Streptococcus pyogenes*) alone is 4.2kb in size resulting already in a 2.7kb payload over the WT virus. Therefore we used a 1kb shorter variant of the Cas9 endonuclease enzyme from *Staphylococcus aureus* (SaCas9) which has been previously used for efficient packaging into other viral vectors with limited coding capacity such as AAVs. We expected this 2kb payload in genome size may potentially impair the efficiency of viral production, but given that other transgenes of similar sizes had been inserted successfully into the rabies virus before, that with altered cell culture conditions we would still be able to produce high titer productions of recombinant rabies viruses with a genome of that size.

Second, there was the structural incorporation of gRNA sequence into the viral genome. The sgRNA is a designed hybrid of native crRNA:tracrRNA duplexes in natural type II cas9 systems with conserved secondary structures which are necessary for effective Cas9 targeting. The bulge and nexus structures are necessary for DNA cleavage, whilst the nexus and hairpin structure dictates Cas9 orthoganality[295]. Structurally, the nexus structure positions both the Cas protein and RNA to receive target DNA duplexes for cleavage. Given the high propensity of sgRNA for secondary structure formation, we were concerned that incorporating its primary sequence into the single stranded RNA genome of the rabies virus may lead to the formation of secondary structures within the viral genome which could impair viral packaging, replication or transcription mechanisms.

Lastly, in addition to the structural consequences of gRNA incorporation into the ssRNA genome, we needed to ensure functional expression of sgRNA. This was arguably one of our major considerations as efficient Cas9 targeting requires high levels of gRNA expression and precise start and stop regulation of transcription. The vast majority of literature regarding CRISPR technology is based either on direct delivery of CAS9 mRNA together with pre-transcribed gRNAs, preassembled ribonucleocomplexes or DNA-mediated expression of CAS9 and gRNA in the nucleus of the cell using DNA viruses or plasmids. These DNA vectors express the Cas9 protein under polymerase II promoters to recruit cellular RNA pol II machinery for the production of mRNA, whilst gRNA is expressed under type III promoters

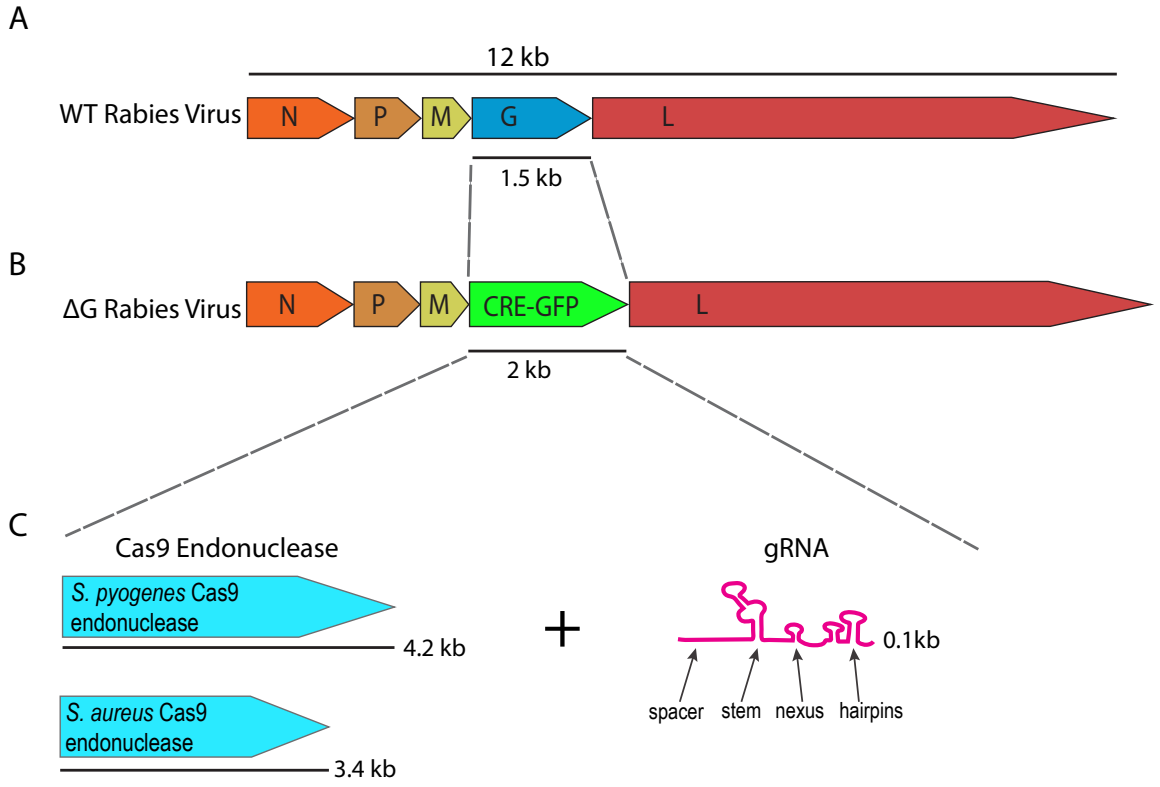


Fig. 4.2 Molecular challenges of expressing CAS9 endonuclease and gRNA from the rabies genome (A) The WT SAD B19 rabies virus is 11.9kb. **(B)** Transgene expressing ΔG rabies viruses used in this thesis have shown that viral productions and infectivity are not impaired by additions of sequences resulting in genetic payloads of 0.5kb. **(C)** Here we try to incorporate into the ΔG rabies virus a Cas9 endonuclease and gRNA coding sequence. Commonly used S. pyogenes Cas9 endonuclease is 4.2kb, but there are smaller alternatives available such as the S. aureus Cas9. gRNA sequences are approximately 100bp and can form complex secondary structures which may interfere with viral genome structure and hence viral function.

(e.g. U6) recognized by the cellular RNA polymerase III machinery which naturally produce small RNAs. However, the rabies virus is a ssRNA virus with no DNA intermediates that resides exclusively in the cytosol and uses its own viral transcription machinery. As such, it is not possible to use reported methods of nuclear DNA driven Cas9 gRNA expression in virally infected cells.

In order to express exogenous Cas9 endonuclease and gRNA from the rabies virus genome, we had to consider existing knowledge about rabies virus transcriptional biology. The rabies virus is a negative strand RNA virus. The ssRNA genome is encapsulated by the viral nucleoprotein (N) to act as a template for transcription and replication by the viral L-P polymerase complex. For transcription, the L-P polymerase complex sequentially produces mRNAs from the 3' to 5' end by recognition of viral transcription start and stop sites to initiate synthesis and polyadenylation respectively. At stop sites, the L-P complexes dissociate from the genome but a proportion are able to re-initiate transcription at the subsequent TSS of the viral genome. Transgenes such as CRE and GFP have been expressed exogenously from the rabies virus before by replacing cell endogenous pol II promoters with viral transcriptional start stop sequences. However, the expression of small RNAs using rabies virus transcriptional start and stop sequences and the viral L-P RNA polymerase was a novel method that was previously unreported in the literature. As such, it was unclear whether the viral transcriptional machinery was capable of producing functional gRNAs. We needed to ensure the virus would be able to process gRNAs functionally and express it to high enough levels.

Taking all these things into consideration, we made a preliminary first design of a WT ΔG SAD B19 rabies virus expressing the *Staph aureus* Cas9 (SaCas9) endonuclease and gRNA NeuN in the G locus (Fig. 4.3A) in order to test whether a rabies virus genome of that size, incorporating a hairpin gRNA sequence could be recovered from plasmids and produced at titers appropriate for *in vivo* injections. In addition to successful viral production, we also tested whether these viruses were functionally expressing both the Cas9 endonuclease and gRNA elements in the cytosol using the viral transcriptional machinery via viral transcriptional start and stop sites

4.2.3 Successful production of WT ΔG Rabies gRNA NeuN SaCas9-3xFLAG, design A

Our preliminary design of a WT B19 ΔG gRNA NeuN Cas9 virus (Design A, Fig. 4.3A) consisted of a WT B19 ΔG viral vector with a *Staph aureus* Cas9 (SaCas9) and gRNA

sequence inserted into the M/L locus of the WT virus with unedited standard rabies virus transcriptional start and stop sites for viral transcription regulation. Due to the already large size of this construct, we did not incorporate any fluorescent proteins for visualisation of viral productions in this design. For eventual visualisation of viral proteins we did add a 3x tandem repeat of the FLAG epitope to the C terminus of the SaCas9 coding sequence. But as FLAG visualisation requires antibody staining, we were unable to use this marker live to follow the production of this virus in real time. As such, productions were conducted "blind" with absence of a fluorescent marker to monitor viral packaging, replication and spreading.

As a positive control for viral production steps, we would make in conjunction with the "blind" rabies virus, a WT B19 ΔG CRE-mCHERRY with the same experimental conditions to assess the success of plasmid transfections. We also were able to monitor "blind" viral production by observing "rounding" of production cells which are indicative of successful rabies virus replication in cells *in vitro*. Once we obtained potentially infective supernatant from blind production cells, we infected mouse neuro2A cells with the supernatant to test first, whether the WT B19 ΔG gRNA NeuN Cas9 virus could be produced, and second, whether this virus retained its ability to infect cells *in vitro*. 3 days post infection, we stained the mouse neuro2A cells with an anti-FLAG primary antibody which we had conjugated to a Cy3 antibody for quick visualization of viral Cas9 expression. We found that addition of 250μl of viral supernatant to a 6 well of neuro2A cells resulted in FLAG+ staining of 70% of cells. This was a high level of infectivity comparable to the titer of our positive control (WT B19 ΔG CRE-mCHERRY). As such, this finding confirmed that the rabies virus could tolerate the addition of Cas9 endonuclease and hairpin gRNA sequences into its viral genome without impairing viral replication or *in vitro* infectivity. Positive FLAG staining also confirmed that the SaCas9-3xFLAG protein was being transcribed effectively by the rabies virus.

4.2.4 WT ΔG Rabies gRNA NeuN SaCas9-3xFLAG can target endogenous loci in vitro

Through FLAG Ab staining we had confirmed that a rabies virus encoding both a Cas9 endonuclease and gRNA sequence could be produced, used to infect cells *in vitro* and express the SaCas9-3xFLAG protein. However, we needed also to assess whether the gRNA was being transcribed effectively by the rabies virus L-P RNA polymerase complex to enable efficient genome targeting. In order to do this, we conducted a mismatch cleavage assay in the NeuN locus in mouse neuro2A cells infected with WT ΔG gRNA NeuN CAS9 viral

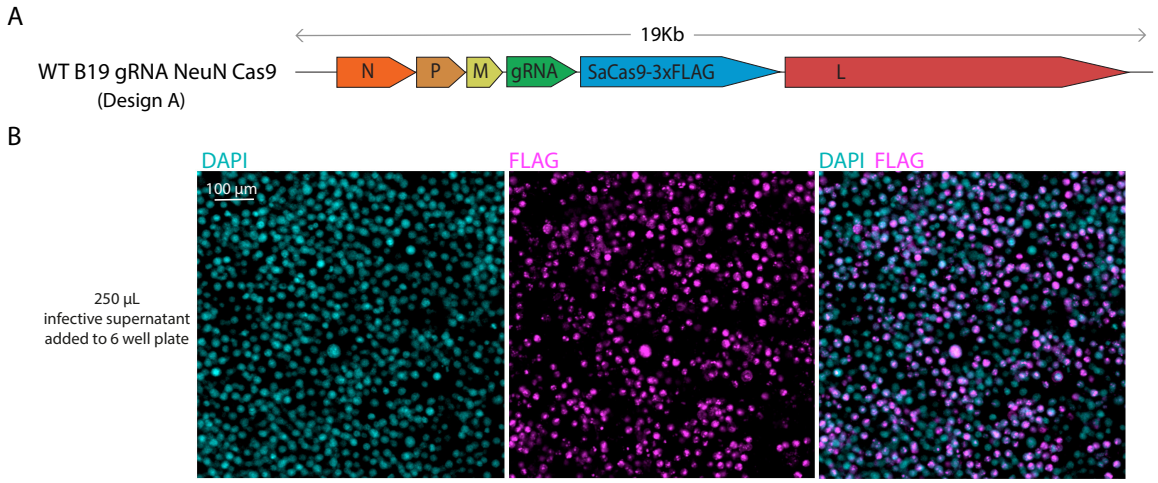


Fig. 4.3 Successful production of WT ΔG Rabies gRNA NeuN SaCas9-3xFLAG, design A (A) Genome map of our preliminary first design of a WT ΔG Rabies virus encoding gRNA and SaCas9-3xFLAG. **(B)** Staining of mouse neuro2A cells infected with 250μl of viral supernatant in 6 well plates with a Cy3 conjugated anti-FLAG antibody (magenta). Cell nuclei were stained with DAPI (blue). Quantification of FLAG+/DAPI+ cells revealed 70% infected cells showing successful viral production, *in vitro* infectivity and SaCas9-3xFLAG protein expression.

supernatant. We first titered the viral supernatant to assess the number of infected neuro2A cells following viral infection. We found the supernatant had an infectivity of 1.3x10^8 i.u./mL, with 50% of cells being infected following addition of 25μl to 80% confluent neuro2A cells in 6 well plates (Fig. 4.4A,B). We then assessed the degree of indel mutation accumulation in the targeted NeuN locus over time using a mismatch cleavage assay in these cells infected with increasing viral titers at 3d and 6d p.i. (Fig. 4.4C,D).

First, we found that the WT B19 gRNA NeuN Cas9 rabies virus could indeed target the endogenous NeuN locus *in vitro*. This was evident by the observation of cleaved NeuN exon 1 PCR fragments in infected cells both 3d and 6d p.i (Fig. 4.4C,D). This indicated that the gRNA sequence that is usually expressed by RNA pol III systems in the nucleus for small RNA transcription could also be transcribed and processed functionally by the viral L-P RNA polymerase complex. It also indicated that even with this preliminary design, the rabies virus could produce enough Cas9 and gRNA to induce clear genome editing *in vitro* as visualised by the T7 endonuclease I assay.

Second, we found that there was increased mutation detection in the NeuN locus at 6d p.i. compared to 3d p.i. in cells infected with the same volumes of viral supernatant (Fig. 4.4C). This showed us that although genome targeting is visible from as early as 3d p.i., the coverage of gene modifications in infected cells can increase with increasing time of up to 6 days p.i.. Extrapolating from this result *in vitro*, we expected that as the period of

transcriptional activity in SiR viruses is 2 weeks, the length of Cas9 and gRNA expression from SiR vectors would not be a limiting factor for efficient genome editing *in vivo*.

Third we found that with increased % of infected cells, we observed more mismatch cleavage. For example, with 25μl viral supernatant and hence 50% infected cells, we saw fragmentation of 5% of PCR amplicons 6 days p.i. whereas with infection of 60% of cells with 100μl viral supernatant we observed fragmentation of 16% of PCR amplicons. This was partly expected as with higher % of infected cells, more of the cells collected for mutation analysis were expressing the viral gRNA-Cas9 cassette. However, between these two dilutions the percentage of cells infected had only gone up by 25%, whilst the fragmentation efficiency had almost tripled. The degree of mutation detection also increased when comparing cells infected with 100 vs 200μl of viral supernatant even though visualisation of infected cells using the FLAG antibody did not reveal significant differences in the number of infected cells (both resulting in the infection of 61% of cells on average). Despite the similar numbers of infected cells, cells infected with higher viral supernatant volumes showed more efficient genome targeting. We hypothesised that this may be due to the higher multiplicity of infection (m.o.i.) associated with infections with more viral supernatant, resulting in more viral particles infecting each FLAG+ cell. For example, if we assume the poisson distribution of viral particles infecting cells in culture and a derive a m.o.i. of 11 vs. 22 in cells infected with 100 vs. 200μl of viral supernatant from our titration analysis, we find that the percentage of cells with multiple infection (i.e. infection with more then 1 viral particle) was 1.5x10^{-}2% vs. 3.3x10^{-}7. It could be that cells infected with more viral particles express higher levels of Cas9 and gRNA resulting in more efficient genome targeting.

Given our results, and our assumptions based on previous literature regarding the requirement for stringent and high gRNA expression for efficient CRISPR-cas9 mediated genome editing, we aimed to investigate directly whether the efficiency of our system could be further improved by improving gRNA expression.

4.2.5 Endogenous gene cleavage efficiency can be boosted with increased gRNA expression

In order to investigate whether the efficiency of genome editing using the rabies virus vector could be increased further by increasing gRNA expression, we compared the % cleavage of NeuN PCR amplicons following the infection of mouse neuro2A cells with WT rabies alone or WT rabies + plasmid driven U6-gRNA expression at 2 different viral titers. Unedited neuro2A cells were used as a negative control, whilst cells transfected with both CMV-Cas9

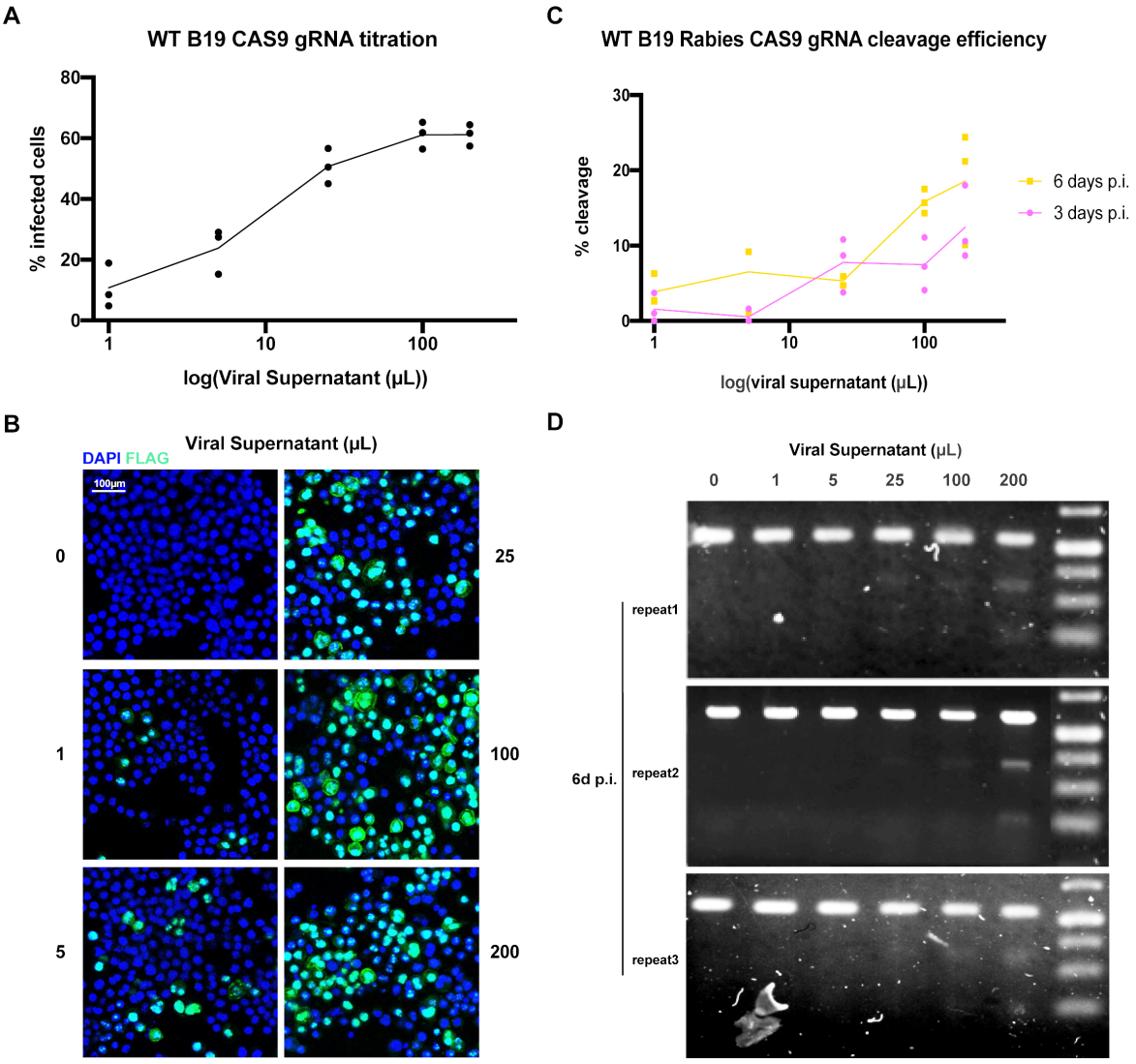


Fig. 4.4 WT B19 Rabies CAS9 gRNA, Design A can target endogenous genes *in vitro* (A) % of FLAG+/DAPI+ cells following infection with increasing volumes of viral supernatant **(B)** Visualisation of FLAG expression (green) by virally infected cells following infection with increasing volumes of viral supernatant. **(C)** % of NeuN locus PCR amplicon fragmentation from the T7 endonuclease I assay at 3d and 6d p.i. of cells infected with increasing volumes of viral supernatant. **(D)** Visualisation of 3 experimental repeats of T7 endonuclease assay gel visualisation at 6d p.i..

and U6-gRNA plasmids were used as a positive control to give an estimate of high efficiency genome editing.

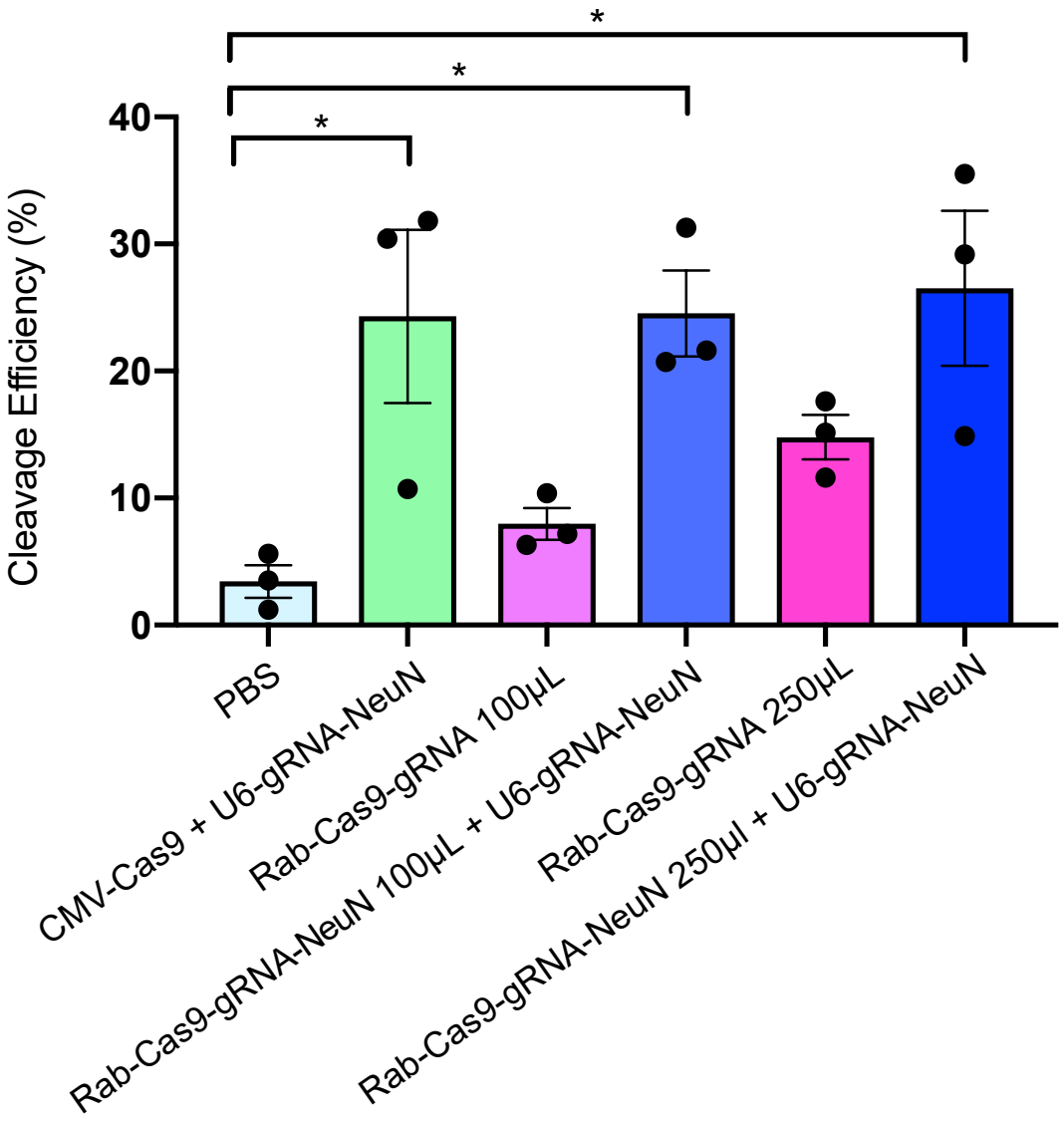


Fig. 4.5 Genome-targeting efficiency of WT ΔG B19 Rabies gRNA Cas9 can be enhanced by higher gRNA expression Graph showing the % of NeuN exon 1 PCR amplicon cleavage from the T7 endonuclease I assay following infection of cells with rabies virus with or without additional U6-gRNA DNA plasmid transfection. PBS was used as the negative control for lack of genome targeting, and co-transfection of CMV-Cas9 and U6-gRNA was used as a positive control for efficient genome targeting. n=3, One-way ANOVA with Tukey correction for multiple comparisons.

We found that at both titers investigated, co-transfection of rabies infected cells with plasmid U6-gRNA DNA resulted in increased cleavage efficiency to levels similar to the

positive control (Fig. 4.5). This suggested that the limiting factor for cleavage efficiency in rabies only infected cells was due to gRNA expression rather then SaCas9 expression efficiency.

In addition, following co-transfection with U6-gRNA, we did not observe differences in cleavage efficiency in cells infected with 100 vs 200 μl viral supernatant suggesting that the differences in genome targeting efficiency observed previously between cell populations with similar numbers of infected cells were due to increased gRNA expression at higher moi.

With this result we decided that in order to increase the efficiency of rabies mediated genome editing, we needed to boost gRNA expression levels. Thus, we subsequently designed, using our limited knowledge about transcriptional regulation of *Rhabdoviruses* (mostly coming from vesicular stomatitis virus), multiple cassettes for the increased functional expression of the CAS9-gRNA elements.

4.2.6 Designs to increase gRNA production from the rabies virus vector

In order to increase gRNA expression levels from the rabies virus genome, we designed 4 variants of WT B19 Cas9 gRNA NeuN, taking into consideration three known factors regarding rabies viral transcription mechanisms. First that transcript levels are known to decrease with increasing distance from the 3' leader sequence of the viral genome (Fig. 4.6). Second that intergenic region lengths and sequence can also heavily influence transgene expression levels[296]. And third, that the addition of exogenous sequences to the rabies virus at different loci can result in unpredictable alterations of transcription levels of subsequent genes in the viral genome[297].

4.2.6.1 Decreasing transcript levels with increasing distance from 3' end in genome

Our knowledge of rabies virus transcription regulation is derived from investigations into the transcriptional regulation of other non-segmented negative strand RNA viruses of the same *Mononegavirales* order or *Rhabdoviridae* family. Study into the vesicular stomatitis virus (a member of the *Rhabdoviridae* family) revealed that a major determinant of transcription regulation in negative ssRNA viruses is the gene order. It is assumed that the viral RNA polymerase initiates transcription from the 3' leader sequence and subsequently transcribes genes towards the 5' direction[298]. At the border of each gene, there is a transcription stop signal for polyadenylation of mRNA transcripts. The viral polymerase dissociates at each gene border and as only a proportion of dissociated polymerase complexes re-initaite

transcription at the subsequent gene there is a progressive loss of mRNA trnscript levels that follows the order of genes from 3' to 5' end[299]. As such, the 5' terminal gene (L) produces the least mRNA transcripts in infected cells[300,301].

We observed the same finding in rabies virus infected neurons *in vivo* following injection of rabies virus into the hippocampus (Fig. 4.6). 1 week p.i., total RNA from infected hippocampi were prepared for RNA-sequencing. Reads were aligned to the viral genome and the density of reads to each gene was compared. We assumed that these reads obtained from the total RNA preparation were derived from both from viral genome RNA as well as virally transcribed gene mRNA molecules. Comparison of read density revealed that there was significantly lower levels of L encoding RNA at the most 5' end of the viral genome in infected hippocampi compared to N encoding RNA at the most 3' end. As such we directly validated from rabies virus infected cells *in vivo*, the known transcriptional regulation determinants of other members of the *Rhabdoviridae* family. Given these results, we decided in our subsequent WT ΔG gRNA NeuN Cas9 designs to alter the locus of exogenous sequence entry to increase gRNA transcription levels.

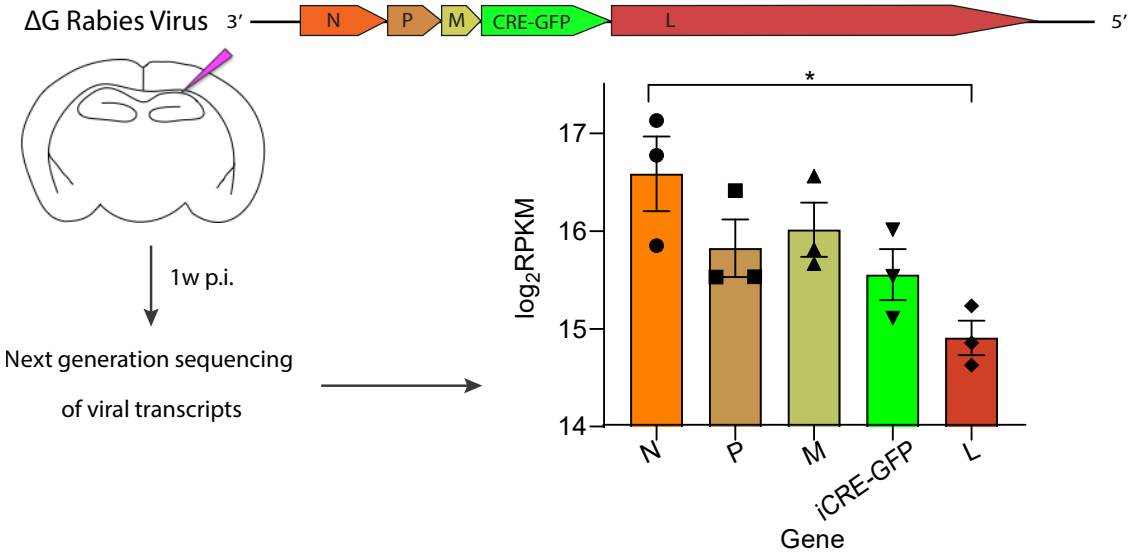


Fig. 4.6 Viral mRNA transcription levels decrease with increasing distance from 3' leader promoter Adult mice were injected with WT B19 CRE-nuGFP virus in CA1 of the hippocampus. 1 week p.i. total RNA was extracted from dissected hippocampi and prepared for Illumina RNA-sequencing. Reads were then aligned to the viral genome and analysed for read density (log2RPKM) for each of the 5 viral genes.

4.2.6.2 Design of 4 Rabies-CAS9-gRNA variants

In order to try to increase gRNA expression from the rabies genome, we designed 4 variants of a WT B19 Δ G virus to produce and screen for genome editing efficiency (Fig. 4.7).

Design A (Fig. 4.7) was our first preliminary design which was used in the previous studies as a proof of principle that the rabies virus could structurally and functionally incorporate gRNA and Cas9 sequences into its genome without impairment of replication, production or infectivity. It also revealed that this rabies vector could successfully target endogenous loci *in vitro*, but that the efficiency of this could be increased by increasing gRNA expression. In this preliminary design, we inserted the gRNA NeuN and SaCas9-3xFLAG sequence into the G locus with unedited M/G intergenic sequences and viral transcription start and stop sites.

In design B, we attempted to boost gRNA expression by altering the intergenic regions (IGRs) between genes. This was due to a previous study showing that the 2nt N/P intergenic sequence yielded the highest amounts of subsequent transgene mRNA transcription compared to all other intergenic sequences. If N/P IGR was considered to result in subsequent luciferase transgene mRNA levels of 100%, the IGRs of the P/M, M/G and G/L locus resulted in 78, 81 and 11% mRNA transcription respectively[296]. As such, we designed a second variant (design B) of the original WT B19 gRNA NeuN Cas9 design with the 5 nt M/G IGR sequence replaced with a 2 nt N/P IGR sequence.

In design C, given our knowledge of increasing transcript levels in the 3' end of the viral genome, we moved our gRNA encoding sequence to the N/P locus (Fig. 4.6). In design D, we incorporated 2 gRNA sequences into two loci of the viral genome (N/P and M/L locus). In theory, if these two loci were used to express the same gRNA we thought this may further increase gRNA expression levels and help increase genome targeting effiency. In addition, we considered that in the future, the expression of 2 gRNAs with a mutant nickase Cas9 gRNA could result in DSB induction with less off-target effects. Or alternatively, the use of 2 different gRNAs with different target genes and a WT Cas9 could result in the targeting of more than 1 gene with one rabies virus vector.

4.2.6.3 Production of the 4 variants

Having designed and cloned these 4 WT ΔG B19 gRNA NeuN Cas9 viruses into pSAD-B19 ΔG genome vectors, we attempted to produce these viruses to see if they could all be produced to high titers using standard WT ΔG B19 rabies virus production protocols (Methods). Similar to the viral production of version A described previously, all these

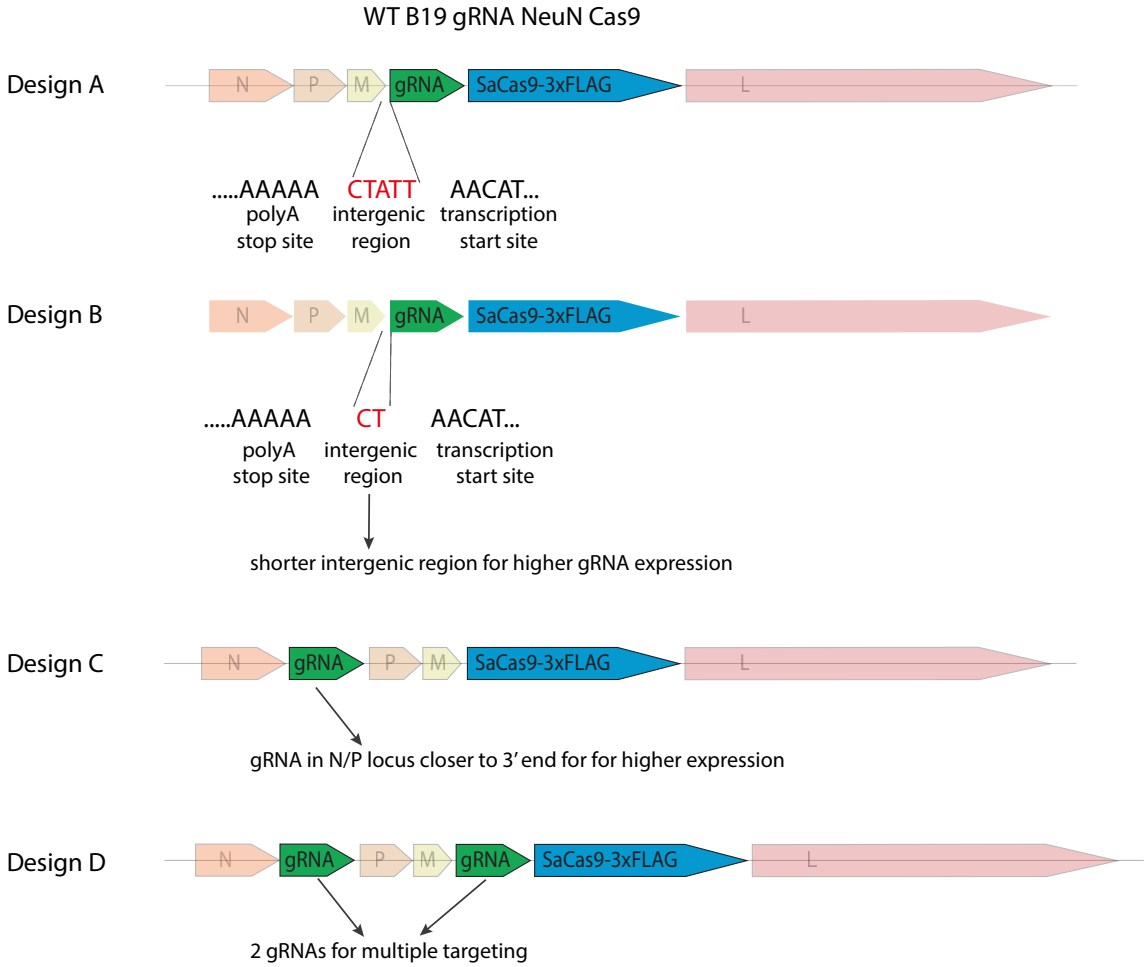


Fig. 4.7 4 designs of WT gRNA NeuN Cas9 for increased gRNA expression Preliminary design A with insertion of Cas9 and gRNA sequences into the G locus with unedited intergenic regions (IGR) and known viral trancsription start and stop sequences. Design B with replacement of the M/G IGR with the N/P IGR to increase gRNA expression. Design C with gRNA sequence in the N/P locus, closer to the 3' end of the viral genome with known higher levels of transcription. Design D with 2 gRNA sequences in the N/P and M/G locus. This design could enable the expression of the same gRNA from two different loci for increased overall expression. The two loci could also be used to express different gRNAs for reduced off-target effects using mutant Cas9 nickase enzymes, or to target two different genes using WT Cas9 endonucleases with one rabies virus vector

productions were conducted "blind" due to restricted viral genome capacity. We did not incorporate fluorescent proteins into the viral genome for visualisation of viral replication and spreading during production *in vitro*. Thus, a WT ΔG B19 mCHERRY-CRE rabies virus was produced in parallel with these "blind" viruses as a positive control for successful plasmid transfection. In addition, the anatomical "rounding" of production cells during viral supernatant collection was another live indicative marker of successful viral production and spreading *in vitro*. Following the collection of viral supernatant from these productions, we infected mouse neuro2A cells in 6 well plates with increasing volumes of viral supernatant for confirmation of viral production and titration. 3 days p.i., cells were fixed and stained for expression of the FLAG epitope on the C terminus of the SaCas9 protein to visualise virally infected cells. We found that version A and B, with modifications only to the G locus could be produced to high titers easily with standard production protocols (Fig. 4.8A,B).

However, versions C and D, with additional gRNA sequences in the N/P locus were not produced with high efficiency and we could only visualise very few virally infected neuro2A cells following infection compared to versions A and B, as well as the WT mCHERRY-CRE rabies positive control (Fig. 4.8C,D). This suggested to us that addition of the gRNA sequence to the N/P locus probably resulted in the disruption of viral transcription or replication mechanisms, which delayed or impaired viral packaging, replication or spreading kinetics *in vitro*. This was in contrast to previous studies showing that the insertion of the GFP coding sequence into the N/P locus of a different rabies virus strain (HEP-Flury) was well tolerated and resulted in higher GFP mRNA transcript levels[297]. However, our different results may be a result of the different strain of rabies virus used, as well as the differences in the size, structure and base composition of the GFP coding sequence vs. gRNA sequence.

Given the difficulty of production of versions C and D, we decided not to go forward with these designs for SiR productions. This was because SiR productions show delayed kinetics compared to WT ΔG productions due to attenuation of the viral N. Therefore, a blind production of an inefficiently produced WT design in a SiR vector would have posed many technical difficulties. Given that designs A and B were produced equally efficiently *in vitro*, and that previous studies show higher transcription levels of viral genes preceded by 2nt N/P IGRs[296,297], we moved forward with design B for future studies with SiR modified viruses *in vivo*.

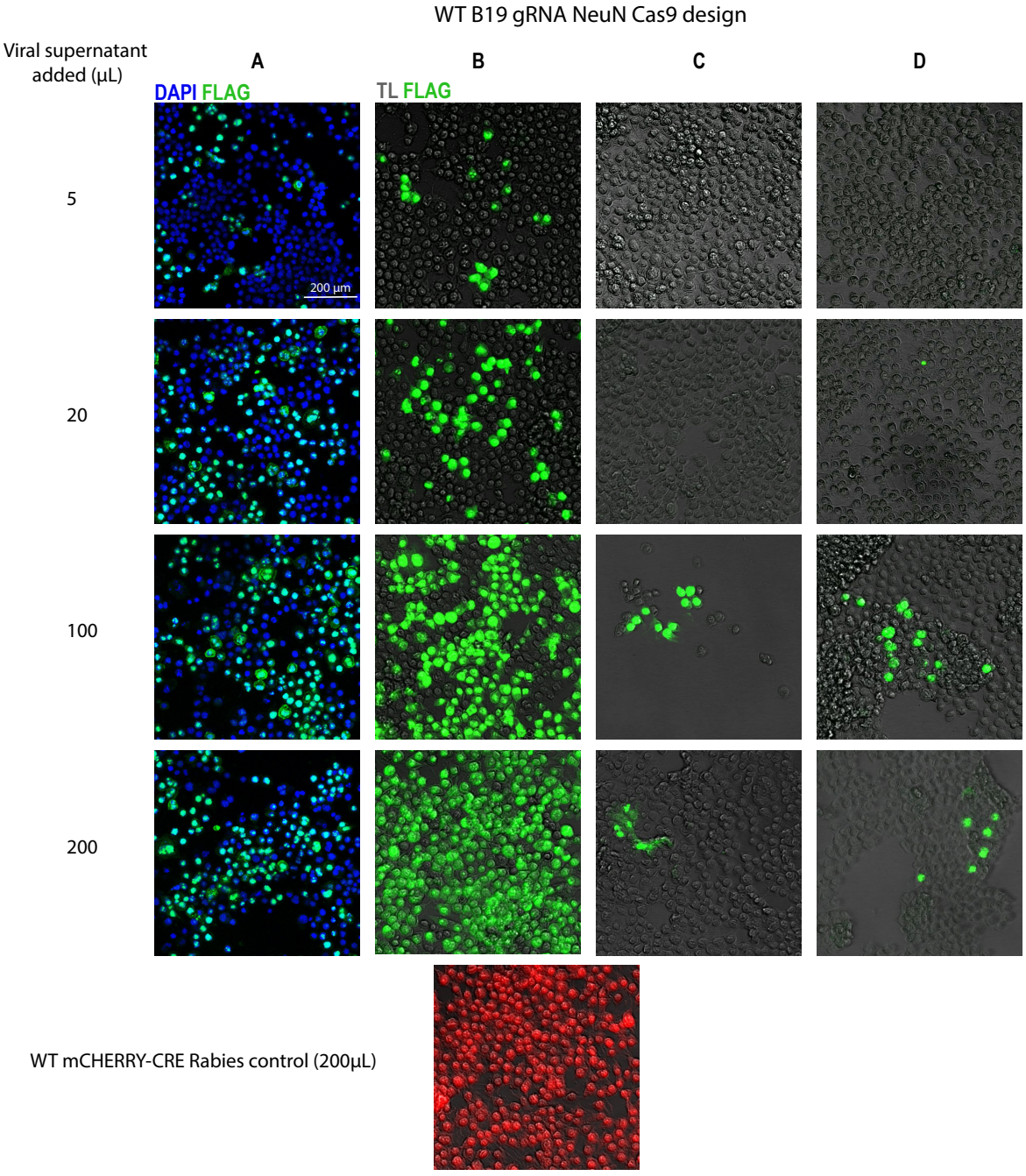


Fig. 4.8 Production of 4 WT gRNA NeuN Cas9 designs for increased gRNA expression The 4 designs of WT gRNA NeuN Cas9 viruses depicted in Fig. 4.8 were cloned into pSAD-B19 ΔG genome vectors and produced following standard B19 ΔG production protocols. Viral supernatant collected from these productions were used to infect mouse neuro2A cells and subsequent FLAG epitope staining (green) to assess successful viral production and viral titration. Due to lack of virally expressed fluorescent proteins in these viruses, a parallel production of a WT B19 mCHERRY-CRE rabies was used as a positive control for successful plasmid transfection and viral spreading during the production process. (TL = transmitted light imaging)

4.2.7 SiR gRNA Cas9 mediated genome editing *in vivo*

4.2.7.1 SiR B19 gRNA NeuN Cas9 virus can be produced to high titers following a standard SiR B19 production protocol

Following confirmation of WT ΔG rabies mediated CRISPR-Cas9 genome editing *in vitro*, we then wanted to produce a SiR virus expressing the same CRISPR-Cas9 expression cassette (design B) that would eventually be necessary for long-term rabies virus mediated genome editing in circuits *in vivo* (Fig. 4.9A). Although derived from the WT B19, the attenuation of the N in the SiR results in delayed viral replication and production kinetics and so we needed to directly validate that a SiR B19 gRNA NeuN Cas9 virus (version B) could also be produced to the high titers necessary for successful *in vivo* injections. Following the standard production protocol of SiR B19 viruses (methods), we successfully recovered infective viral supernatant following initial plasmid transfection in HEK-TGG production cells over a period of 3 weeks. Subsequently, the viral supernatant was amplified further in new HEK-TGG production cells for a week before purification of high-titer viral productions.

Following viral purification, the production was titered by infecting neuro2A cells in culture with serially diluted purified virus. 3 days following infection, the infected cells were fixed and stained with a Cy3 conjugated anti-FLAG primary antibody for visualisation of virally infected cells (Fig. 4.9B). Titers of approximately $2x10^9$ i.u./mL were obtained following the standard protocol.

4.2.7.2 SiR B19 gRNA NeuN Cas9 shows effective genome editing *in vivo*

The purified SiR B19 gRNA NeuN Cas9 virus was subsequently injected into the CA1 of mouse hippocampi to test whether this novel virus with 2kb payload over the WT genome and a hairpin gRNA sequence retained the ability to infect neurons *in vivo* following intracerebral injection. In addition, we wanted to assess in this preliminary experiment, whether the SiR virus could target genomic loci efficiently in infected neurons *in vivo*. Again, although the SiR is derived from the WT B19 virus, due to the attenuation of the N protein we observe the SiR has lower transcriptional levels and kinetics. Therefore, there was a possibility that the lower gRNA and SaCas9 expression in SiR would decrease the efficiency of genome editing. However, compared to our tests of genome targeting efficiency of the WT virus over a period of 3-6 days *in vitro*, the SiR vector could be used to express CRISPR-Cas9 components for a much longer period of 1-2 weeks (the period of viral transcriptional activity following initial infection but before viral self-inactivation). This longer expression would in theory also

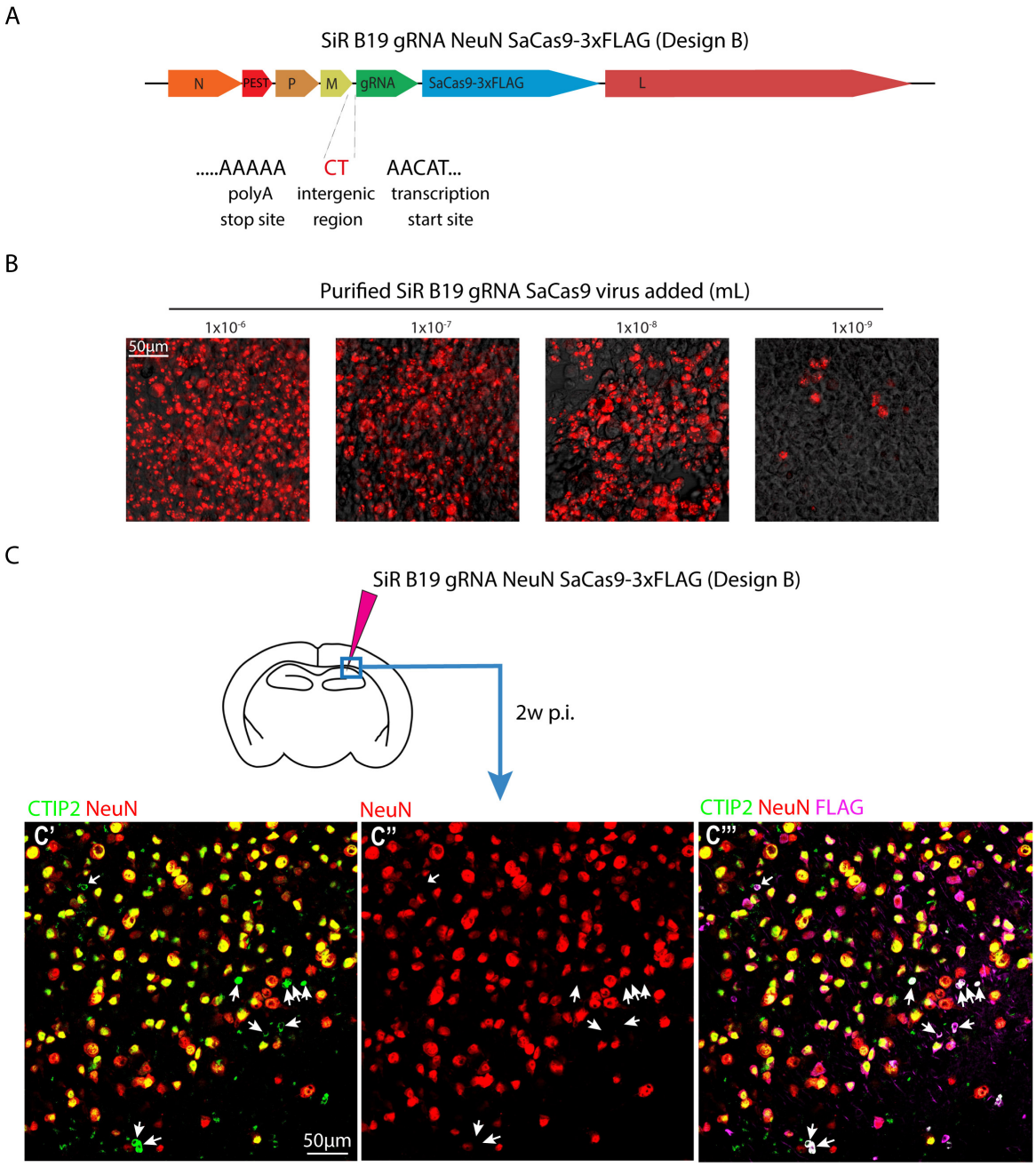


Fig. 4.9 SiR B19-gRNA-NeuN-SaCas9 can be produced to high titers and mediate genome-editing of endogenous genes *in vivo* **(A)** Gene map of SiR B19-gRNA-NeuN-SaCas9-3xFLAG. **(B)** Functional titration of purified SiR B19-gRNA-NeuN-SaCas9-3xFLAG virus in neuro2A cells using FLAG antibody staining (red) to visualise virally infected cells. **(C)** SiR B19-gRNA-NeuN-SaCas9-3xFLAG was injected into the CA1 region of adult mouse hippocampi. 2 weeks later, hippocampal sections were stained for CTIP2 (CA1 neuronal marker, green), NeuN (gRNA targetted locus and pan-neuronal marker, red) and FLAG (virally infected cells, magenta). Neurons in CA1 should be both CTIP2+ and NeuN+. Images show the presence of CTIP2+NeuN- cells (C', C'') which are all FLAG+. This indicates the knockdown of an otherwise expressed NeuN protein in virally infected cells.

increase the on-target genome targeting efficiency of the SiR compared to the WT virus *in vitro*.

2 weeks following injection, adult mice were perfused and hippocampal sections stained for FLAG (magenta) to visualise virally infected cells with still transcriptionally active rabies virus, NeuN (red) to visualise whether the virus had successfully targetted the NeuN locus resulting in a decreased or absent staining of the NeuN protein in infected cells, and CTIP2 (green) which is an alternative neuronal marker for CA1 neurons[302]. In theory, all neurons in CA1 should be both CTIP2 and NeuN positive. However, we observed in our SiR-B19-gRNA-NeuN-Cas9 infected sections neurons which were CTIP2 positive but NeuN negative (Fig. 4.9C', C") which were all FLAG+ (Fig. 4.9C"'). This indicated SiR mediated knockdown of an otherwise expressed endogenous NeuN gene *in vivo*.

For further robust quantification of this observation, we are planning further experiments that would allow us to permanently label SiR gRNA NeuN Cas9 infected cells. In this preliminary experiment, we can only visualise transcriptionally active SiR virus from its expression of FLAG, and so cannot ensure labeling of the majority of virally infected cells unless working within short time frames of 1 week p.i.. In addition, we are in the process of producing a SiR gRNA scramble Cas9 control virus to confirm that indeed the decreased NeuN staining observed in these preliminary experiments are a result of targeted Cas9 modification as opposed to being an off-target effect of Cas9 expression in neurons from the SiR vector or a general consequence of viral infection.

4.2.7.3 CRISPR-Cas9 mediated genome-editing using the SiR 2.0 vector

As well as producing SiR-B19-gRNA-Cas9 viruses, we also wanted to produce SiR 2.0 CRISPR-Cas9 vectors. This was because the SiR 2.0 vector is less immunogenic then the SiR 1.0 vector, and reduced immunogenicity is a very important advantageous trait for any potential gene therapy vector. In addition, the SiR 2.0 vector allows more efficient targeting of spinal motor neuron networks following peripheral injection (a commonly affected disease network in motor disorders) that could potentially be a target for SiR mediated gene therapy. Therefore, we initiated the viral production of two SiR 2.0s: a SiR-N2C-gRNA(scramble)-SaCas9-FLAGx3 and a SiR-N2C-gRNA(tdTOMATO4)-Cas9-FLAGx3 virus. We selected the tdTOMATO4 gRNA sequence for targeting the tdTOMATO locus of transgenic mice following experimental validation *in vitro* (Fig. 4.13C, D).

Again, despite the addition of a very similar gRNA-SaCas9-FLAGx3 expression cassette, we needed to directly verify that the SiR 2.0 virus could also be produced to high titers for *in vivo* use whilst incorporating long SaCas9 and hairpin gRNA sequences. This was

because SiR 2.0 displays much reduced transcriptional kinetics compared to the SiR 1.0 and may therefore respond to the same exogenous sequences in different ways. 3 weeks following viral recovery from plasmids, production cells were stained with a Cy3 conjugated anti-FLAG primary antibody for confirmation of successful viral production. Following further viral amplification and high-titer viral purification, functional titration of the SiR-N2C-gRNA(scramble)-SaCas9-FLAGx3 virus using FLAG antibody staining of virally expressed proteins revealed that SiR 2.0 viruses encoding SaCas9 and gRNA sequences can be produced to titers of $5X10^8$ i.u./mL (Fig. 4.10). These functional titers were similar to those obtained for the production of standard WT ΔG N2C viruses (Fig. 3.6). Such titers are more then sufficient for efficient infection of neurons following intracerebral or peripheral intramuscular injection.

These 2 viruses can be used to assess the efficiency of SiR 2.0 mediated CRISPR-Cas9 targeting of the tdTOMATO locus in transgenic mice in future experiments. Our motivation for targeting the tdTOMATO gene using the tdTOMATO4 gRNA sequence was due to its extremely high on-target efficiency as observed experimentally *in vitro* (Fig. 4.13C and D). In addition to being approximately 3 times more efficient at on-target DSB induction compared to the NeuN3 gRNA sequence, it can target the tdTOMATO coding sequence in two different locations within the target gene. This is because the tdTOMATO coding sequence is a tandem repeat of the dTomato fluorescent protein. Therefore, we expect that CRISPR-Cas9 mediated knock-out of the tdTOMATO genomic locus can be visualised more sensitively *in vivo* compared to that of the NeuN marker using gRNA NeuN3.

4.2.8 Tool developments to enable more efficient SiR gRNA Cas9 virus production and screening *in vitro* and *in vivo*

4.2.8.1 Development of a split FLP system to enable live and permanent visualisation of SiR-gRNA-Cas9 infected cells

One major technical difficulty faced during this project was the inability to observe viral productions live. Although possible, "blind" productions are more labour intensive then the production of viruses which express proteins that can be visualised directly in live cells. In addition, when moving *in vivo*, our viral designs did not enable permanent labeling of SiR infected cells. This was a limitation due to the fact that the SiR viruses self-inactivate over a period of 1-3 weeks, and so for more longitudinal experiments, there is no method of labeling cells that have once been infected with the genome-editing SiR virus. Even for shorter time points before complete viral self-inactivation, one cannot ensure that all virally infected cells

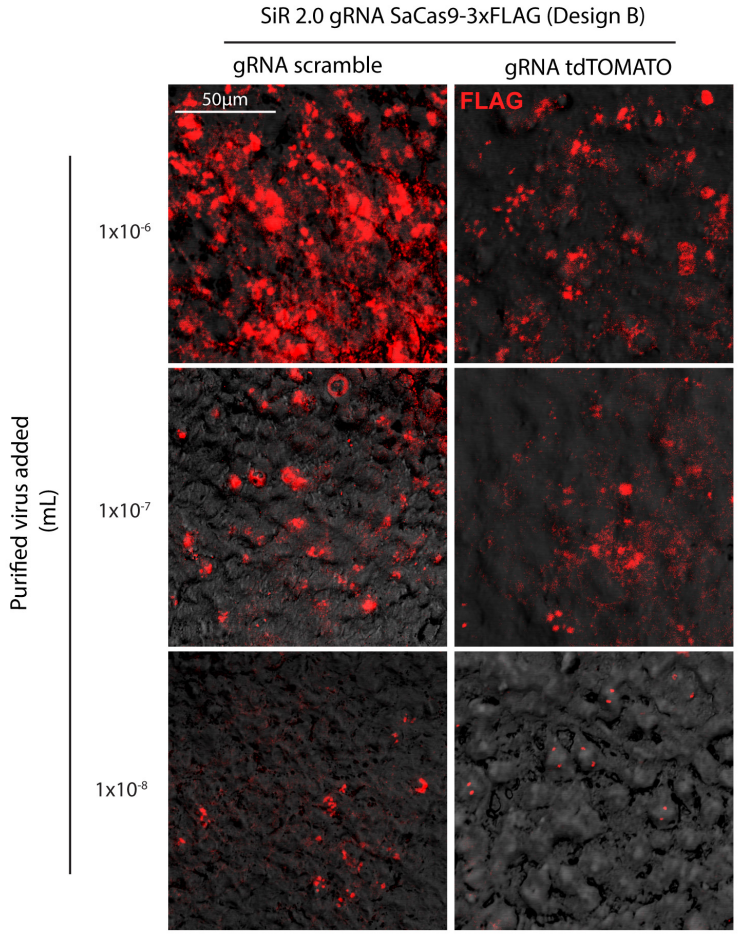


Fig. 4.10 High-titer production and functional titration of SiR N2C-gRNA-SaCas9-3xFLAG viruses Purified viral productions of two different SiR N2C-gRNA-SaCas9-3xFLAG viruses (one with a scrambled gRNA sequence, and another with the tdTOMATO4 gRNA sequence) were produced blind and used to infected mouse neuro2A cells in culture for functional titration. Virally infected cells were visualised 3d p.i. using a Cy3 conjugated anti-FLAG primary antibody. Titers of $5X10^8$ i.u./mL were obtained.

are being visualised due to variability of viral infections between neurons surrounding the injection site in the same animal.

In order to overcome both these problems, we designed a split-FLP system whereby a short C terminal fragment of FLP (flippase, a site directed recombinase) was added into the viral cassette and could be complemented in infected cells by expression of the longer N terminal FLP fragment(Fig. 4.11). This split FLP design was modified from a known functional split site of the FLP enzyme (between AA 396-397) previously used in conjunction with chemically inducible heterodimerization domains in order to reconstitute split-FLP activity using phytohormone abscisic acid (ABA)[303]. In our case, we were looking to reconstitute FLP activity when both fragments were co-expressed in a cell. Therefore we added mutated Fos and Jun leucine zipper heterodimerization domains[304] and SV40 nuclear localisation signals to the C and N terminus of the AA1-396 and 397-423 FLP fragments respectively via a flexible 23AA glycine/serine F6 linker for efficient spatial reconstitution.

This resulted in a short modified C terminal FLP sequence (297bp) that could be added to the rabies virus genome to enable both live visualisation of viral transcription during viral productions, and also permanent labeling of SiR infected cells *in vivo*.

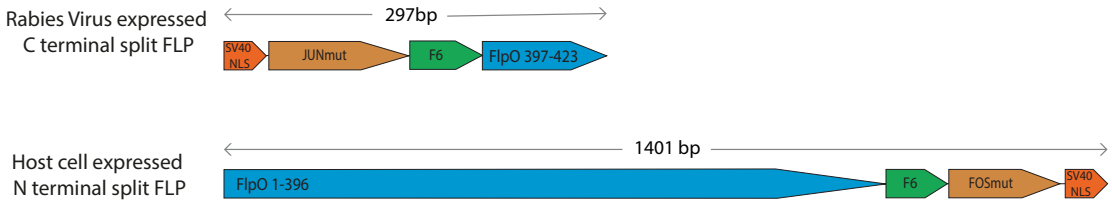


Fig. 4.11 Design of a split FLP reporter system The design of a co-expression FLP reporter system using known FLP split sites between AA 396-397. The FOS and JUNmut heterodimerisation domains and SV40 nuclear localisation signals were added to the split FLP fragments to enable functional reconstitution of FLP activity when co-expressed in the same cell. Dimerisation domains were added via a flexible 23AA glycine/serine linker (F6) to the N terminus of the C terminal FLP fragment, and the C terminus of the N terminal FLP fragment.

4.2.8.2 Development of a bidirectional fluorescent cellular assay for quick and sensitive assessment of genome targeting efficiency *in vitro*

Another technical challenge we faced during the course of this project was the multi-step time-consuming nature and low sensitivity of the T7 endonuclease assay for assessing genome targeting efficiency of rabies virus variants *in vitro*. Use of the T7 endonuclease I assay to assess genome targeting of Cas9-gRNA expressing rabies viruses would require 7

intermediate experimental steps: the staining and fluorescence-activated cell sorting (FACS) of virally infected cells, genomic DNA extraction from infected cells, PCR of the targeted genomic locus, denaturation and reannealing of the PCR amplicon, incubation with the T7 endonuclease I enzyme which can selectively cleave distorted duplex DNA formed via cross-annealing of mutated and wild-type sequence, and finally cleavage analysis by gel visualisation of PCR fragments (Fig. 4.12A).

In addition to this multi-step and time-consuming experimental workflow, we found from our experience that the T7 endonuclease I assay resulted in quite high variability (e.g. three experimental replicates depicted in Fig. 4.4D). In addition, it did not show high sensitivity to WT B19 rabies virus induced CRISPR-Cas9 genomic mutations *in vitro* (Fig. 4.4C), thus limiting its use for the comparison of genome editing efficiency of different viral cassettes. The low sensitivity of this assay is partly a result of its methodology. Even with 100% gene editing efficiency, opposing strands of PCR amplicons can re-anneal resulting in perfect complementarity and thus lack of detection by resolvase enzymes, despite successful gene targeting[305]. Studies in the literature have shown that compared to other biochemical methods, the T7 endonuclease assay fails to measure weaker levels of Cas9 cleavage activity[295].

Given the experimental difficulties and lack of sensitivity of the T7 endonuclease system, we adopted a previously reported cellular double fluorescent reporter system where successful gRNA directed DSBs *in vitro* result in a two way amplified fluorescent signal[306]. In this design, reporter cells initially express GFP under the control of a CAG promoter (Fig. 4.12C). As the GFP coding sequence is flanked by two gRNA target sites on either side, if there is efficient gRNA targeting it results in excision of the GFP coding sequence and the subsequent expression of a TagBFP2 fluorescent protein under the control of the same CAG promoter. As a consequence, efficient genomic editing results in both a decrease in the GFP signal, and an increase in the TagBFP2 signal. This system would in theory require just 2 steps for measurement of genome editing efficiency by different rabies virus designs: antibody staining of virally infected cells followed by FACS analysis of the two reporter fluorescent proteins GFP and BFP in virally infected cells.

4.2.8.3 Screening of gRNA NeuN variants 1-3 and tdTomato variant 1-3 using double reporter cell system

In order to test the double reporter cell system, we used it to screen the 3 NeuN targeting gRNA variants that we had previously assessed using the T7 endonuclease assay (Fig. 4.13A). Using this method we had not detected significant NeuN locus PCR amplicon cleavage following the expression of gRNA NeuN 1 or 2 compared to gRNA GFP controls.

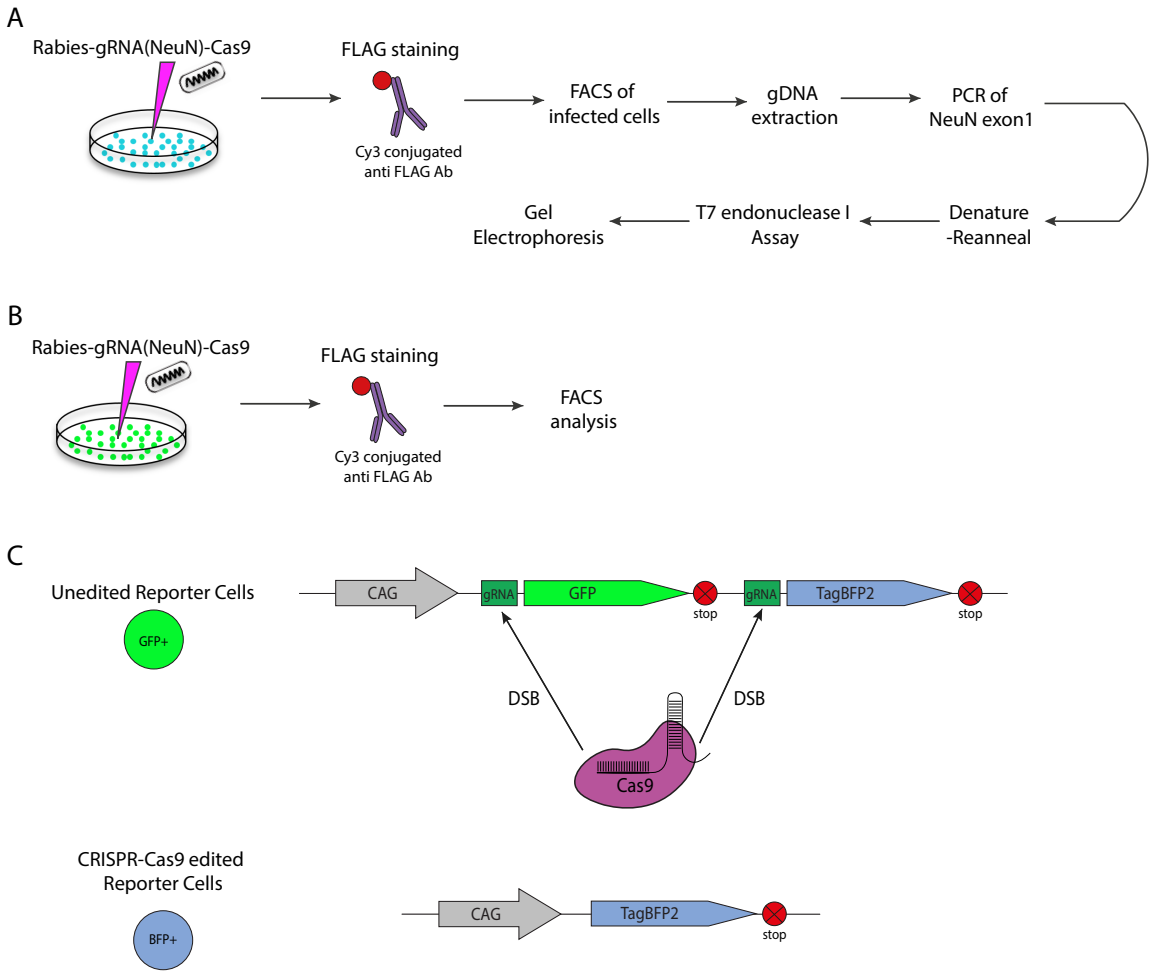


Fig. 4.12 Alternative assays for the detection of rabies virus mediated genome editing (A) Labour intensive experimental workflow for the measurement of virus mediated genome editing efficiency using the T7 endonuclease I assay. **(B)** Shorter 2-step experimental workflow for the measurement of virus mediated genome editing efficiency using a double reporter cell system. **(C)** Mechanism of the double reporter cell system. Unedited reporter cells contain expression cassettes that drive the expression of GFP under a CAG promoter. The GFP coding sequence is flanked by gRNA target sequences such that efficient DSB induction results in excision of the GFP sequence, and subsequent expression of TagBFP2 under the same CAG promoter. This gives a faster and more sensitive read out of genome targeting efficiency compared to the T7 endonuclease I assay.

HEK293T cells were transfected in culture with plasmids encoding the double reporter system construct containing target sequences for gRNA NeuN 1, 2 or 3. Cells were also co-transfected with CMV-Cas9 and U6-gRNA NeuN 1, 2 or 3 encoding plasmids. In order to test CRISPR-Cas9 induced reporter activation, the reporter construct needed to contain target sequences complementary to the gRNA sequence being tested. Using this double reporter cell system we found that all 3 NeuN gRNA designs did show significant genome targeting activity compared to controls (Fig. 4.13B,D). This was unlike the lack of detection of targeted genome editing by gRNA NeuN 1 and 2 using the T7 endonuclease assay.

We also used this system to screen 4 new gRNA sequences designed to target the tdTOMATO coding sequence. This revealed that gRNA tdTOMATO4 showed extremely high on-target DSB cleavage efficiency that was approximately triple that detected for gRNA NeuN3 (Fig. 4.13C,D). We thus incorporated this gRNA tdTOMATO4 in the production of a tdTOMATO targeting SiR2.0-gRNA-Cas9 virus (previous section).

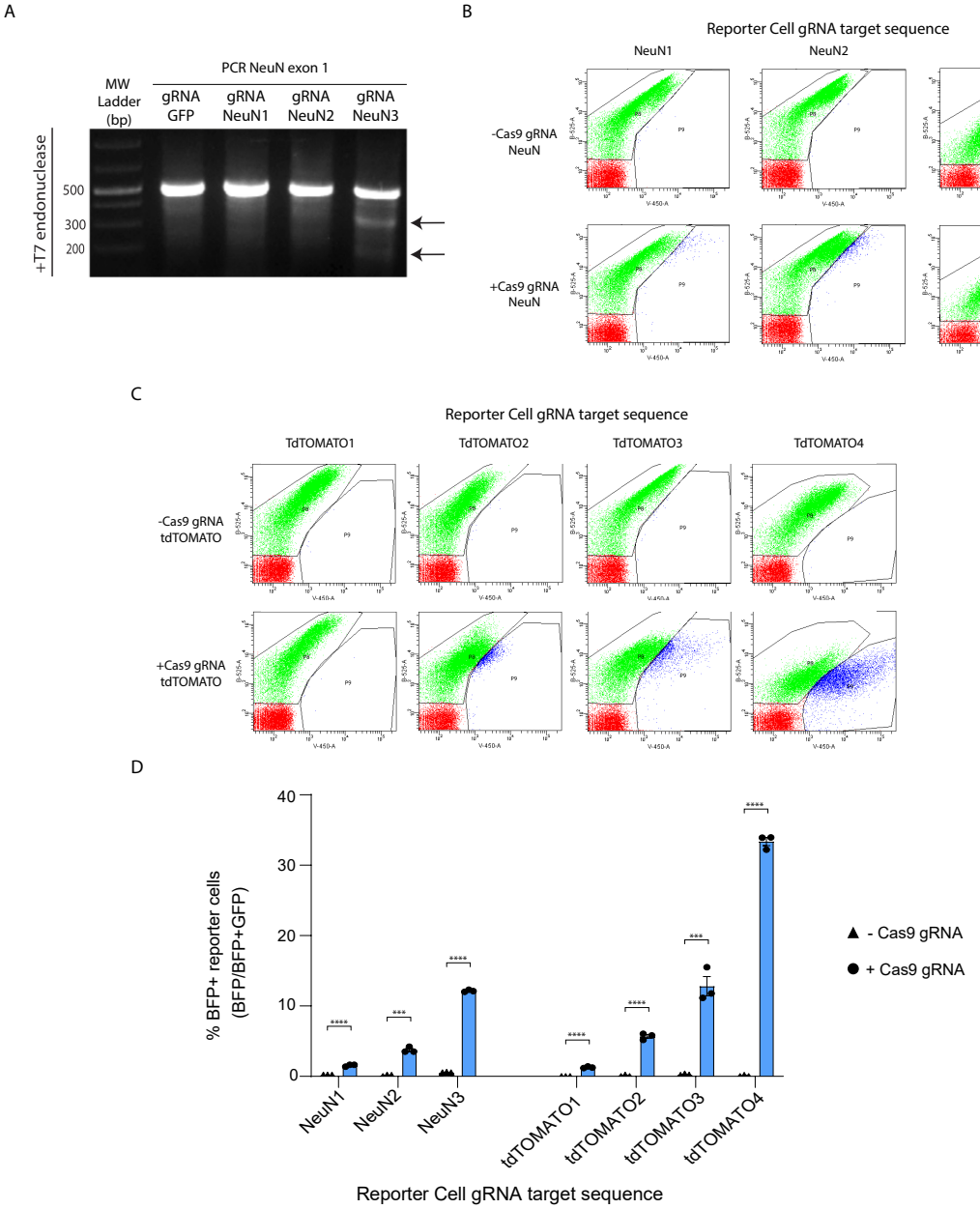


Fig. 4.13 Testing of gRNAs with double reporter cell system (A) Screening of gene targeting efficiency gRNA NeuN 1-3 using T7 endonuclease I mismatch cleavage assay. **(B)** Screening of the same 3 NeuN gRNAs using double reporter cell system by FACS. **(C)** Screening of 4 tdTOMATO targeting gRNAs using double reporter cell system by FACS. **(D)** Quantification of reporter cell activation by % of BFP+ reporter cells for gRNA NeuN 1-3 and gRNA tdTOMATO 1-4. n=3, T-test between -/+ Cas9gRNA conditions for each gRNA target sequence with correction for multiple comparisons using the Holm-Sidak method.

4.3 Summary

In this experimental chapter I have worked to express the two CRISPR-Cas9 components (a Cas9 endonuclease and a gRNA) from the rabies virus in order to initiate the creation of a circuit-specific genome editing vector that can be used both for long-term experiments into the roles of genes within specific circuits for brain function, as well as for potential gene therapy.

I have shown that despite various molecular caveats, the WT B19 ΔG rabies virus can be modified successfully to express both Cas9 endonuclease and gRNA without impairing normal viral replication, production or infectivity. This rabies virus can successfully target endogenous gene loci such as the NeuN locus in mouse neuro2A cells *in vitro*. We also found that increasing gRNA expression could significantly increase the efficiency of WT ΔG rabies virus mediated CRISPR-Cas9 genome editing, and so we designed 4 different Cas9-gRNA viral expression cassettes to increase gRNA expression levels using different known methods of viral transcription regulation. We then screened these 4 designs for efficient production *in vitro*.

Ultimately, we want to be able to genomically edit circuits *in vivo* to investigate the roles of genes in behaviour of experimental animals, as well as for potential gene therapy. The WT ΔG rabies virus would not be suitable for this purpose as it would result in the death of genomically edited circuits within 1-4 weeks p.i.. Therefore, having selected a suitable gRNA-Cas9 cassette design from the screening of WT viruses *in vitro*, we aimed to produce high titers of both SiR 1.0 and 2.0 gRNA-Cas9 viruses for *in vivo* use. We found that both such viruses could be produced to high-titers that would be necessary for successful infection of neurons *in vivo*. In addition, a preliminary proof of principle experiment demonstrated SiR B19 vector mediated NeuN locus targeting in the CA1 region of mice hippocampus 2 weeks p.i.. This indicates that SiR-gRNA-Cas9 vectors can be used to target endogenous genes in circuits *in vivo*.

Chapter 5

Interneuronal Protein Transfer

5.1 Introduction

In this project we aim to test the trans-synaptic circuit hypothesis of transneuronal protein transfer. This model implies that proteins may move from one neuron to another much like the electrical signal that flows through neural networks. Despite the accumulating evidence for transneuronal protein transfer in both physiological and disease processes (covered in chapter 1.3), very little is understood of its mechanism, control or function. In addition, although there is an observed correlation between protein movement and synaptically connected neuronal pathways, it is still unclear whether this protein transfer is truly trans-synaptic.

5.1.1 The screening of experimental systems to study interneuronal protein transfer

In order to investigate interneuronal movement, we needed to establish experimental paradigms that would enable us to visualize and mechanistically probe this biological process. For example, in the case of commonly observed pathological protein aggregate propagation through circuits, this is a very slow process that accumulates over decades in patients and over time-scales of months to years in experimental animals[119,172]. As such, use of such systems to iteratively screen novel protein-tags and identify transcellularly moving proteins would not be practically feasible in reasonable time-scales. Other faster examples of interneuronal protein movement have been published. However, the observation of trans-synaptic transfer varies hugely depending on the experimental conditions such as: the method of establishment of donor cells (e.g. transgenic, viral injection, protein injection), the method of detection of protein movement in receiver cells (e.g. radioactive labelling, immunofluorescence, immunocytochemistry), the age or period of development of the experimental animal and the neural circuitry being investigated. As such, studies using protein tracers such as WGA-HRP or CtxB have shown highly varied results as to whether they exhibit or do not exhibit trans-synaptic transfer ability[167,168].

Therefore, we decided to consider many different experimental systems to test if they could be used, for our purposes, to observe and thus investigate interneuronal protein transfer at relatively quick time-scales. Potential experimental systems are suggested in Fig. 5.1, and those that were directly experimentally investigated are highlighted in red. Some of the experimental paradigms that we investigated had not previously been used to study interneuronal protein transfer but we thought in theory they could be good candidates (e.g. in utero electroporation). Others were experimental systems that had repeatedly been used to reveal interneuronal protein transfer in many different contexts (e.g. the primary visual

sensory pathway through which trans-synaptic protein transfer has been detected following intraocular injection of tritiated amino acids[174], WGA-HRP[307] and biotinylated Otx2[169]). Ideally, we needed to develop an experimental system that would enable us to observe robust interneuronal protein movement in the time-scale of days. In this way, we could use this system to iteratively screen for the efficacy of novel protein movement visualisation tools. We could also use this system to dissect in more details the mechanistic details of the process being visualised within a shorter time-frame.

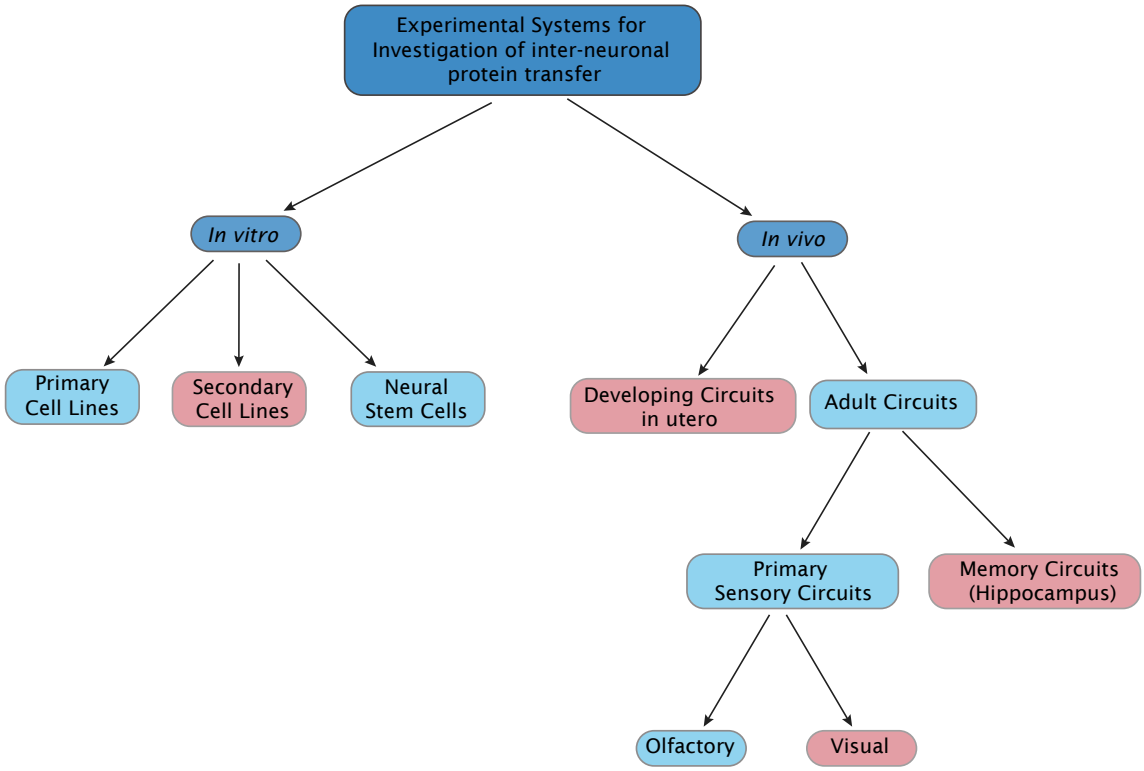


Fig. 5.1 Flow chart of possible experimental systems to investigate interneuronal protein transfer Experimental systems highlighted in red are those that have been directly investigated.

5.1.2 Aims and experimental strategy

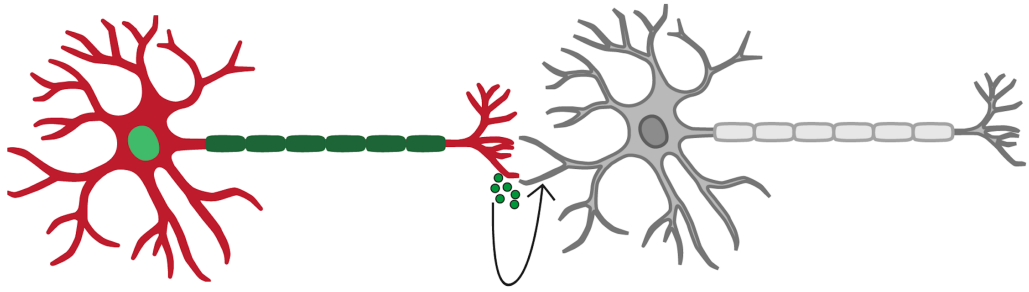
In such a context, the main aims of this project are:

1. To develop experimental paradigms that will allow us to efficiently probe the mechanisms behind interneuronal protein transfer *in vitro* and trans-synaptic protein transfer *in vivo* through neural circuits (Fig. 5.1).

2. To determine the universality of this process by using top-down proteomic approaches to label all proteins produced from a donor cell population. The two methods of whole proteome labelling that I investigate are: a) the use of a promiscuous biotin ligase to biotinylate all cytoplasmic proteins in donor cells (BioID2) and b) the use of a mutant methionyl tRNA synthetase (mutRS) which enables the incorporation of an azide labelled methionine analogue called azidonorleucine (AnI) into all newly synthesised proteins.

3. To develop a small split CRE tag that we can append onto candidate proteins (e.g. Tau, Otx2, WGA) that amplifies the signal of interneuronal protein movement and genetically modifies receiver cells of trans-synaptic protein movement.

For aims 2-3, we adopt a general strategy where we express labelled proteins in a donor neuron population and observe their movement into acceptor neurons which are cellularly distinct to the donor neuron population (Fig. 5.2).



	DONOR neuron	ACCEPTOR neuron
Bottom up	Split CRE tagged Candidate Protein (e.g. Tau, Otx2, WGA)	Split CRE + floxed fluorescent marker
Top down	Labelling of entire proteome (e.g. mutant tRNA synthetase, BioID2)	--

Fig. 5.2 Project strategy to investigate mechanistic details of interneuronal protein transfer
We propose two parallel approaches. First, a bottom-up approach involving the addition of small recombinase inducing tags to proteins known to demonstrate interneuronal movement. Second, a top down approach involving the unbiased labelling of the entire proteome of donor neurons, enabling the identification of novel interneuronally transferred proteins.

5.2 Results

5.2.1 The mouse visual circuitry as an *in vivo* model of interneuronal protein transfer

The visual pathway has repeatedly been shown to be an exemplary *in vivo* circuit to investigate trans-synaptic protein transfer. For example, the intraocular injection of tritiated amino acids, WGA-HRP or biotinylated homeoprotein Otx2 into monkeys, mice or rats have all demonstrated the presence of tritiated, HRP positive or biotinylated proteins respectively in the contralateral visual cortex within time scales of hours to days[169,174,307]. For example, the intraocular injection of 1% WGA-HRP resulted in the labelling of all known primary visual pathways and ocular dominance columns in the visual cortex as well as neuronal soma in the superior colliculus of monkeys and rats as revealed by tetramethyl benzidine histochemistry 3 days post-injection[307]. The visual cortex is anatomically almost on the opposite side of the central nervous system to the eyes and is not known to receive direct input from retinal ganglion cells. As such, the fact that one can observe specific contralateral positive staining of the visual cortex in WGA-HRP injected animals but not HRP injected controls strongly suggests that WGA-HRP is transferred trans-neuronally through the visual circuit (from retinal ganglion cells to cells in the lateral geniculate nucleus which project to the visual cortex). For our purposes, this data also reveals that visual pathways *in vivo* can demonstrate clear interneuronal trans-synaptic transfer within a time scale of 3 days after intraocular administration of proteins to donor retinal ganglion cells.

As such, we decided to investigate the use of the mouse visual pathway for further experiments. We first tested whether we could indeed observe interneuronal protein movement through the visual circuitry by injecting labelled proteins directly into the eye. This resulted in the endocytic uptake of labelled proteins into retinal ganglion cells (RGCs) and subsequent transfer of these proteins from RGCs into the visual circuit. Having verified intraocular protein injections as an efficient method of visualising interneuronal protein transfer, we then proceeded to design recombinant AAVs to transduce retinal ganglion cells. This was because we wanted to investigate the movement of all or any RGC protein through the visual circuitry. The preparation of highly concentrated purified protein preps for intraocular injections was not an ideal system in which to screen proteomic level protein movement in an unbiased manner. In contrast, AAV transductions are a more flexible system which allows us to label entire proteomes by the integration of transgenes into RGC genome. Another advantage of AAVs is that after successful infection and integration of transgenes into the donor cell

genome, the donor RGC can permanently express labelled donor proteins at high levels for the entire duration of the experiment. As such, AAV transduction enables more robust and long-term investigations of trans-synaptic protein transfer. In addition, we could add fluorescent or recombinase markers into the AAV genome in order to permanently label and thus quantify or experimentally modify starting retinal ganglion cell donor populations.

5.2.1.1 Human A53T alpha synuclein-Dylight594 shows interneuronal movement from RGCs into the visual cortex within 4 days

The direct injection of proteins into the eye to visualise protein movement through visual networks has been conducted successfully in many different contexts. For example the injection of WGA-HRP or biotinylated Otx2 directly into the eye results in transneuronal movement of the labelled protein through the visual network[169,307]. In addition, intraocular injection of tritiated amino acids into the vitreous humor results in the uptake and assimilation of these amino acids into proteins within RGCs that then demonstrate transneuronal movement[174]. To test the efficiency of this process of endocytic protein uptake directly from the vitreous humor into the visual circuit, we injected fluorophore-tagged proteins cholera toxin subunit B (CtB)-AF488 or aggregated human mutant (A53T) alpha synuclein-Dylight 594 into the eye (Fig. 5.3A). In this experiment, CtB-AF488 was used as a negative control for transneuronal protein movement. CtB has not been shown to move transneuronally in visual pathways following intraocular injection and is routinely used for non-synaptic retrograde labelling of neurons in anatomical studies. CtB attaches to cells by binding to ganglioside GM1 and is taken up into neurons by receptor mediated endocytosis. Aggregated human mutant A53T alpha synuclein was used as a test protein as alpha synuclein aggregates demonstrate interneuronal movement in disease[108]. We wanted to observe whether it could also demonstrate interneuronal movement through mouse visual pathways.

4-7 days p.i., mice were perfused and retinal flat mounts prepared to assess successful intraocular injection whilst serial coronal sections were taken throughout the entire brain to visualise the transport of the protein into the visual pathway (Fig. 5.3A). We were able to see uptake of injected protein into RGCs in the retina for both CtB and alpha synuclein, demonstrating efficient endocytic uptake of intraocular proteins into RGCs (Fig. 5.3B, C). In the case of CtB, we could also strongly visualise CtB in RGC axons and their direct projection sites within the brain such as the suprachiasmatic nucleus, lateral geniculate nucleus and superior colliculus (Fig. 5.3D, F, H). The superior colliculus (SC) and lateral geniculate nucleus (LGN) which are known to be the main outputs of RGCs showed especially strong fluorescence signals as expected. As previously reported, we did not observe any

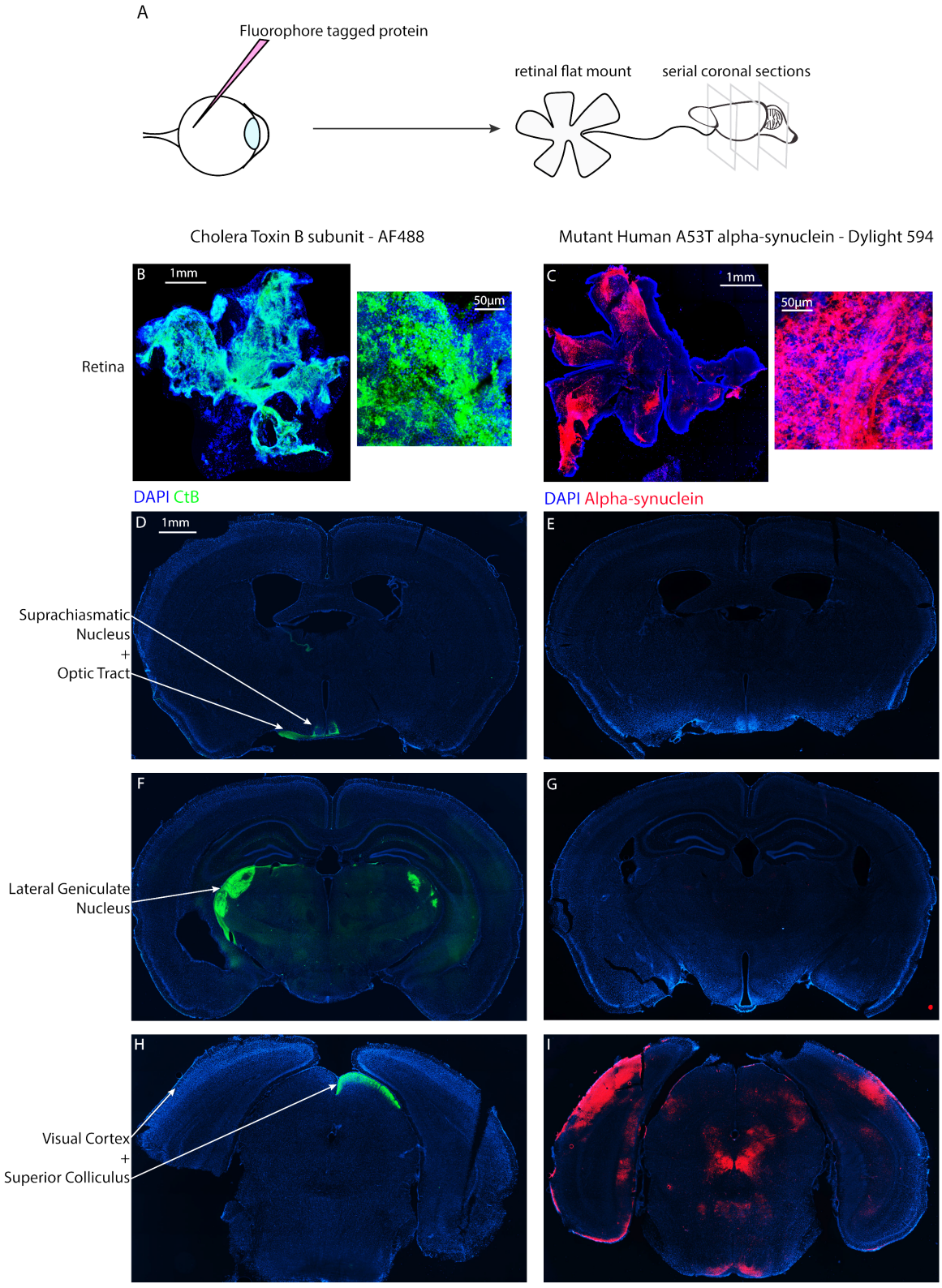


Fig. 5.3 Human mutant A53T alpha synuclein shows trans-neuronal movement through visual circuitry (A) Fluorophore conjugated protein was injected into the right eye of adult mice. Retinal flat mounts **(B,C)** and serial coronal sections throughout the whole brain **(D-I)** were prepared 3-7 days p.i. *(Figure continued on next page...)*

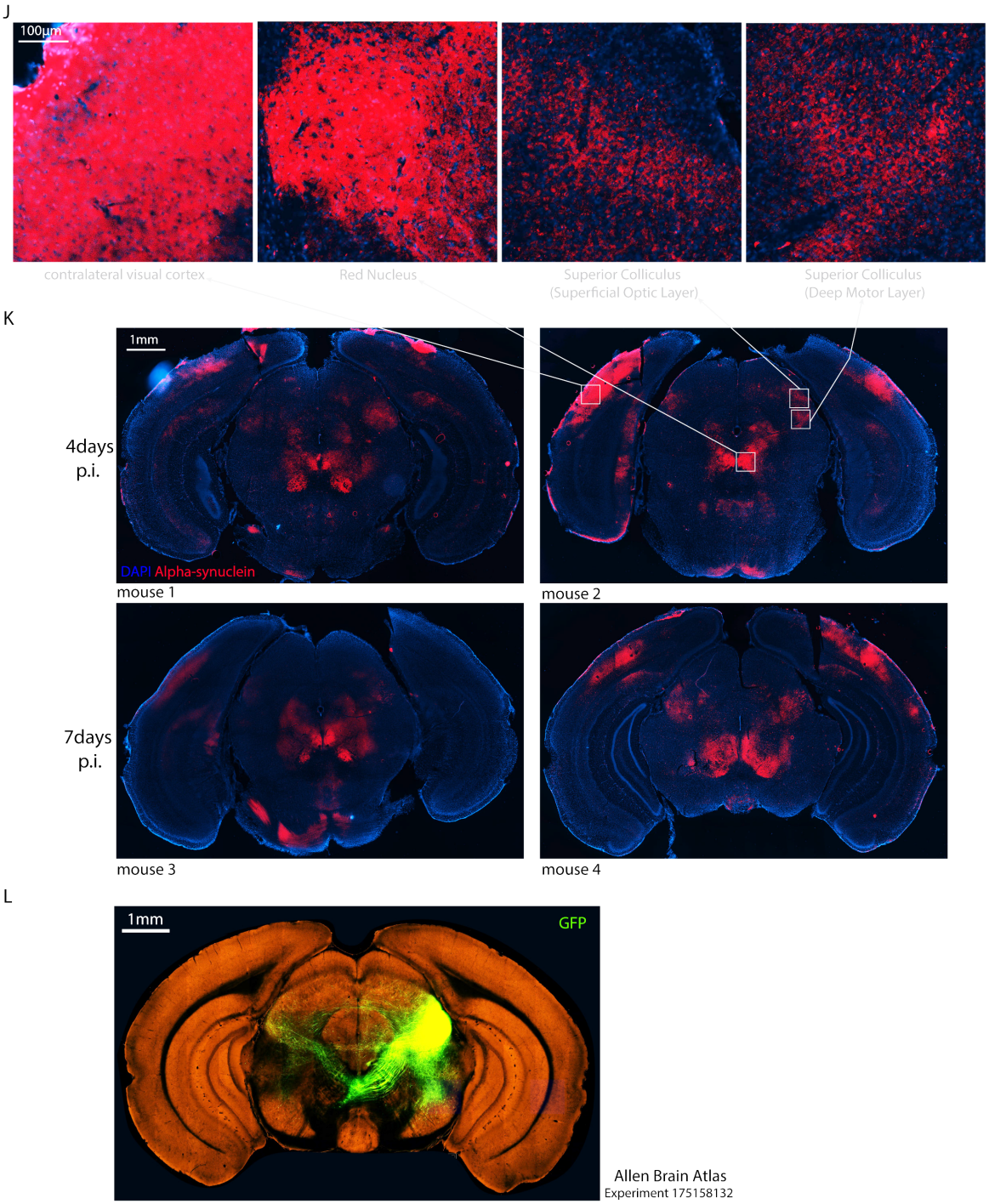


Fig. 5.3 Human mutant A53T alpha synuclein shows trans-neuronal movement through visual circuitry (J) Zoomed-in images of alpha synuclein-Dylight594 positive visual cortices, red nucleus and superior colliculus in the posterior brain section. **(K)** Posterior coronal brain sections from 4 different animal replicates showing repeatable patterns of alpha synuclein-Dylight594 positive areas. **(L)** Snapshot of Allen's mouse connectivity map showing anterograde projections of AAV GFP infected deep superior colliculus neurons to the red nucleus.

transneuronal transfer of CtB into second order neurons. For example, there was no AF488 fluorescence visible in the visual cortex (Fig. 5.3H). In the case of alpha synuclein, we were able to detect some faint protein fluorescence in the SC but not in the LGN. More strikingly we were able to detect significant alpha synuclein within the visual cortex (Fig. 5.3I-K). As there is no known direct output of RGCs into the visual cortex, the presence of alpha synuclein-Dylight 594 signal in the visual cortex following intraocular injection is most probably a result of the transneuronal transfer of this protein from RGCs into second order neurons projecting to the visual cortex from the lateral geniculate nucleus. This was a very promising result as it demonstrated experimentally detectable interneuronal protein transfer through a neural network *in vivo* within a short time scale of 4 days. Hence, this system would allow us to investigate the mechanisms behind the process fairly efficiently within experimentally tractable timescales.

One interesting point to note however was that, quite oddly, the LGN (one of the main projection sites of RGCs) did not show positive fluorescence signals following intraocular injections of alpha synuclein-Dylight 594. This was especially odd given that the alpha synuclein-Dylight 594 was visible in the visual cortex that receives input from lateral geniculate nucleus neurons. In addition, the RGC cell bodies in the retina at the same time points p.i also showed strong presence of alpha synuclein-Dylight594. Thus, the staining pattern which we observed through the RGC to LGN to visual cortex pathway was the presence of alpha synuclein in the retinal ganglion cell bodies and visual cortex, but not in the LGN which receives direct input from RGCs. This pattern was shown independently in 4 out of 4 mice injected intraocularly with alpha synuclein-Dylight594 and analysed at 4 and 7 days p.i.(Fig. 5.3K). This pattern of fluorescence is probably at least partly due to the fact that although there is a high load of alpha synculein in RGCs, the amounts of protein that are transported axonally into projection sites in the brain are much lower and thus harder to visualise. The more puzzling question was however, why there was more alpha synuclein present in the visual cortex than in the lateral geniculate nucleus. One hypothesis could be that the movement of alpha synuclein down the optic tract into the LGN is a limiting step in transneuronal protein transfer through the visual circuit. As such, alpha synuclein is inefficiently transported from RGCs into the LGN (resulting in high levels of protein in RGC cell bodies) but any alpha synuclein that does reach the LGN is quickly transported to the visual cortex thus not allowing a large build up of alpha synuclein in the LGN. In contrast, CtB could be visualised in these first order projection sites such as the LGN and SC as it cannot transfer interneuronally, so can accumulate in RGC axon terminals (Fig. 5.3F and H respectively).

5.2.1.2 Human A53T alpha synuclein-Dylight594 shows interneuronal movement from RGCs into the red nucleus within 4 days

Although we had initially planned this experiment to visualise potential protein movement through the RGC-LGN-visual cortex pathway, we found that repeatedly in our 4 mice replicates alpha synuclein was visible also in the deep motor layers of the superior colliculus and in the red nucleus (Fig. 5.3J,K). The deep superior colliculus has previously been shown to be a major projection site of RGCs in the mouse brain[308]. Subsequently, deep superior colliculus neurons are known to project directly to the red nucleus (Fig. 5.3L, Allen Brain Atlas). Exp.175158132 from the Allen Brain Atlas shows the the strong presence of GFP+ axon terminals in the red nucleus following injection of an AAV-EGFP into the deep layer of superior colliculus.

Given the repeatedly observed pattern of alpha synuclein detection in these sites, and our current knowledge of mouse visual circuit connectivity, we thus concluded that the injection of alpha synuclein into the eye resulted in the uptake of the protein into RGCs projecting to the deep superior colliculus. Subsequently, the protein was trans-synaptically transferred into the deep superior colliculus cell bodies which then project to the red nucleus. Further anterograde transport of alpha synuclein to the axon terminals of these neurons resulted in high alpha synuclein detection in the red nucleus.

To conclude, the strong fluorescence of RGCs post-injection of fluorescently labelled proteins into the eye verified to us that direct intraocular protein injection was an efficient method of incorporating labelled proteins into the RGCs as "donor neurons". In addition, the visualisation of fluorescence in second order projection sites such as the visual cortex and red nucleus within a few days after injection demonstrated that this system could be used within very short experimental time scales to probe the mechanistic details of observed interneuronal transfer.

5.2.1.3 AAV transduction of retinal ganglion cells for unbiased protein screening

Having confirmed that the introduction of labelled proteins into RGCs was an effective method for visualising interneuronal protein transfer through the mouse visual circuitry *in vivo*, we then wanted to develop a more scalable, quantifiable and long-term method of establishing RGCs as protein donor cells. Essentially, we wanted to increase the numbers of proteins that could be investigated for interneuronal protein transfer without relying on the production of highly concentrated purified protein preps. In addition, due to the eventual degradation of injected proteins by cells, the suitability of intraocular protein injection for

longer-term manipulation studies of donor cell properties was limited. Therefore, we considered transducing RGC genomes with recombinant AAV vectors in order to produce "donor RGCs" which would permanently express labelled donor proteins for further investigation.

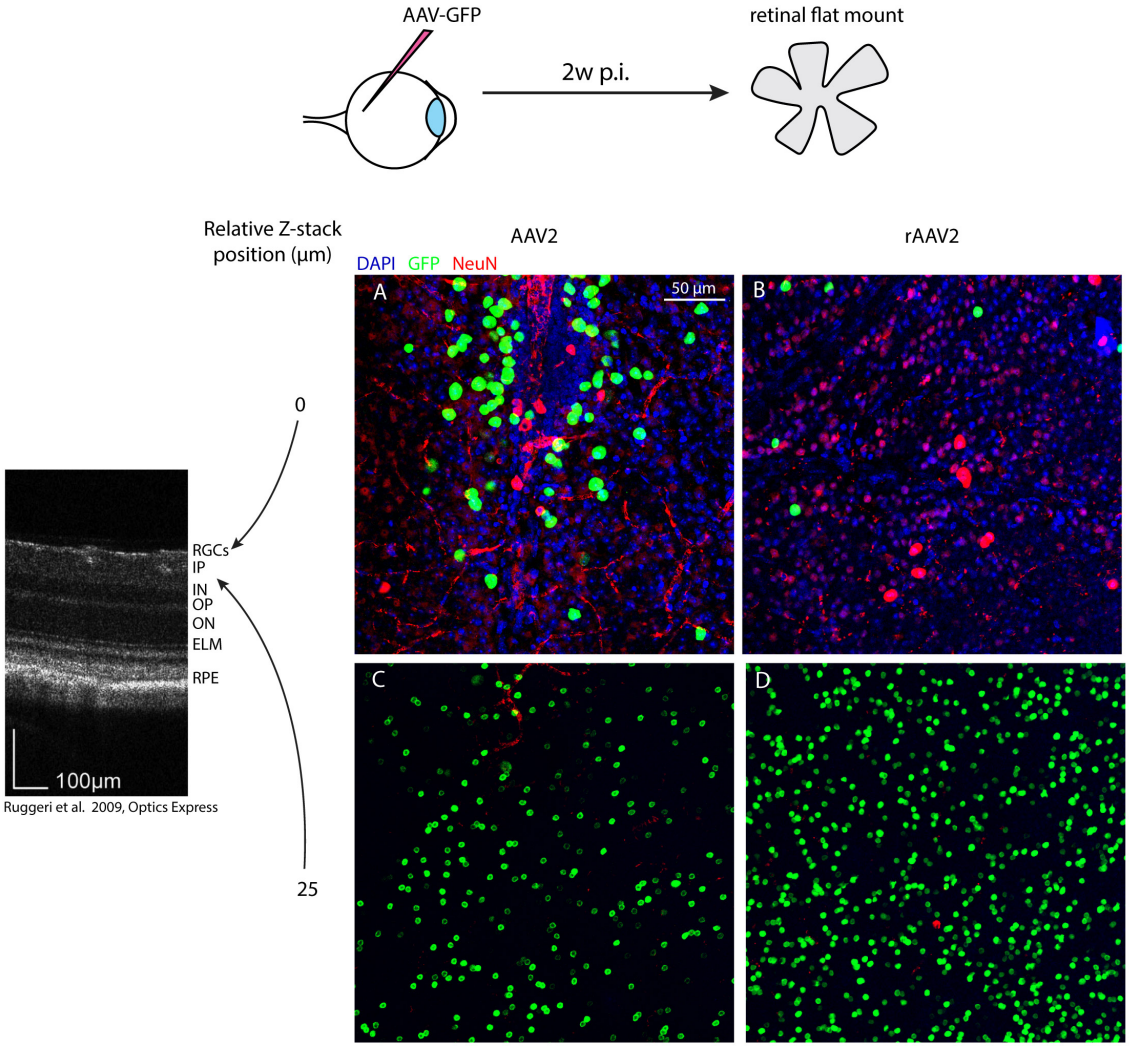


Fig. 5.4 AAV2 infects RGCs at higher efficiency then rAAV2 viruses 2μL of either AAV2-GFP or rAAV2-GFP virus were injected into the vitreous humor of adult mice at a titer of 8 x 10^{12}. 2 weeks p.i., retinal flat mounts were prepared and stained for GFP (green, virally transduced cells), NeuN (red, retinal ganglion cells) and DAPI (blue, nuclei). **(A,B)** Confocal imaging of the retinal ganglion cell layer. **(C,D)** Confocal imaging of deeper inner plexiform/nuclear layers. Left inset, cross-sectional image of WT BL6 mouse retina taken from Ruggeri et al. 2009[309]. RGC: retinal ganglion cell layer; IP: inner plexiform layer; IN: inner nuclear layer; OP: outer plexiform layer; ON: outer nuclear layer; ELM: external limiting membrane; RPE: retinal pigment epithelium.

Within the mouse retina, there are over 30 different types of RGCs. 90% of them project to the superior colliculus (SC), 30-40% project to the lateral geniculate nucleus (LGN) and a

smaller percentage project to other subcortical areas. For example, melanopsin expressing RGCs project to the suprachiasmatic nucleus (Fig. 5.3D) and are involved in circadian rhythms[308,310]. AAVs have been used extensively for the efficient infection of RGCs. Of the many different AAV serotypes that have been used and modified, AAV serotype 2 (AAV2) has become a standard vector for transducing retinal ganglion cells with recombinant genes. AAV2 can efficiently transduce over 85% of RGC populations in adult rodents after intraocular injection[311–313]. Therefore we decided to test a AAV2-GFP construct to observe the efficiency of virally mediated RGC targeting. We found that the AAV2-GFP was indeed efficient at infecting RGCs within our experimental conditions (Fig. 5.4A). We also decided to test the efficiency of a retro-AAV2-GFP virus (rAAV2). This was because, to our knowledge, this serotype had not yet been tested for RGC transduction efficiency. This relatively new serotype published in 2016 was developed through *in vivo* directed evolution for increased retrograde targeting efficiency and shows enhanced retrograde targeting efficiency compared to AAV2 in other intracerebral neural circuits[314]. In our case, we found that the rAAV2 was less efficient then AAV2 at infecting RGCs (Fig. 5.4B) but more efficient at infecting cells in the deeper layers of the retina (Fig. 5.4D). As such, we decided to use AAV2 for future experiments requiring viral targeting of RGCs.

5.2.2 Identifying novel proteins exhibiting protein transfer movement

Having identified a suitable *in vivo* experimental system to visualise interneuronal protein movement, we now wanted to implement into this system the ability to screen for novel interneuronally transferred proteins. In order to do this we needed a method of labelling all the proteins of a donor cell in an unbiased manner, so that if any protein were to be transported intercellularly into an acceptor neuron we would be able to identify it as originating from a donor cell. We investigated two methods of cell specific proteomic labelling.

1. By expression of a promiscuous biotin ligase enzyme that indiscriminately biotinylates proximate proteins within a 10nm radius (BioID2)[315]. This is most commonly used to detect protein–protein associations in subcellular locations by fusing BioID2 to other proteins for subcellular targeting. However, in our case as we would like to biotinylate indiscriminately all proteins within the cell, we will express BioID2 freely in the cytoplasm of donor cells. The BioID2 system has been used successfully both *in vitro* and in mice *in vivo*[316].

2. By expression of a mutant murine methionyl-tRNA synthetase[317] (MutRS) which incorporates the noncanonical methionine analogue azidonorleucine (Anl) to elongator

$tRNA^{Met}$ (Fig. 5.7A). In this way, proteins synthesised in donor cells that express the MutRS will incorporate azide tags which can be labelled using click chemistry for fluorescence imaging or enriched on an affinity resin to facilitate identification by mass spectrometry (Fig. 5.7B). The method does not require expression of orthogonal tRNAs or depletion of canonical amino acids. In addition, this method has been used previously *in vivo* to enable cell-specific proteomic mass spectrometry analysis[318].

5.2.2.1 *In vivo* BioID2 enables visualisation of interneuronal transfer of endogenous mouse RGC proteins through visual circuitry

For the top-down unbiased approach, we wanted to identify all the proteins being exchanged intercellularly between neurons in an unbiased manner. To this aim, we introduced the system of promiscuous biotin ligation[315,316] into mouse RGCs to observe the transfer of the biotin-tagged donor proteome through the mouse visual circuit. As a preliminary test of this system, we designed a donor construct which expressed both cytoplasmic GFP and a promiscuous biotin ligase (BioID2). We first transfected this construct into neuro2A cells in culture to test if the free cytoplasmic expression of BioID2 would enable efficient unbiased biotinylation of all cellular proteins. We found that the addition of 50-100 μM biotin resulted in a clear increase of biotinylated protein content in GFP+ transduced cells within 24 hours (Fig. 5.5). In addition, we verified that there was no particular subcellular localisation of biotin labelled proteins following the expression of BioID2 from our donor construct. This was important as we wanted to use this method to label all cellular proteins in donor cells the most unbiased manner possible.

Following this *in vitro* test, we then produced a recombinant AAV vector encoding the same donor construct which expresses both cytoplasmic GFP and BioID2 for experiments of intercellular donor protein transfer *in vivo*. As the mouse visual system has repeatedly been shown to be an efficient system for the visualisation of interneuronal protein transfer from RGCs to LGN neurons projecting to the visual cortex, we injected this AAV intraocularly into the right eye of adult mice (Fig. 5.6A). 3 weeks following injections when transduced RGCs were likely to be expressing high levels of rAAV inserted BioID2, we administered biotin into the mouse to initiate biotinylation of donor cell proteins. We administered biotin both intraocularly (enabling efficient access into donor cells) and intraperitoneally (which enabled longer-term repeated daily biotin administration). Following 8 days of biotin administration to allow for donor protein labelling and intercellular protein transfer, mice were sacrificed and coronal sections taken in the posterior half of the brain to visualise the LGN and visual cortex.

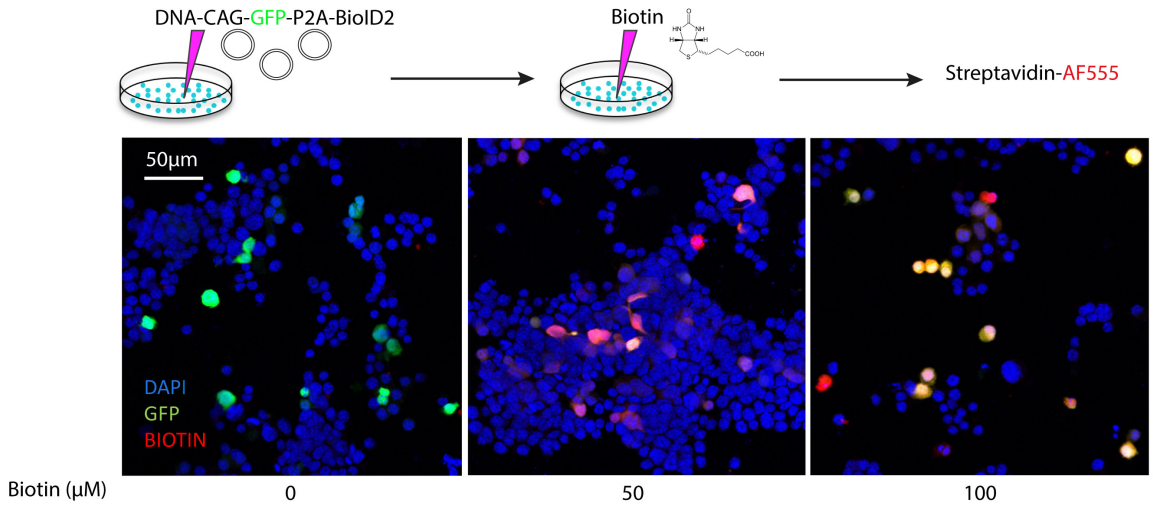


Fig. 5.5 BioID2 enables fast biotinylation of cellular proteins *in vitro* Mouse neuro2A cells were transfected at 80% confluency with a DNA expression plasmid encoding GFP and biotin ligase (BioID2). The following day, the culture medium was exchanged with fresh medium with 0, 50 or 100 μM biotin. 24 hours later, cells were fixed and stained with Streptavidin-AF555 to visualise biotinylated proteins.

By design of the donor construct, successfully transduced RGC donors express both a cytoplasmic GFP marker and BioID2. As RGCs send major projections into the LGN, we were able to observe GFP+ axon terminals within the LGN demonstrating successful RGC transduction following intraocular injection (Fig. 5.6B). More strikingly, the visualisation of biotinylated proteins using a streptavidin-HRP induced TSA Plus Cyanine 3 system (red) revealed a distinct pattern of staining within the LGN that was different to that of GFP staining. Whilst the GFP staining localised to show strong staining of LGN projecting RGC axons, the staining of biotinylated proteins was present not in the axonic projections of RGCs but within the more medial cell bodies of LGN neurons. More strikingly, we were then able to observe the specific staining of biotinylated proteins in the contralateral visual cortex (Fig. 5.6C). This strongly supported the interpretation that these donor-RGC derived biotinylated proteins had transferred intercellularly into LGN neurons which then projected to the visual cortex.

Much like the finding with the intercellular transfer of alpha-synuclein through the visual pathway, we noted that biotinylated proteins were visible only in the second order neurons of the pathway but not in axons of LGN projecting RGCs themselves. This suggested that intercellular protein transfer from RGC axon terminals into LGN cell bodies was an efficient process which did not result in the build up of biotinylated donor proteins in the axon terminals. However, if we looked at the axonic projections of RGCs in the superior colliculus,

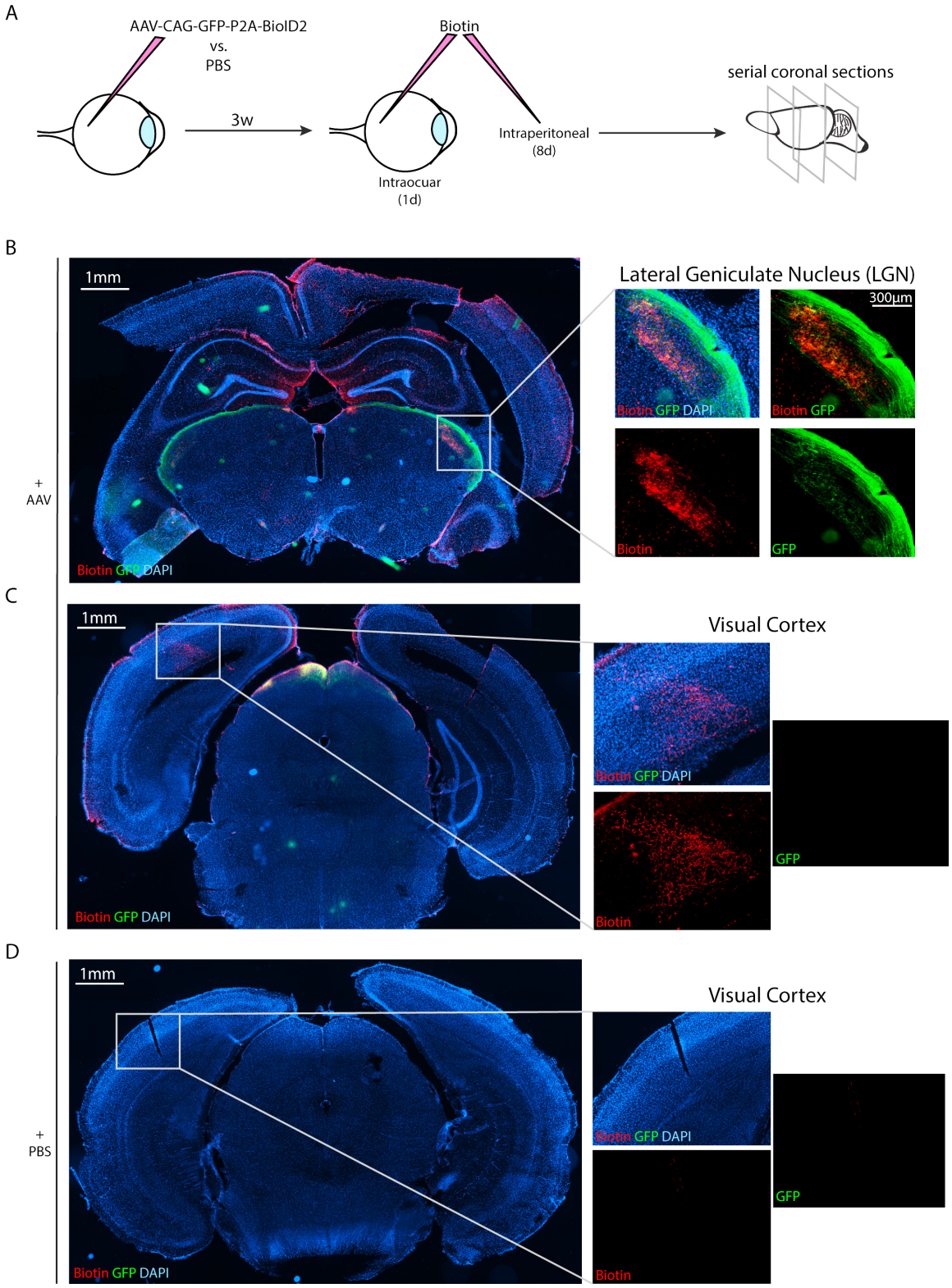


Fig. 5.6 *In vivo* BioID2 enables visualisation of interneuronal transfer of endogenous mouse RGC proteins through visual circuitry (A) A flowchart of the *in vivo* BioID2 protocol. Whole coronal section imaging of the LGN **(B)** and visual cortex **(C)** in AAV infected mice. **(D)** Whole coronal section imaging of PBS injected control mice subject to the same biotin administration protocol.

we were able to see biotinylated proteins within the first order axonic terminals (Fig. 5.6C). This difference may be due to the fact that whilst only 30-40% of RGCs project to the LGN, 90% of RGCs project to the SC. Therefore, our method of detection of biotinylated proteins may not be sensitive enough to visualise biotinylated proteins in LGN projecting axonic tracts, but enable visualisation of higher volume SC projecting axonic tracts.

In summary, this result strongly supported the case for the use of this BioID2 system *in vivo* to enable further top-down investigation into the volume and identities of endogenous RGC proteins moving interneuronally into LGN neurons.

5.2.2.2 Investigation of intercellular protein transfer *in vitro* using the MutRS system

The second method of unbiased total donor cell protein labelling that we investigated was by expression of a mutant murine methionyl-tRNA synthetase[317] (MutRS). This enables the incorporation of the azide-tagged methionine analogue azidonorleucine (Anl) into all newly synthesised proteins of cells expressing the MutRS enzyme. Therefore, MutRS expressing cells could be established as donor neurons and mixed with wild type acceptor neurons which could be investigated for the presence of azide-labelled donor-originating proteins. Azide-tagged proteins can be labelled using click chemistry for fluorescence imaging or enriched on an affinity resin to facilitate identification by mass spectrometry (Fig. 5.7B). To this aim, we decided to create an *in vitro* model of donor and acceptor neurons using the neuro2A mouse neuroblastoma cell line. Therefore we produced a stable donor cell line expressing both a mCHERRY marker and the mutRS for azide incorporation into all newly synthesised donor proteins

5.2.2.2.1 Maximising AnI incorporation into newly synthesised donor proteins In order to maximise the detection of moving labelled proteins from donor to acceptor, we first wanted to find the optimal conditions for the fast incorporation of AnI into newly synthesised proteins in donor cells expressing MutRS. Undifferentiated donor neuro2a cells were grown in 1 or 2mM AnI medium and collected over a period of 24 hours for analysis. AnI incorporation was measured by AF488 signal intensity following the chemical labelling of azide proteins with AF488-alkyne. We covalently attached the AF488 fluorophore specifically to azide-labelled proteins via click chemistry for assessment of the speed and degree of AnI incorporation. AnI incorporation was significantly faster at 2mM and was seen to increase up till the last 24 hour time point (Fig. 5.7C,D). Following this result, we decided to quantify AnI incorporation into donor cells beyond the 24 hour time point for a period of 6 days. We found that continued exposure of donor cells to 2mM AnI beyond 2 days resulted in cell

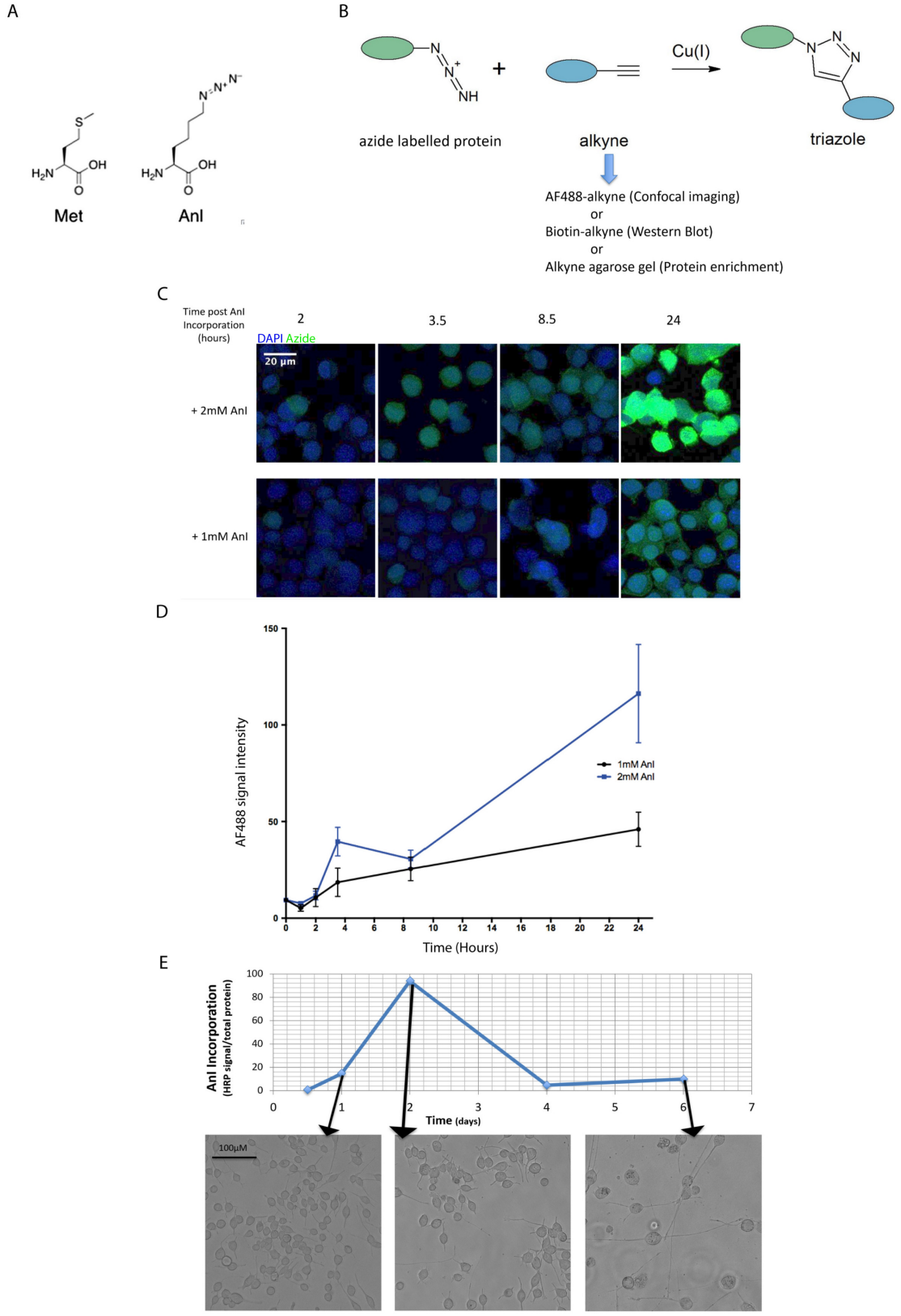


Fig. 5.7 Time course of AnI incorporation into newly synthesised donor proteins Legend on following page...

Fig. 5.7 Time course of AnI incorporation into newly synthesised donor proteins (A) Chemical structures of canonical amino acid methionine (Met) and the non-canonical methionine analogue azidonorleucine (AnI). **(B)** Use of click chemistry to covalently modify azide tagged donor proteins. **(C,D)** 24 hour time course of AnI incorporation into newly synthesised donor proteins. Azide incorporation was quantified by assessing AF488 signal intensity following click chemistry of azide tagged proteins with AF488-alkyne. **(E)** AnI incorporation (Azide specific HRP signal/total protein) over 6 days. Images show evidence of AnI-induced cytotoxicity beyond the 2 day time point.

death and a decrease in apparent AnI incorporation as defined by AnI+ proteins/total cellular protein(Fig5.7E). Given the high incorporation of AnI into cellular proteins by the 48 hour time point, we planned that for future experiments we would label donor proteins using a short 24 hour pulse of 2mM AnI.

5.2.2.2.2 Investigation of intercellular transfer of azide labelled donor proteins *in vitro* Having selected a relatively efficient AnI incorporation protocol, we wanted to verify that only the donor cell line expressing mutRS, but not wild type acceptor cells, could incorporate AnI into newly synthesised proteins. We therefore incubated the donor neuro2a cell line and the WT acceptor neuro2a cell line separately in 2mM AnI culture for 24 hours and assessed AnI incorporation. Only mCHERRY positive donor cells were found to have incorporated AnI into their cell proteins when imaged by confocal microscopy following click reaction with AF488-alkyne (Fig. 5.8A). In addition, donor cells grown in culture medium without any AnI were found to have no AF488 signal thus supporting the azide-specificity of the click reaction. The specificity of AnI incorporation only into donor cells was also verified by mass spectrometry (Fig. 5.9).

In order to investigate intercellular protein movement, we then carried out a preliminary experiment where donor and acceptor neuro2a cells were mixed at a 1:1 ratio and co-cultured for 24 hours in medium containing 2mM AnI. Post-fixation, azide-containing proteins were chemically reacted to AF488-alkyne by click chemistry and donor cells were stained with anti-mCHERRY antibody (red). This rendered all donor cells expressing the mutRS red and azide-labelled proteins green (Fig. 5.8B). By confocal imaging, we were able to see acceptor cells that did not express mutRS (red) but contained azide-labelled protein (green). From a total of 4500 non-red acceptor cells, 27 were found to be AF488 positive. Without cell mixing, 0% of acceptor cells were found to be AF488 positive using the same imaging and analysis settings. Given the specificity of azide incorporation and the click reaction, the visibility of green signal in non-red WT acceptor cells following cell mixing suggested that azide labelled proteins synthesised in donor cells had somehow moved into the acceptor cell.

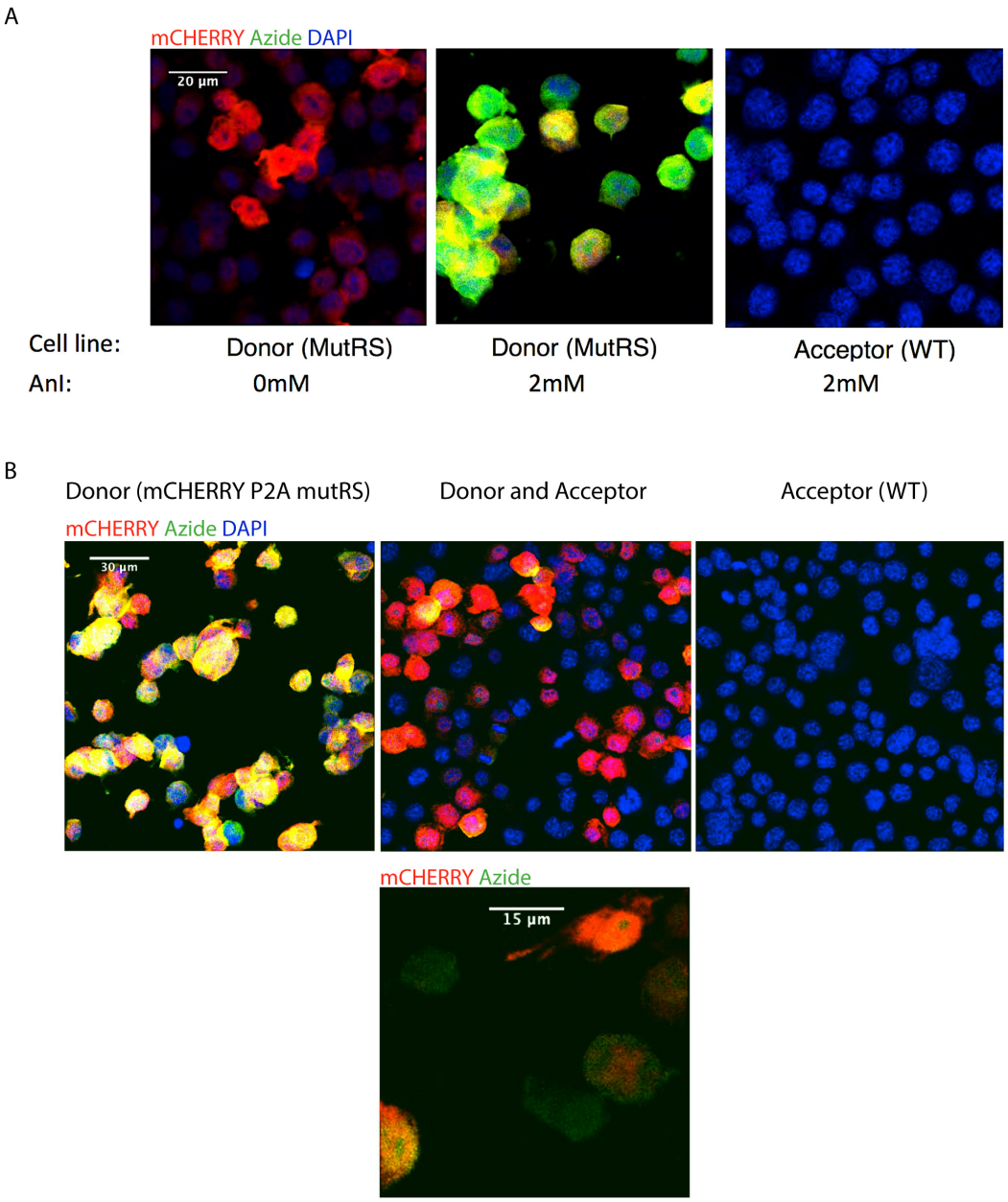


Fig. 5.8 **(A)** Neuro2Acells were grown with or without 2mM AnI added to the culture medium for 24 hours pre-fixation. Donor cells stably express the CAG-mCHERRY-P2A-mutRS construct. Post-fixation, azide positive proteins were chemically labelled with AF488-alkyne and donor cells stained with anti-mCHERRY antibody (red). **(B)** Confocal imaging of donor and/or acceptor cells following 24 hour incubation in 2mM AnI medium.

5.2.2.2.3 Identification of azide-tagged donor proteins by mass spectrometry Using this method of unbiased cell protein labeling, we eventually wanted to quantify and identify the proteins being transferred intercellularly from donor cells into acceptor cells by mass spectrometry. We predicted that the ratio of azide labelled donor proteins to unlabelled acceptor proteins would be extremely low in acceptor cells. Therefore, we needed to develop a very sensitive and stringent azide-enrichment protocol. We therefore considered the need for optimal conditions for efficient click chemistry during the attachment of azide-tagged proteins to alkyne resins, and stringent washes following the click chemistry to decrease non-specific binding (Fig. 5.9A). We found that by using commercial kits with stringent washing conditions (as described in the methods section), we could enrich and identify a total of 1527 azide-tagged proteins in donor cells incubated for 24 hours in 2mM AnI medium as opposed to only 196 proteins in WT acceptor cells incubated for 24 hours in 2mM AnI medium. This was a reasonably low level of background using a preliminary commercial protocol which we presume could be optimised further by altering some of the experimental conditions. More interestingly, we found that of the many azide-tagged proteins identified in donor cells, many were neurodegeneration related proteins such as Tau, EWSR1 and Fus. This was a notable result as we were previously unaware that these proteins were expressed by neuro2A cells in culture and indicated that such proteins, which have been shown to demonstrate intercellular transfer in human disease, could be labelled to test for intercellular protein movement using this system.

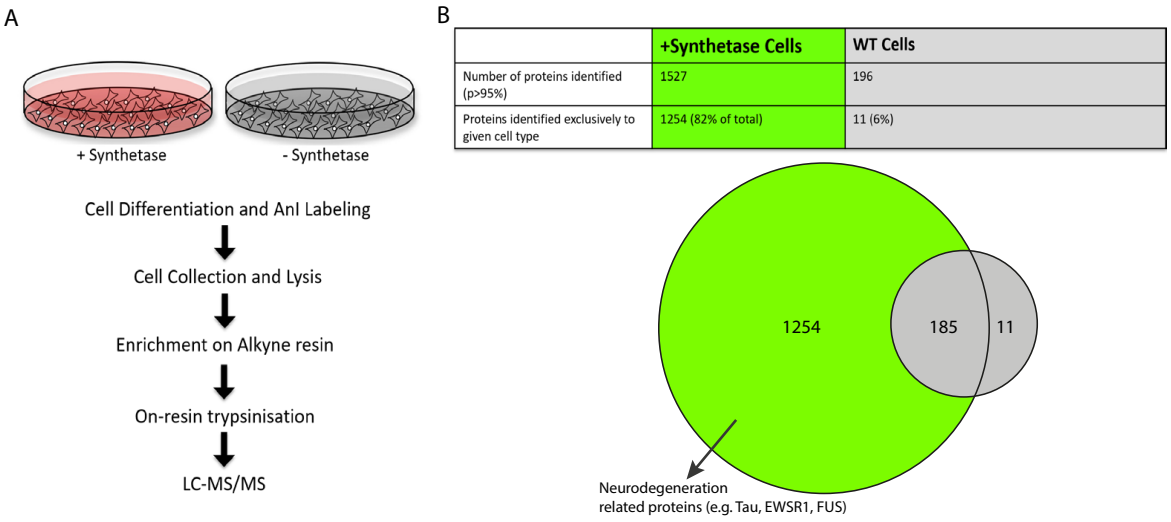


Fig. 5.9 Identification of azide-tagged donor proteins by mass spectrometry (A) Flowchart of steps required for selective identification of azide-tagged proteins from total cell protein preparations by mass spectrometry **(B)** Table and venn diagram showing the numbers of proteins identified in +Synthetase donor cells or WT cells incubated for 24 hours in 2mM culture medium.

5.2.2.2.4 *In vitro* differentiation of secondary neuronal cell lines As the ultimate aim of this project is to study the transfer of proteins between neurons rather then neural precursor cells, we also worked to establish an efficient protocol to differentiate the mouse neuroblastoma neuro2A cell line into into neurons *in vitro*. There are various methods of producing *in vitro* cultures of neuronal cells, some of which are more laborious then others and recapitulate the properties of neuronal cells *in vivo* to different degrees. As a general rule, cells defined as "neurons" in culture should have neuronal morphology, express neuron specific markers and also be electrically excitable[319]. Primary neuronal cell lines are thought to best represent neurons *in vivo* as they are produced directly from dissected nervous tissue from live animals. However, the preparation and culture of primary neurons can be challenging and only yields a limited number of cells as cultured neurons are not immortal[320]. In contrast, the differentiation of neuronal cultures from induced pluripotent stem cells (iPSCs), which is especially useful in the modeling of human disease, can take up to 3-5 weeks[321].

One particular difficulty of culturing neuronal cells is that mature neuronal cells do not undergo cell division and so are not amenable to amplification in cell culture. Secondary neuronal cell lines derived from neuronal tumours are much easier to use for *in vitro* studies as they can be grown easily in cell culture, amplified to give more homogenous and abundant cell populations for experiments and can relatively easily and quickly be differentiated into neurons by manipulation of culture conditions. They can be differentiated by serum deprivation and/or treatment with environmental factors such as cAMP analogs, retinoids and neurotrophic factors (e.g. nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF) or glial cell-derived neurotrophic factor (GDNF)). The disadvantage of secondary cell lines is that they are tumor derived and have most likely accumulated many physiological differences compared to true neurons *in vivo* during passaging after initial collection[322].

However, to our intents and purposes, we decided that the differentiation of secondary cell lines would be a sufficient strategy to investigate interneuronal protein transfer through cell mixing experiments. The mouse neuro2A cell line is a fast-growing mouse neuroblastoma cell line that has been used extensively as an *in vitro* model of neurons to study neuron differentiation, neurite growth and synaptogenesis. The neuronal differentiation of neuro2A cells can be induced by serum deprivation and/or addition of factors such as transforming growth factor1 (TGF1), bone morphogenic protein 4 (BMP4), glial cell-derived neurotrophic factor (GDNF), dibutyryl cyclic adenosine monophosphate (dbcAMP) or retionic acid (RA)[322,323]. Neuro2A cells respond very quickly to serum deprivation and differentiating factors by activation of ERK signalling pathways[324–326] resulting in neurite growth and expression of

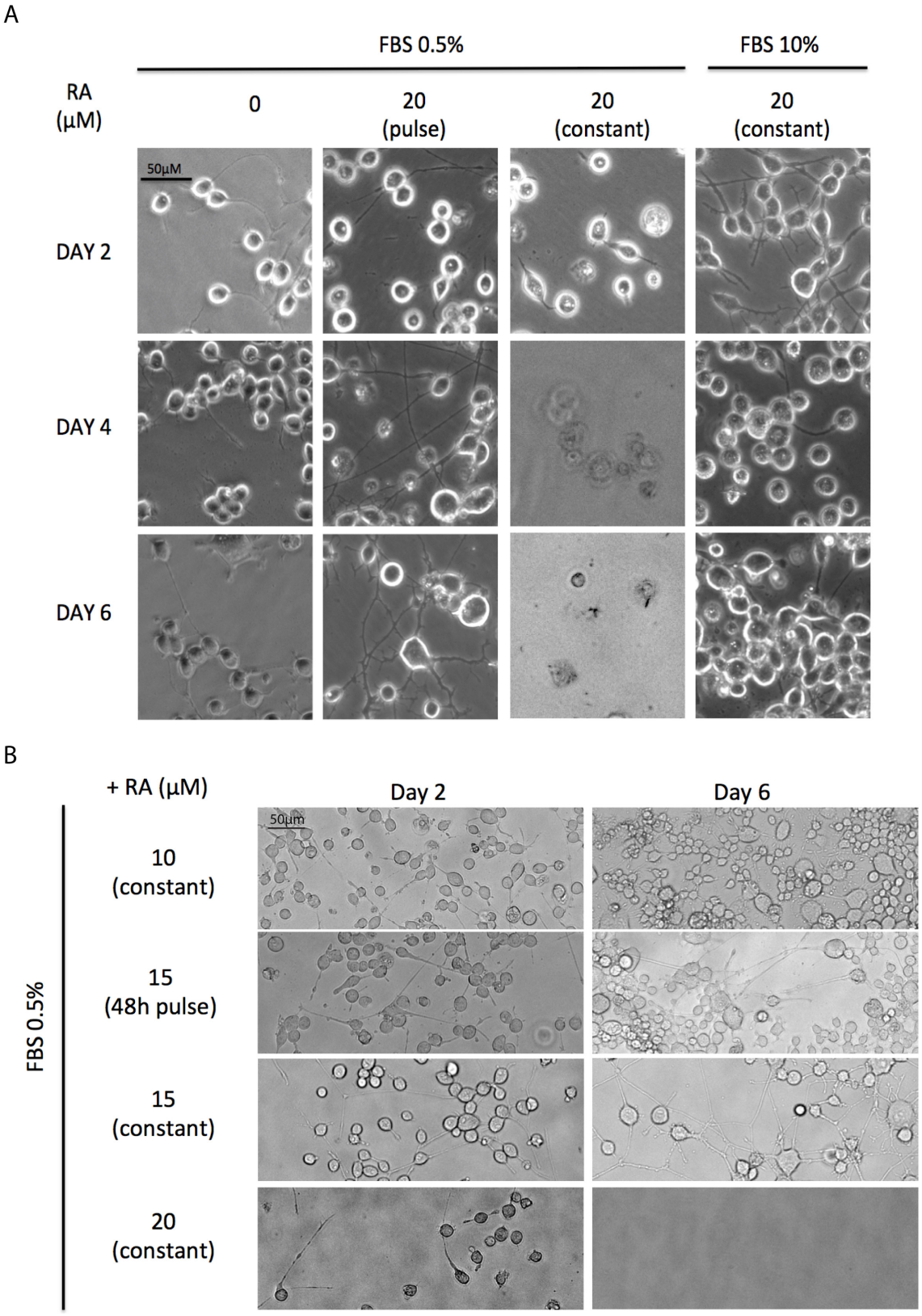


Fig. 5.10 Differentiation of mouse neuro2A cells Mouse neuro2A cells were seeded at 40% confluency and differentiated by serum deprivation (FBS 0.5%) and/or addition of retinoic acid at specifiied concentrations. **(A)** Neuro2A cells differentiated with serum deprivation (left), serum deprivation and 48h or constant addition of RA (middle two insets) or just with RA added to usual 10% medium (right). **(B)** Investigation into different concentrations and lengths of RA addition to serum deprived media and its effects on neuro2A neuronal differentiation. We find that constant 15 μM RA addition gives optimal parameters for quick and effective neuronal differentiation without observable toxicity.

neuron specific markers such as NeuN and choline acetyltransferase (ChAT)[323]. Differentiated neuro2A cells can form randomly connected neuronal populations in cell culture[327,328] or form pre-designed neural networks on micropatterned neurochips that display functional excitability and interneuronal communication[329].

We attempted to test two published methods of neuro2A cell differentiation to assess which would be most suited to our purposes: serum deprivation and/or retinoic acid (RA) induction (Fig. 5.10). We found that either serum deprivation or addition of 20M RA resulted in neurite growth within 2 days (Fig. 5.10A). A combination of both a pulse of 20μM RA addition and serum deprivation resulted in a higher density of neurite growth in cell culture compared to either of the methods in isolation. However, the constant addition of RA with serum deprived media resulted in cell toxicity. Subsequently, we attempted to find a suitable RA concentration and delivery length within serum deprived medium for efficient neuro2A differentiation. We found that constant addition of a lower concentration of 15M RA resulted in long neurite processes in the majority of cultured cells with minimal cell toxicity. Thus, we decided to use these conditions for our future cell mixing experiments with the MutRS expressing neuro2A donor cell line (Fig. 5.10B).

5.2.3 Candidate based study of the mechanisms behind interneuronal protein movement

In addition to conducting an unbiased screen for the identification of novel interneuronally moving proteins, we also aimed to investigate the details of the mechanism of interneuronal protein transfer using a candidate based approach. In this instance, we aimed to design a molecular tag which could be appended to known interneuronally moving proteins and would enable the selective manipulation of donor or acceptor proteins.

5.2.3.1 *In vivo* protein transfer of V5 Tau through mouse hippocampal circuits

To test the possibility of a candidate protein based approach, we decided to investigate the interneuronal movement of the candidate protein tau through canonical hippocampal circuits. This was for a variety of different reasons. First, tau has already been shown to move interneuronally through the mouse hippocampal circuit both in WT and mutant aggregate form[119,172,173]. Second, the synaptic connectivity and circuitry of the hippocampus is very well established, as such increasing the confidence with which we can deduce specific trans-synaptic protein transfer. Third, the hippocampal memory and learning neural circuits have extensively been shown to demonstrate synaptic plasticity with the expression of the

known trans-synaptic retroviral gag like protein Arc. For example, after exposure to a novel environment, there is a brief burst of Arc transcription in hippocamapal memory networks which quickly returns to baseline levels within minutes to an hour[184]. Given Arc's crucial role in synaptic plasticity and learning, and given the ability of Arc to travel inter-cellularly by forming capsid like structures[170,187], we considered that the hippocampus may potentially be a promising circuit within which to study other candidate proteins such as Arc in the future.

The CA1 region of the mouse hippocampus is relatively easy to target through stereotaxic intracerebral injections due to its relatively large coverage of all anterior-posterior, medio-lateral and dorso-ventral axes. In addition, its connectivity within the hippocampal circuitry and its major projection to the subiculum is very well established. As a preliminary test, we designed a lentiviral vector that expressed a iCRE-GFP marker and V5 tagged human wild type 0N4R Tau. In this design, the donor neuron could be identified by expression of the nuclear GFP marker whilst the expression of V5-Tau expressed from the donor neuron and any subsequent interneuronal transport of V5-Tau could be visualised by staining for the V5 tag. We also designed the donor neuron to express a CRE recombinase so that successfully transduced donor neurons could be selectively re-targeted for future experiments if this protocol was successful. For example, if interneuronal protein movement could be visualized from the donor neuron, we could investigate whether this movement was activity dependent by retargeting the donor neurons with a AAV vector encoding flexed DREADDs (Designer Receptors Exclusively Activated by Designer Drugs). In another instance, we could re-target donor neurons with a flexed TVA and oG encoding AAV vector for subsequent retrograde trans-synaptic rabies virus tracing, allowing us to unambiguously correlate any retrograde interneuronal protein movement with synaptic connectivity *in vivo* at the cellular level.

As a preliminary test of this approach, we injected the lentiviral vector into the CA1 of adult WT mice. CA1 neurons are known to project to the subiculum within the canonical hippocampal circuitry[330]. Subsequently, subicular neurons project to the entorhinal cortex (Fig. 5.11A). 1 month p.i. we were clearly able to see the expression of GFP and V5-Tau from transduced donor CA1 neurons (Fig. 5.11B). When looking into the first and second order projection sites of CA1 neurons to test for interneuronal protein movement, we were also able to see weaker V5-Tau staining both in the subiculum and entorhinal cortex (Fig. 5.11C). This specific pattern of V5-Tau staining strongly suggested that there had been inter-neuronal transfer of V5-Tau from CA1 donor cells into the second order subicular neurons which then projected to the entorhinal cortex. Although further experimental repeats are required, this result strongly suggests that the hippocampal circuitry can also be targeted

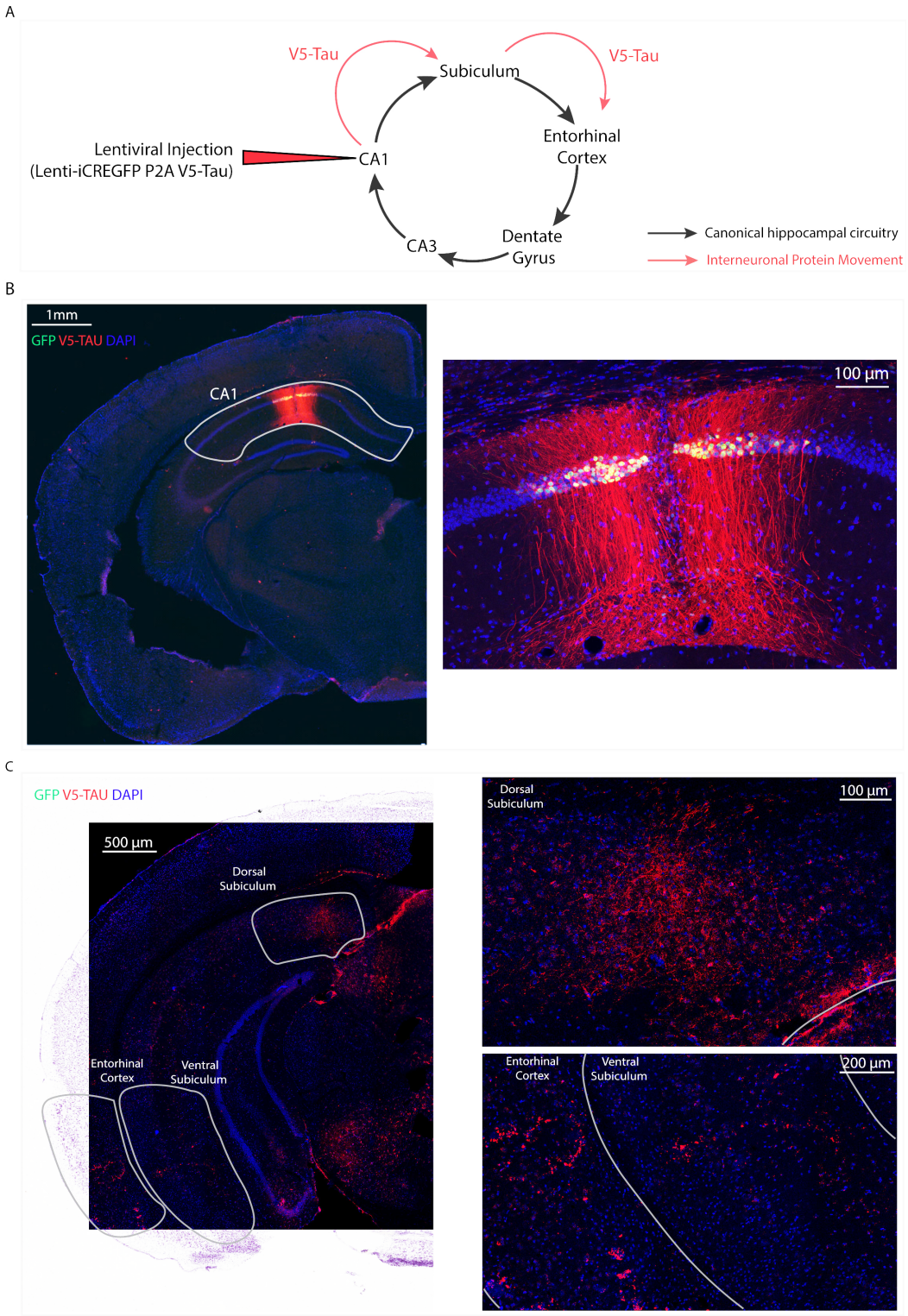


Fig. 5.11 Movement of V5-Tau through the hippocampal circuitry *in vivo* **(A)** Schematic of the canonical hippocampal circuitry and hypothesised movement of V5-Tau through the circuit following lentiviral transduction of CA1 donor cells. Lentivirus encoding iCREGFP and V5Tau was injected into CA1 of adult mice. One month post injection, brain sections were stained for GFP (green, donor neuron), V5-Tau (red) and DAPI (blue, nuclei). **(B)** An anterior section shows the expression of nuclear GFP and V5 Tau in donor CA1 neurons transduced by the lentivirus. **(C)** A posterior section shows the movement of V5 Tau from CA1 neurons into the subiculum and entorhinal cortex.

for the investigation of interneuronal protein transfer, albeit at longer time-frames (1 month) than for the visual pathway (4 days).

5.2.3.2 Design of a preliminary split-CRE design to investigate circuit logic of intercellular protein spreading

Having observed that the mouse hippocampal circuit could be used to study the interneuronal movement of tau, we wanted to develop a small split CRE tag that we could add onto tau to enable further investigation of the mechanisms enabling this process. This was stated as Aim 3 in the introduction of this chapter and the purpose of this aim was to A) enable amplification of the signal of interneuronal protein movement by reconstitution of CRE recombinase in the acceptor neuron and B) genetically manipulate acceptor neurons for further mechanistic characterisation of the underlying biology behind protein transfer. For example, in this case we could conduct retrograde trans-synaptic tracing from the labeled acceptor neurons to unambiguously determine the trans-synaptic specificity of any protein movement in the anterograde direction from donor neurons.

To begin, we designed a preliminary split-CRE system as depicted in Fig. 5.12A. This design was derived from the known functional CRE split site between amino acids 59 and 60. It was previously shown that functional CRE activity could be reconstituted on co-expression of a short N terminal fragment of CRE (AA19-59) and longer C terminal fragment (AA60-343) in the same cell. This reconstitution persisted following the addition of rapamycin-dependent dimerisation domains onto the N terminus of both split-CRE fragments by short flexible linkers (F2)[331]. In our case, as we did not require ligand-induced heterodimerisation, we replaced the FKBP and FRB heterodimerisation domains used in this study with known heterodimerising coiled coil domains of the GABA B receptor subtypes 1a and 2 (GBR1acc and GBR2cc)[332]. We also added a 3XFLAG molecular tag onto the CCRE system for visualisation of acceptor cells. In order to add the NCRE fragment onto candidate proteins, we decided to use a long 43AA flexible glycine-serine (GS) linker with a TEV cleavage site. The rationale behind this was to then express TEV protease in the acceptor cell, enabling cleavage of the NCRE tag from the candidate protein and enabling effective CRE reconstitution on heterodimerisation with CCRE.

We first tested whether these new additions to the N termini of NCRE and CCRE fragments had impaired the ability of these fragments to functionally reconstitute CRE activity following co-expression in the same cell (Fig. 5.12B). As a negative control, we transfected cells with only the CCRE <mCH> plasmid. This plasmid expresses the CCRE fragment as well as encoding a flexed mCHERRY marker and TEV protease. As the

positive control, we co-transfected the CCRE <mCH> plasmid with CRE-expressing plasmid constructs. As expected, we observed CRE activity in these cells as measured by positive mCHERRY fluorescence. In order to test NCRE and CCRE reconstitution, we co-transfected the CCRE <mCH> plasmid with donor construct plasmids expressing NCRE appended to one of four different candidate proteins. We found that all 4 NCRE constructs enabled functional CRE activity when co-expressed with CCRE. This clearly demonstrated that the new modifications to the split CRE fragments had not impaired their ability to conduct DNA recombination.

The rationale for adding the NCRE tag to the 4 proteins Tau, Otx2, Gsad and GFP are as follows. Gsad-NCRE was designed as a positive control for protein movement as the G protein of the SAD-B19 rabies virus is known to transfer intercellularly *in vitro*. In contrast NCRE-GFP was the negative control as we did not expect GFP to move intercellularly *in vitro*. NCRE-Tau and NCRE-Otx2 were test constructs which we wanted to test for intercellular protein movement.

Following confirmation of effective CRE reconstitution following co-expression of NCRE and CCRE within the same cell, we then conducted cell mixing experiments in order to test the efficiency of this system for the visualisation of intercellular protein transfer (Fig5.12C). In these cell mixing experiments, we transfected donor cells with plasmids expressing a TVAeGFP marker and a NCRE tagged candidate protein. In contrast, acceptor cells were transfected with plasmids encoding CCRE, TEV protease and flexed mCHERRY on a separate cell culture plate. After transfection, donor and acceptor cells were mixed in culture following DNAseI treatment as described in the methods section. 3 days post-transfection we assessed the degree of mCHERRY activation of acceptor cells as a measure of intercellular protein transfer. This was because the CCRE fragment on its own cannot activate the flexed mCHERRY marker in acceptor cells. Therefore, increased mCHERRY activation could only be enabled following the reconstitution of CRE activity due to the movement of NCRE from donor cells into CCRE-expressing acceptor cells. In these experiments, donor neurons expressing Gsad-CRE were used as a positive control for intercellular protein transfer (Gsad) and activation of flexed mCHERRY (CRE) in acceptor cells. The results are depicted in Fig. 5.12D. One can observe that if CCRE <mCH> acceptor cells are mixed with GFP and Gsad-CRE expressing donor cells, there is a notable activation of mCHERRY in acceptor cells. However, none of the NCRE-candidate protein expressing donor cells caused activation of mCHERRY in acceptor cells. In particular, the observation that Gsad-CRE but not Gsad-NCRE donor cells activated acceptor cells indicated to us that the NCRE design was somehow reducing intercellular protein movement, or more likely, the split CRE system

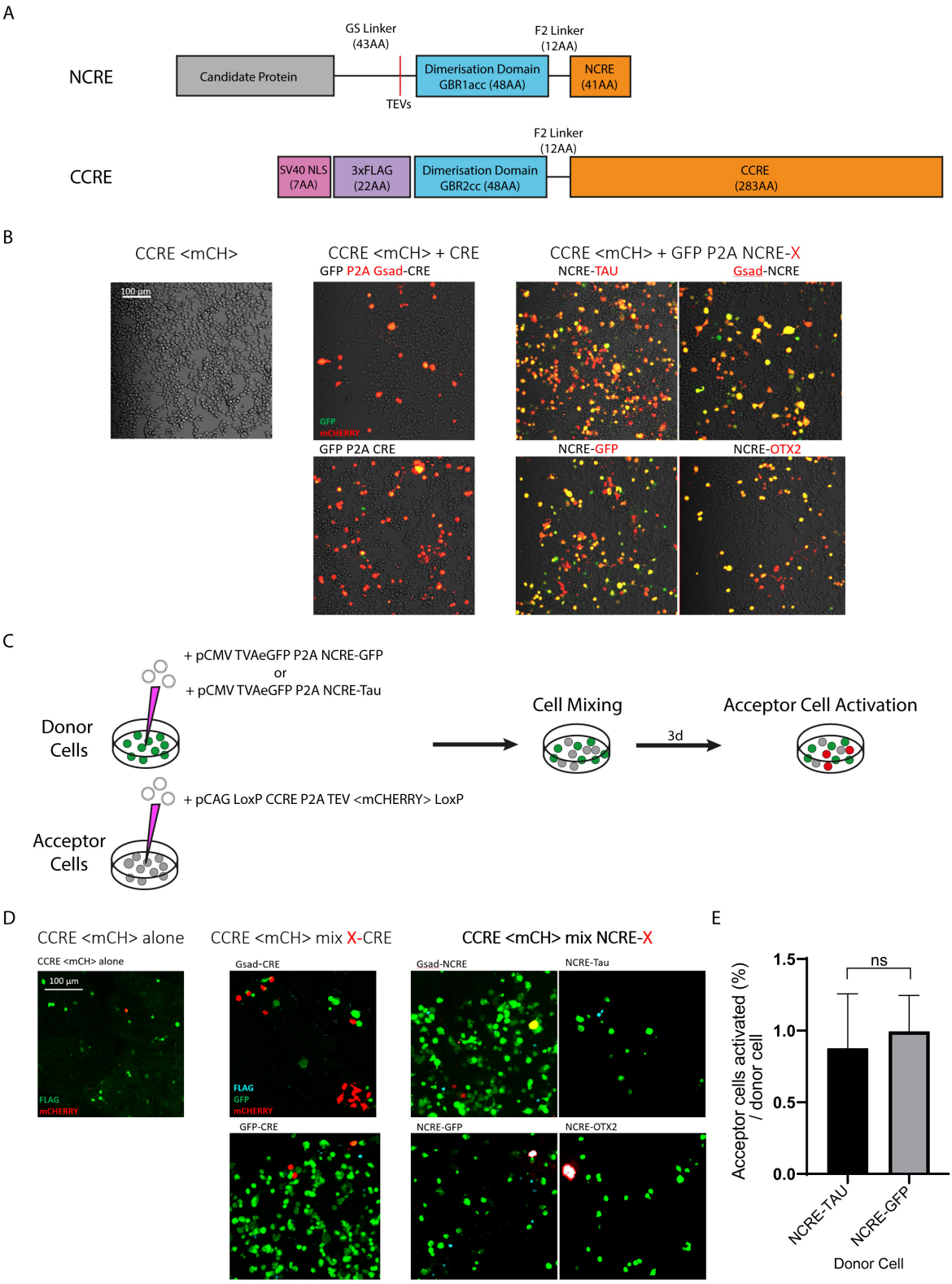


Fig. 5.12 ***In vitro*** **test of preliminary split NCRE/CCRE design (A)** Gene map of preliminary NCRE and CCRE designs **(B)** Transfection of CCRE <mCH> acceptor plasmid into neuro2A cells in isolation (left) or with CRE expressing plasmids (middle) or with NCRE-candidate protein (X) expressing plasmids. **(C)** Scheme of cell mixing experiments. mCHERRY activation of acceptor cells is taken as a measure of intercellular NCRE-candidate protein movement. **(D)** Confocal imaging of mixed cell culture 3 days post-mixing. Only acceptor cells mixed with Gsad-CRE expressing donor cells demonstrate clear mCHERRY activation. **(E)** Quantification of mCHERRY+ acceptor cells 3 days post-mixing. Two-tailed unpaired student's T test.

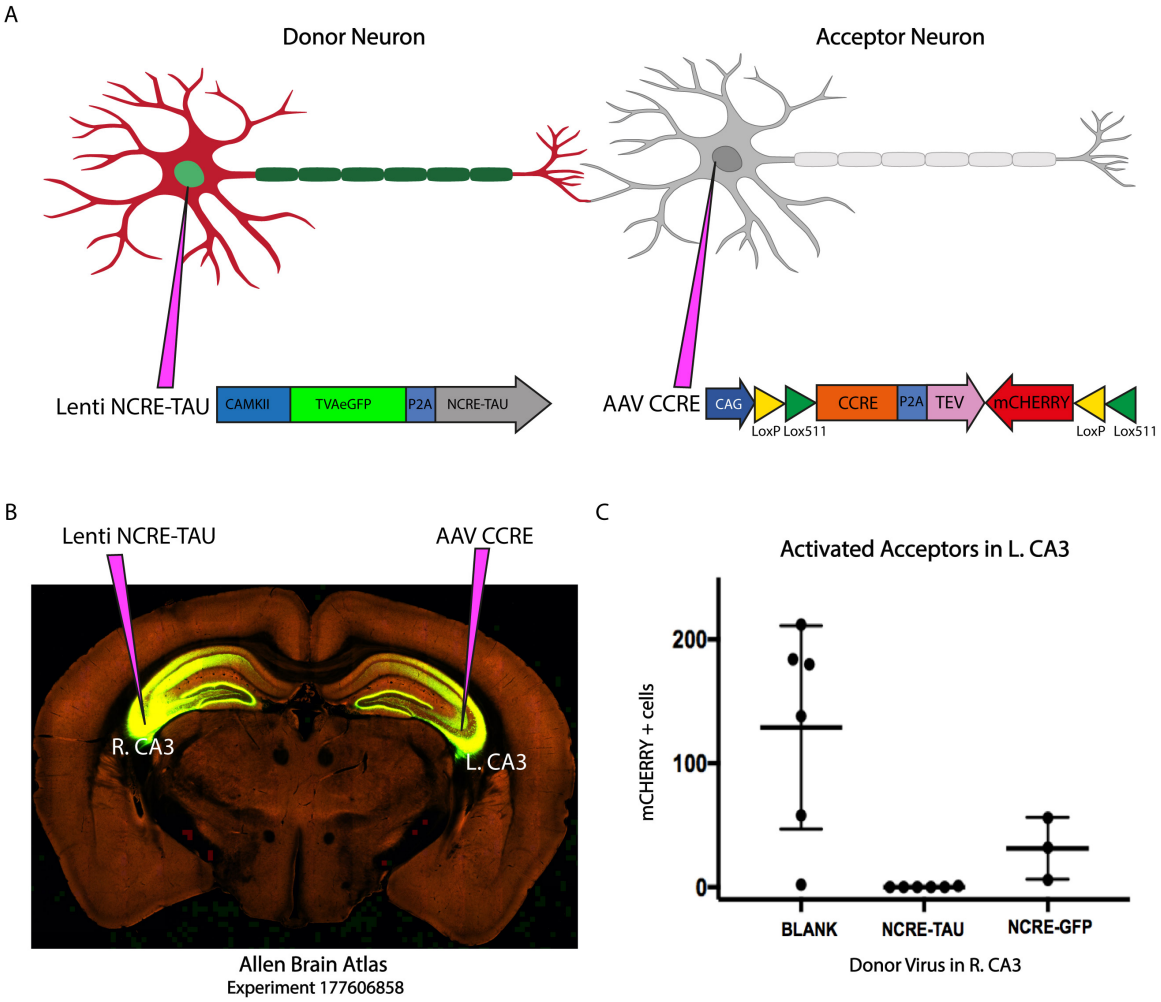


Fig. 5.13 ***In vivo*** **test of preliminary split NCRE/CCRE design (A)** Schematic representation of dual viral injections for the establishment of NCRE-Tau expressing donor neurons and CCRE expressing acceptor neurons *in vivo*. **(B)** Allen Brain Atlas, Experiment 177606858 showing the projections of R CA3 neurons infected with AAV-EGFP into the L CA3. **(C)** Quantification of mCHERRY+ acceptor cells in the left CA3 following injection of PBS, Lenti NCRE-Tau or Lenti NCRE-GFP into the right CA3. Results demonstrate no significant differences in mCHERRY+ cell numbers between all three experimental conditions.

was not efficient enough to reconstitute functional CRE activity following the low levels of intercellular transfer of Gsad-NCRE into CCRE expressing cells.

We also tested this preliminary split CRE system *in vivo* (Fig. 5.13A). First to trial the experimental feasibility of this approach using double viral injections *in vivo* and second to investigate whether longer term expression of NCRE-Tau from donor cells within neural circuits *in vivo* would result in higher levels of intercellular protein transfer and thus activation of CCRE <mCHERRY> acceptor neurons. To do this, we injected a lentivirus expressing the donor construct as described in the *in vitro* experiments into the right CA3 of adult mice. As there are known to be contralateral synaptic projections between CA3 regions (Allen Brain Atlas, Experiment 177606858) we then injected an AAV encoding the acceptor construct into the left CA3 (Fig. 5.13B). Quantification of mCHERRY+ acceptor neurons in the L CA3 showed that our preliminary split CRE design did not result in activated acceptor neurons *in vivo* (Fig. 5.13B).

Given these negative results, we decided to further modify and optimise the efficacy of our split-CRE system.

5.2.3.3 Design of multiple split-CRE variants to enhance the CRE reconstitution efficiency of the preliminary split-CRE system

Given the finding that Gsad-CRE but not Gsad-NCRE could activate mCHERRY in CCRE expressing acceptor cells *in vitro*, we concluded that the efficacy of NCRE/CCRE reconstitution needed to be improved such that functional CRE activity would be present with even very low levels of NCRE/CCRE dimers. With this aim, we rationally altered 4 aspects of the preliminary split-CRE design (Fig. 5.14A).

First, we decided to alter the nuclear localisation signal (NLS) on the CCRE fragment. In the preliminary design, it was thought in theory to increase CRE reconstitution efficiency by increasing the transport of NCRE/CCRE dimerised complexes to the nucleus. However, we also considered the possibility that nuclear translocation of non-heterodimerised CCRE to the nucleus may reduce the chances of CCRE dimerising with incoming NCRE molecules to enable efficient reconstitution in the first place. Therefore, we designed a CCRE construct which did not encode a NLS. On the other hand, as CCRE is translated in the cytoplasm one may argue that heterodimerisation with incoming NCRE should not be impaired significantly by the addition of a NLS given the continuous expression of CCRE in the cytoplasm acceptor cells. As such we also considered the opposite possibility that the NLS needed to be stronger in order to enable more efficient nuclear transport of any heterodimers that may form in the cytoplasm. Therefore, we also designed a CCRE construct where the

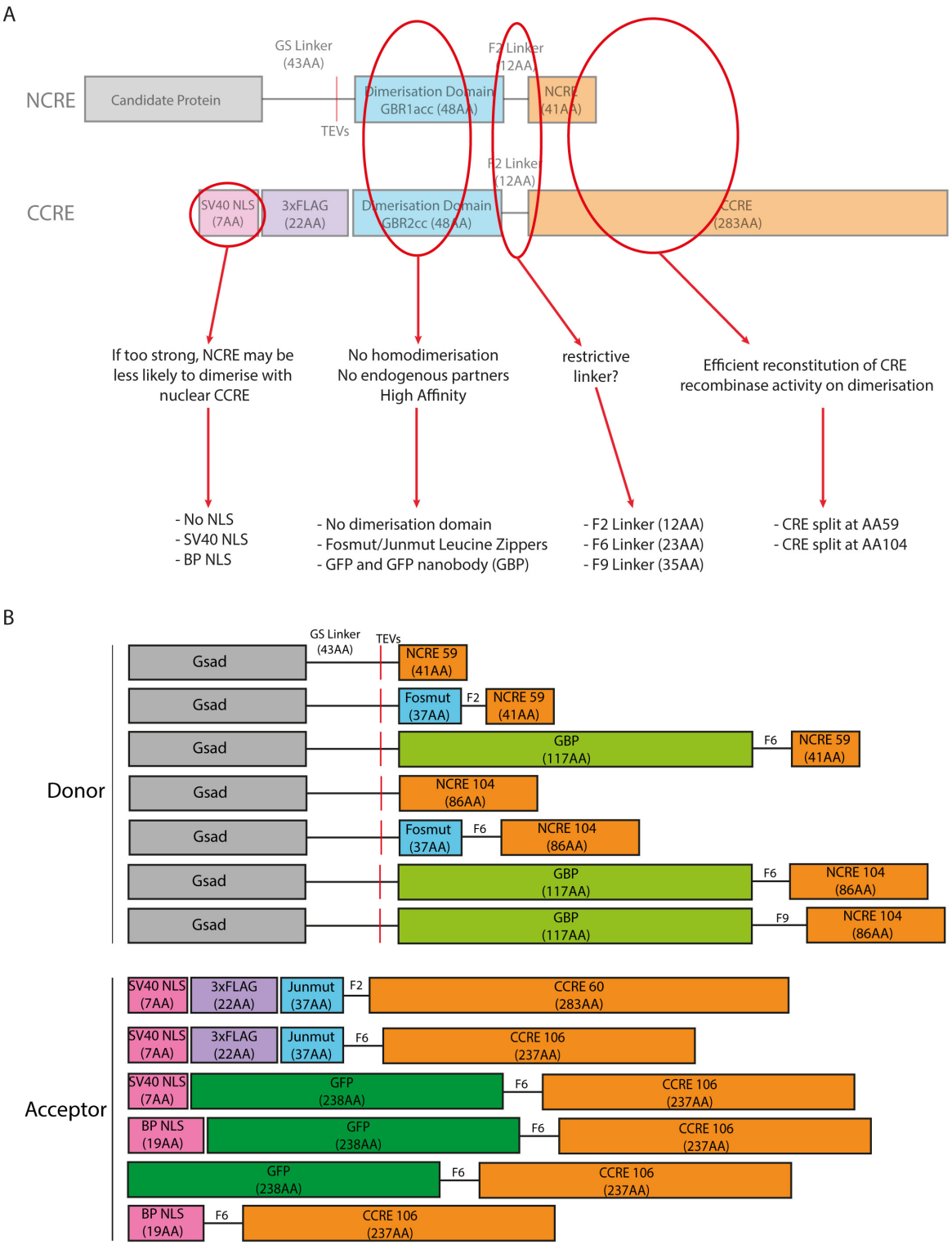


Fig. 5.14 Modifications to the preliminary SPLIT-CRE design (A) Gene maps of preliminary SPLIT-CRE design with superimposed red circles indicating the 4 points at which we decided to introduce adjustments. **(B)** Gene map of the alternative split-CRE designs that were cloned for *in vitro* tests.

SV40 NLS sequence was replaced with a stronger bipartite SV40 NLS (BPNLS - sequence KRTADGSEFESPKKKRKVE) which had previously been shown to cause greater nuclear localisation of proteins compared to the SV40 NLS sequence[333].

Second, we altered the heterodimerisation domain used in the preliminary split-CRE design. Our two main criteria in the search of new heterodimerisation domains was that A) the dimerisation domains had high affinity interactions and B) they did not have any obvious endogenous ligands which could interfere with the dimerisation of our split-CRE system. We especially noted that the heterodimers used in the preliminary design (the coiled coil domains of GABA B receptor subtypes 1a and 2[332]) could also bind endogenous GABA B receptors in neurons which could interfere not only with effective NCRE/CCRE dimerisation but also with intercellular protein movement mechanisms. As such we considered two alternatives. One was a mutated leucine zipper Fosmut and Junmut heterodimersiation system which had been developed by the complementary mutation of charged residues responsible for wild type Fos-Jun binding. This resulted in a mutated Fos and Jun which were still able to bind one another at high affinity but could not bind endogenous Fos and Jun proteins[334]. In addition, we considered the use of the GFP:GFP nanobody system which, by nature of being an antibody-derived system has both high affinity and specificity[188]. One caveat however was that the GFP nanobody (or GBP for GFP Binding Protein) is 117AAs, which is significantly bigger than the Fosmut: Junmut leucine zipper system which was only 37AAs in size.

Third, we considered altering the split site of the CRE protein. Although the AA59-60 split site used in the preliminary design has been shown to be functional, rational design of an alternative split site between AA104-106 has also been shown to have functional reconstitution ability[335]. In fact, the level of CRE recombination was actually found to be higher with the AA104-106 split site compared to the AA59-60 split site. One disadvantage of this alternative split site was that this would increase the size of the NCRE-tag which we wanted to keep as small as possible.

Fourth, we also considered increasing the length of the F2 linker on the N terminus of the N and CCRE fragments. It had been shown previously that especially for the new AA104-106 CRE split site that we had now incorporated, longer linker arms increased recombination activity[335]. We thought the need for longer linkers was also greater due to the increased size of some of our new heterodimerisation domains (e.g. the GBP). Thus we also attempted replacing the F2 linker with F6 or F9 linkers. These are flexible linkers similar in design to the F2 linker but 23 and 35AAs long respectively.

As such, we designed and cloned many combinatorial variants of these 4 different factors into the original split-CRE design as shown in Fig. 5.14B for further testing *in vitro*.

5.2.3.4 *In vitro* screening of different split-CRE designs

Following the design and cloning of multiple combinatorial variants of the original split-CRE design, we then wanted to assess whether any of the modifications made had actually impaired the ability of NCRE and CCRE fragments to reconstitute CRE activity when co-expressed in the same cell. Given the increasing size of the dimerisation domains, linker lengths and split CRE sites, we deduced that some of our modifications could potentially impair DNA recombination by steric hindrance. On co-transfection of various NCRE and CCRE pairs (Fig. 5.15A), we observed that the use of the GFP:GBP dimerisation system impaired DNA recombination resulting in an absence or vast reduction of mCHERRY activation in co-transfected cells. In contrast the Fosmut:Junmut dimerisation system seemed to enable efficient DNA recombination for either of the two CRE split sites. Interestingly, we also observed functional reconstitution of CRE activity on the co-expression of overlapping NCRE104 (AA19-104) and CCRE60 (AA60-343) fragments.

Given these results, we conducted a cell mixing experiment in differentiating neuro2A cells with donor cells expressing Gsad-Fosmut F6 NCRE 104 and acceptor cells expressing NLS Junmut F6 CCRE 106 <mCH> (Fig. 5.15B). We observed a possible increase of DNA recombination activity in mixed cells (right) compared to the non-mixed acceptor cells control (left). However, the system still appeared inefficient, especially when compared to the Gsad-CRE mixed positive control which resulted in a much larger number of more strongly mCHERRY positive acceptor cells.

5.3 Summary

In this final experimental chapter, we aimed to initiate investigations into the understudied phenomenon of interneuronal protein transfer. In order to do this we first tested different *in vitro* and *in vivo* systems that could be manipulated to establish protein donating "donor" neurons into a neural network. We found that both the mouse visual circuitry and hippocampal circuitry could be targeted relatively easily for this purpose. In both cases, we were able to obtain promising preliminary data showing intercellular protein movement into the axonic projection sites of second order neurons within the circuit.

Secondly, we investigated two independent methods of total proteome labelling to enable top-down identification of proteins demonstrating intercellular transfer in an unbiased manner. Most notably, we show that the expression of the promiscuous biotin ligase BioID2 in RGCs enables the biotinylation of RGC proteins which then appear to move intercellularly into

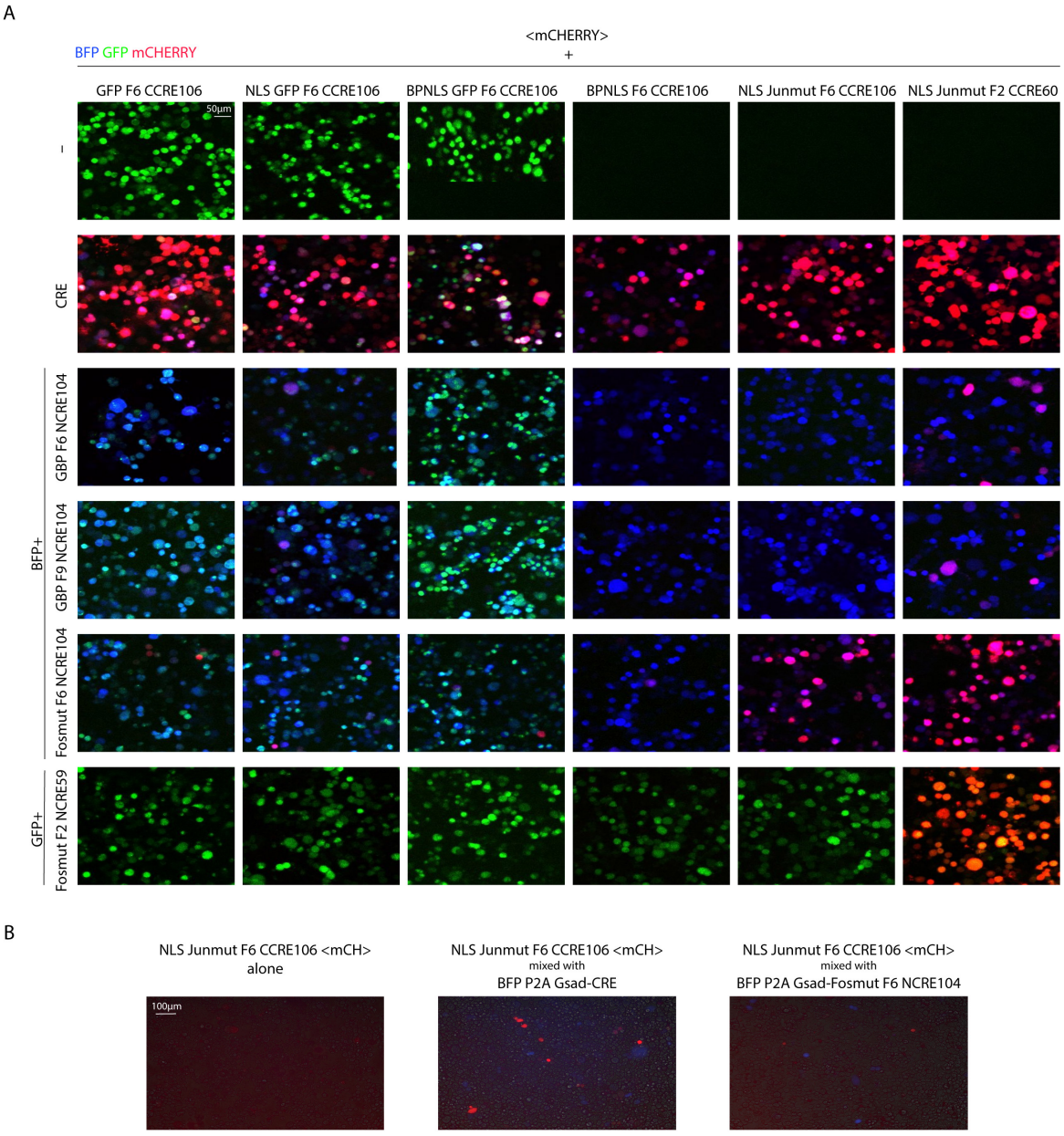


Fig. 5.15 ***In vitro*** **screening of multiple split-CRE designs** **(A)** Cotransfection of different NCRE:CCRE design pairs to assess efficiency of DNA recombination on co-expression in the same cell. The efficiency of DNA recombination is assessed by the number and intensity of mCHERRY+ cells. **(B)** Representative cell mixing experiment with the Fosmut F6 NCRE105 and NLS Junmut F6 CCRE106 split-CRE pair.

LGN neuron cell bodies and transport anterogradely into the visual cortex. This is an exciting result which can be investigated further to identify novel RGC endogenous proteins that demonstrate interneuronal protein transfer.

Thirdly, we aimed to develop a small split-CRE tag which would enable more sensitive visualisation of intercellular protein transfer and also enable genetic manipulation of protein-receiving "acceptor" neurons in the circuit. Although we were able to slightly improve the efficiency of the split-CRE tag through the screening of multiple designs, we were yet unable to obtain an efficiency of functional split-CRE reconstitution high enough to visualise the small amounts of NCRE-tagged proteins moving intercellularly.

Chapter 6

General Discussion and Conclusions

6.1 The design and *in vivo* characterisation of a second generation Self-Inactivating Rabies vector (SiR 2.0)

In the first experimental chapter I describe the design, production and *in vivo* characterisation of a second generation Self-inactivating Rabies virus (SiR 2.0 or SiR N2C). This new SiR 2.0 virus incorporates a C-terminal TEVs-PEST addition design to the N protein of the WT ΔG CVS-N2C rabies virus. SiR 2.0 therefore incorporates the already existing rabies self inactivation technology into a less immunogenic and more neuroinvasive rabies virus strain. By design, the improved SiR 2.0 vector abolishes WT rabies virus toxicity like the SiR 1.0 but enables more efficient circuit targeting and reduced immunological impairment of targeted circuits which is advantageous both for research and for potential therapuetic use. I show that the SiR 2.0 viruses can be functionally produced from plasmids in HEK 293T cells within an experimentally tractable time frame (5-6 weeks). These viral preparations are produced to high functional titers that can successfully infect neurons following injection *in vivo*. Thus, we confirm that the TEVs-PEST modification does not inadvertantly impair the production, infectivity or trans-synaptic transfer ability of the CVS-N2C strain, and that the SiR 2.0 can be used functionally *in vivo*.

6.1.1 The SiR virus does not accumulate revertant mutations to any appreciable level *in vitro* or *in vivo*

Given the common knowledge that ssRNA viruses are subject to a high rate of mutations and a recent report concerning the two viral batches of SiR virus that had accumulated WT reverting mutations[192], we conducted a thorough analysis of the genomic stability of multiple SiR virus preparations. We show through sequencing analyses of the viral genomes collected from both non-injected SiR virus preparations and from SiR infected neurons *in vivo*, that the TEVs-PEST modification that is crucial for the self-inactivation and thus reduced toxicity of the SiR virus is not selectively mutated either during production in cell culture *in vitro* or following infection of neurons *in vivo*.

Next generation sequencing of viral RNA from infected hippocampi 1 week p.i. does not identify any mutations within the TEVs-PEST modification, despite identifying multiple mutations elsewhere in the genome. In addition, trans-synaptically labelled premotor neurons also demonstrate positive staining for the TEVs epitope that can only be expressed by non-revertant bona fide SiR viral particles. As such, this data suggests that SiR viruses,

if produced as previously described[191], show a very low risk of accumulating revertant mutations during production *in vitro* and following *in vivo* injection.

6.1.2 The SiR 2.0 virus as a potential first step towards a therapeutic rabies vector

In this chapter I demonstrate that the SiR 2.0 has many characteristics different to the SiR 1.0 virus that increase its suitability as a potential therapeutic vector for circuit-based therapy.

6.1.2.1 Increased genomic stability of the CVS-N2C strain

First I show data regarding the increased genomic stability of the CVS-N2C strain compared to the SAD-B19 strain. Although no stringently controlled comparison of mutation rates have been conducted, I show, through the comparison of multiple parallel productions of WT or SiR SAD-B19 vs. WT or SiR CVS-N2C, that the SAD-B19 productions unanimously demonstrate a higher accumulation of mutation events. This is despite the fact that viral production conditions are in fact biased for the increased mutation accumulation of the CVS-N2C strain as these viruses take longer to produce *in vitro* enabling a longer time for mutation accumulation. Nonetheless, our results show that even under these favourable conditions, the SAD-B19 demonstrates higher mutation accumulation.

These comparisons are informative as they analyse the genomic stability of SAD-B19 or CVS-N2C viral preparations that have been produced as described in the methods section for quick high-titer productions. Therefore, they accurately reflect the genomic heterogeneity of viral populations that would be used for *in vivo* injections of either of the two strains. Undoubtedly, the ability to produce more homogenous viral productions is an advantageous quality for the large-scale production of therapuetic viral vectors for therapy as they must be quality controlled to a very high level.

6.1.2.2 Simple SiR design modifications that could be implemented to abolish revertant mutation risk for safe therapuetic use

Another important point regarding SiR genomic stability and potential therapeutic use is that, for patient safety-considerations, the risk of revertant SiR viruses in therapy should effectively be zero. Although our sequencing data shows that the risk of revertant mutations accumulating in SiR viruses is extremely low if produced in the correct conditions, it may

also be worth to consider other simple measures that could be taken to reduce this potential risk even further.

For example, one could consider adding the TEVs-PEST addition to the N terminus of the N protein. In this way, the introduction of frameshift or nonsense mutations resulting in premature truncation of the protein would result in a lack of N expression altogether. Such mutations would therefore inactivate the virus without posing a risk for patients. In contrast, frameshift or nonsense mutations of C terminal protein modifications result in potential toxic revertant mutations as is the case for our current C terminal TEVs-PEST additions.

In addition, further study into the mechanism of the viral RNA polymerase complex could reveal particular genomic signatures that result in higher or lower mutation accumulation. Therefore, with such an understanding, the sequence of the TEVs-PEST addition could be modified to reduce this risk as much as possible.

6.1.2.3 SiR 2.0 enables more efficient targeting of the motor-premotor networks implicated in ALS

A direct comparison into the neurotropism of the SiR 2.0 compared to the SiR 1.0 vector shows that the SiR 2.0 demonstrates a significantly higher efficiency of motor neuron terminal targeting following intramuscular injection compared to the SiR 1.0 vector. This finding confirms previous studies describing the higher neurotropism of the CVS-N2C strain compared to the SAD-B19 strain.

However, our experimental design enables, for the first time, the direct comparison of the neurotropism of the recently optimised oG coated SAD-B19 virus with the neurotropism of the CVS-N2C virus within the same circuit in the same experimental conditions. Previous reports specify that the CVS-N2C strain has roughly 10-20 fold higher trans-synaptic tracing efficiency compared to the SAD-B19 in corticostriatal and spinal motor circuits[190]. Comparison of the original WT ΔG SAD-B19 strain with a new optimized glycoprotein (oG) coated SAD-B19 strain has also been shown to improve monosynaptic tracing efficiency by around 10- to 20- in thalamocortical circuits[244]. Therefore this would suggest that the CVS-N2C and the oG coated SAD-B19 wuold show roughly similar tracing efficiency. However, we show that in the case of motor neuron terminal targeting following intramuscular injection, the CVS-N2C still demonstrates an almost 30 fold higher neurotropism despite oG modifications to the SAD-B19 strain. We also show that the trans-synaptic transfer kinetics and efficiency of the CVS-N2C strain is significantly greater then the oG coated SAD-B19 strain.

The significance of this finding is relevant for therapy due to the need for efficient circuit targeting *in vivo*. As described in the introduction, efficient circuit-targeting is key for neuro-

logical therapy as it reduces the risk of immunological responses to therapeutic vectors and off-target effects. Equally importantly, efficient circuit targeting enables higher therapeutic doses into the networks actually implicated in the neural pathology. We demonstrate this targeting efficiency in a therapeutically relevant motor-premotor network that is affected in the onset of ALS. ALS pathology is known to initiate in the spinal cord α motorneurons and Betz cells of the motor cortex[121]. Therefore, the testing of the SiR 2.0 vector for high tropism to motor neurons and efficient retrograde trans-synaptic transfer from spinal motor neurons demonstrates the advantages of the SiR 2.0 vector directly within ALS-implicated circuits.

6.1.2.4 SiR 2.0 is significantly less immunogenic compared to the SiR 1.0

In this chapter, I also demonstrate through immunohistochemistry and through RNA-Seq experiments in two different circuits *in vivo* that the immune response to rabies virus infection is markedly reduced in the new SiR 2.0 vector when compared to its WT or SAD-B19 counterpart. This is a relevant result with view to therapy as viral immunogenicity is a key factor determining the success of viral therapeutic vectors. For example, the effective use of systemic AAV delivery for CNS targeting is hindered significantly by the fact that the high systemic doses of recombinant AAV required for efficient CNS targeting triggers host immune responses. This is despite the fact that recombinant AAVs and lentiviruses are already the preferred therapeutic vectors of choice due to their low immunogenicities compared to other viruses.

6.1.2.5 Other potential future modifications that could dampen rabies virus immunogenicity

As our immunofluorescence and RNA-seq data implicated the activation of the PRR RIG-I as a sensor for rabies infection, we considered that rabies virus immunogenicity could be reduced even further by the expression of RIG-I inhibitors by the rabies genome. This could be performed through multiple different mechanisms. For example, rabies mediated expression of RIG-I targeting siRNA could reduce RIG-I induction in rabies infected cells. In a similar vein, the ablation of RIG-I using antisense morpholino oligonucleotide in mouse neuro2A cells is known to negatively regulate RIG-I activation following infection by Japenese Encephalitis Virus (JEV)[245]. JEV is also a RNA virus which is recognised by the PRR RIG-I.

In addition, one could consider expressing RIG-I inhibiting proteins from the rabies genome. For example, the viral nonstructural protein 1 (NS1) of the influenza virus negatively regulates RIG-I activation at many levels[336]: it interacts directly with RIG-I, binds dsRNA to reduce RIG-I activation and prevents the induction of downstream IFN responses by inhibiting the activation of transcription factors such as IRF3.

6.1.3 Considerations for the use of SiR 2.0 in research

Here we show that the SiR 2.0 demonstrates higher neurotropism, higher trans-synaptic transfer efficiency and decreased immunogenicity compared to the SiR 1.0. These changes are undoubtedly unanimously advantageous for the development of rabies vector derived therapies.

However, for the use of SiRs for research, choice of vector depends on a wider variety of factors. For example, the production of SiR 2.0 is more difficult then the production of SiR 1.0 mainly due to the inability to visually track viral production *in vitro*. Therefore, in cases where reduced immune perturbation of targeted circuits is not critical, one may prefer to use the SiR 1.0 vector. In addition, although we demonstrate that the SiR 2.0 is more neurotropic when injected into peripheral organs (and this is clearly advantageous for non-invasive therapeutic use), we do not find that the SiR 2.0 is significantly more neurotropic when injected intracerebrally (e.g. into mouse CA1). Therefore, although the SiR 2.0 may be preferred for more efficient and complete trans-synaptic targeting of peripherally infected circuits, it would not have the same advantage when targeting circuits by intracerebral injection. In addition, in cases where a shorter length or higher level of exogenous gene expression is important, one may prefer the SAD-B19 derived SiR 1.0 which shows more transient but higher levels of transgene expression following infection.

6.2 The use of SiR 2.0 for CRISPR/Cas9 mediated genome-editing

In the second experimental chapter I present data showing the functional incorporation of both a Cas9 endonuclease and a gRNA sequence into the WT ΔG rabies virus genome. Through screening of multiple designs, we show that such a virus can be produced and used to target endogenous loci *in vitro*. We also produce SiR rabies viruses incorporating the same design cassette, and show through a preliminary experiment, that this virus can perform SiR mediated CRISPR/Cas9 genome editing *in vivo*. Given the repeated observation

that neurodegenerative disorders are characterized by the network-level progression of pathological protein aggregates, and given that familial forms of these diseases are caused by specific gene mutations, circuit-level genome editing vectors in theory would be a promising therapeutic strategy for these diseases.

6.2.1 SiR-Cas9-gRNA could potentially target many different disease circuits by non-invasive peripheral injection

As the rabies virus can infect a multitude of different peripheral sensory and motor neurons, trans-synaptic transfer of rabies virus from these different primary neurons opens up the possibility of using the rabies virus to target many different disease circuits. The key example I focus on throughout this thesis is the targeting of lower and upper motor neurons by single intramuscular injection of rabies virus, and the clear advantages this has in targeting all of the two areas initially involved in the pathology of ALS[121].

However, this principle could in theory be applied to many other diseases. For example, Parkinson's disease is known to present in patients initially by involvement of the olfactory circuitry. In fact, loss of a sense of smell is very commonly the very first symptom that Parkinson's patients experience[121]. Therefore, it is tempting to consider the intranasal administration of a trans-synaptic rabies virus to the olfactory epithelium to enable the targeting of olfactory circuits for therapy.

In addition, the lateral hypothalamus and arcuate nucleus are highly involved in the neural control of appetite[337,338]. As hypothalamic pathways are known to project polysynaptically into peripheral adipose tissue, one could consider targeting these hypothalamic neural networks by peripheral rabies injection into peripheral adipose tissue[339,340]. Here, the injection of rabies virus into the adipose tissue or the stellate ganglion would enable retrograde targeting of crucial feeding centers of the brain and thus be used to modulate feeding behaviour for the treatment of eating disorders.

Given our preliminary finding that the SiR can conduct genome editing *in vivo*, we propose that the continued improvement of this vector could one day enable the execution of these proposed examples of network-targeted therapy.

6.2.2 Improving the production and quantification methods of SiR-Cas9-gRNA viruses that do not express fluorescent markers

Although we show preliminary *in vivo* data showing SiR mediated targeting of endogenous genomic loci, we need to design an experimental protocol that would enable robust quantification of genome-editing efficiency to successfully develop a genome-editing SiR. However, the current design presents some time-limiting challenges for such developments. First we find that the production of unlabelled viruses is much more difficult and labour intensive then for rabies viruses expressing fluorescent markers, as live visualisation of viral replication and amplification is not possible. Second, following injection *in vivo*, there is no method of permanently labeling once SiR infected cells. Therefore, it is hard to quantify exactly what percentage of SiR infected cells show successful SiR-mediated genome editing. As such, we have already described a preliminary design of a split-FLP tag in order to permanently label SiR infected cells. We know from our studies of split-CRE reconstitution that the design of efficient split-recombinase enzymes can be difficult. However, given the fact that virally driven split-FLP expression will be very high and that co-transfection of split-CRE fragments to high levels does result in clearly visible recombinase activity, we suspect that high efficiency of reconstitution will not be a key limiting factor in this development.

6.2.3 Other potential modifications to boost gRNA expression from the rabies genome

Further study into the structural elements of single gRNA has revealed that some parts of the gRNA are not necessary for efficient Cas9 targeting[295]. In addition, there is also the potential for rational modifications to the gRNA structure to improve gRNA stability or Cas9 targeting. Such changes may help to increase the efficiency of SiR mediated CRISPR/Cas9 genome editing. For example, the sequence flexibility of the lower stem, upper stem and hairpin 1 suggest that gRNAs can be designed that are more compact and more functionally and structurally stable[295].

6.3 Identification and mechanistic elucidation of neuronal proteins demonstrating interneuronal transfer through neural circuits

In the final experimental chapter I present exploratory work aiming to identify efficient interneuronal protein transfer systems within well-established neural circuits *in vivo*. I also aim to implement existing cell-specific proteomic labeling techniques into such circuits. Here our most notable findings are the observation that intraocularly injected proteins (human A53T mutant alpha synuclein) as well as BioID2-tagged RGC donor proteins demonstrate interneuronal protein transfer from RGCs into the visual cortex in a time scale of days. Given the relative ease of RGC targeting by intraocular injection, and the fast time-scale of protein movement observed in this system, it should be pursued for future mechanistic elucidation of intercellular protein movement mechanisms. We also note that although considerably slower, the movement of V5 Tau through the hippocampal circuitry also lends itself nicely to further mechanistic elucidation. Lastly, although not demonstrated directly in this thesis, the use of candidate based or proteomic labeling techniques can also be implemented in neural circuitry *in vitro* following the differentiation of neuro2A cells.

6.3.1 Mechanistic investigation into the observed transfer of V5-Tau through hippocampal circuitry

So far, we have successfully designed and produced a lentivirus which, following injection into the mouse hippocampus, establishes a V5-Tau expressing donor neuron population. Analyses of brains injected with lentivirus-CAMKII-CREGFP-P2A-V5Tau in mouse CA1 showed that our viral design was suitable for generating donor neuron populations within prescribed restricted regions in the mouse brain. We have also observed what appears to be specific V5 staining both in the subiculum (the first order projection site of CA1 neurons) and the entorhinal cortex (the projection site of second order neurons from the subiculum). This strongly suggests that V5-Tau was transferred interneuronally through the canonical hippocampal circuitry. However, the addition of a cytoplasmic or membrane marker to donor neurons may aid further assertion that the V5 staining is present in bona fide acceptor neurons which are cellularly distinct to lentivirally infected donor neurons. This could be done by injecting the same construct into a transgenic lox-stop-lox membrane GFP mouse, such that

CRE-GFP expression in donor neurons will not only label donor neurons with a nuclear GFP signal, but also by GFP staining of its plasma membrane.

In addition, to expand this approach, we could replace tau in this design to test other candidate proteins in order to compare their relative mobilities within the same mouse hippocampal network. These other candidate proteins may be proteins which are unrelated to neurodegeneration but have been shown to move transneuronally in other systems (eg. neuronal transcription factors such as Otx2, or retroviral gag like proteins such as Arc). This would offer insight into the specificity or generalisability of protein transfer within particular circuitry. For example, it may be the case that although Otx2 demonstrates interneuronal movement through the visual circuitry, it may not do so within the hippocampal circuitry.

The expression of CRE in donor cells also enables the versatile subsequent manipulation of donor cells to elucidate various aspects behind the mechanism of protein transfer. For example, CRE dependent optogenetic manipulation of neuronal activity or knockout of genes in specific molecular pathways would offer new insight into the mechanisms underlying interneuronal protein transfer. Alternatively, through the CRE dependent expression of DREADDs in donor cells, one could investigate whether protein transfer increases with increased synaptic activity.

Lastly, by inducing candidate protein expression in different distinct classes of neurons by modifying the promoter of our donor construct, we can also assess the contribution of different molecularly defined neuronal classes to protein propagation, which remains, so far, unclear.

6.3.2 The use of neuro2A cell mixing experiments to study interneuronal protein transfer *in vitro*

Differentiated neuro2A cells are known to form randomly connected neuronal populations in culture[327,328]. Following the establishment of such *in vitro* neural circuitry we could subsequently use sparse lentiviral labelling to create a donor population. Additional rabies tracing or optogenetic methods as described for the *in vivo* experiments above would be equally feasible in such an *in vitro* system. In addition, the cell mixing experiments between WT and mutRS expressing neuro2A cells in Fig. 5.8 suggest that there may be some intercellular transfer of azide labelled proteins from donor cells into acceptor cells *in vitro*.

However, a key consideration that should be taken when studying protein transfer *in vitro* is the possibility that there may be non-specific uptake of donor proteins released from dead donor cells into acceptor cells. To address such issues, one could consider

the use of microfluidic chambers to separate the culture medium of donor and acceptor cells. Alternatively, one could use culture medium from donor neurons and add them onto separately cultured acceptor neurons for control experiments. The establishment of a reliable and representative *in vitro* neural network model could potentially be a very powerful platform for future studies. For example, the system could be used to perform large scale screening tests to investigate whether the systematic knock out of various genes inhibits or enhances the intercellular transfer of proteins. This could reveal various pathways or genes that are implicated in the intercellular protein transfer mechanism. In a medical context, this sort of system may be useful to screen drug targets that could help prevent the propagation of pathological proteins through neural networks in neurodegenerative disease.

6.3.3 Identification of trans-synaptically transported endogenous RGC proteins in the visual cortex by mass spectrometry

Lastly, we demonstrate that AAV driven expression of BioID2 in RGCs followed by biotin administration (intraocularly and intraperitoneally) results in efficient biotinylation of the RGC proteome. We also observe that this protocol results in specific biotin-staining in the cell bodies of the LGN and the contralateral visual cortex. This strongly suggests that biotinylated RGC proteins are being transferred transsynaptically into LGN neurons projecting to the visual cortex. Therefore protein purification from visual cortex samples followed by enrichment of biotinylated proteins and mass spectrometry identification would enable us, in an unbiased manner, to identify interneuronally transported endogenous RGC proteins.

Very recently, biotinylation of the RGC proteome was in fact used to subsequently identify RGC proteins transported to various first order projection sites of RGCs including the SC and LGN[337]. In this report, RGC proteome biotinylation was achieved by intraocular injection of N-hydroxysuccinimidobiotin (NHS-biotin) which binds covalently to lysines and N-terminal amino acids of proteins. Interestingly, they identified some proteins within the RGC transportome that overlap with post-synaptic density enriched protein datasets such as SYN1, STX1B, STX1BP, NSF, and SNAP25. These proteins support vesicle fusion in pre- and post-synaptic compartments, potentially suggesting a role for these axonally transported RGC proteins in interneuronal protein transfer.

One potential advantage of using AAV-transduced RGC systems over direct intraocular injection of proteins or protein tags is the ability for long-term biotinylation enabling larger biotinylated protein volumes transported from the RGC over longer time scales. In our case,

where we want to identify biotinylated proteins from a second order projection site, this may be largely helpful. In addition, establishment of RGCs by AAV transduction enables both the quantification and selective modification of RGC donor neuron properties. This also enables further mechanistic investigation into the molecular processes involved in intercellular protein transfer.

6.4 Final Summary

Neural circuits underlie human brain function. In this thesis I have aimed to develop novel tools to target and manipulate neural circuits *in vivo*.

First, I have developed and characterised a second generation self-inactivating rabies virus vector (SiR 2.0) which expands the trans-synaptic tracing vector toolbox for circuit neuroscience research and implements key improvements over the SiR 1.0 vector for its further development as a therapeutic vector. SiR 2.0 enables more efficient targeting of peripheral neural circuits, enhanced trans-synaptic labelling efficiency and markedly reduced immunogenicity when compared to its previous WT and SiR 1.0 counterparts.

Second, I have incorporated into the SiR 2.0 genome the two components of the CAS9/gRNA system and shown a proof of principle that the rabies virus genome can functionally encode these two components without compromising viral production or infectivity *in vivo*. I also discuss the potential experimental and therapeutic applications of a circuit-specific genome-editing vector for both research and therapy. In addition, I discuss the technical challenges that need to be overcome for the future development, improvement and *in vivo* characterisation of these SiR-Cas9-gRNA viruses. In theory, we propose that the rabies virus could be modified and optimised further to express multiple gRNAs to target multiple genomic loci.

Third, in my final experimental chapter, I identify different possible experimental systems that would enable further investigation into the mechanisms behind interneuronal protein transfer. We work to design and implement tools into these different experimental systems that would enable either a candidate based bottom up mechanistic approach or an unbiased top-bottom proteomic approach. In this work I have identified the mouse visual circuitry as a particularly accessible and efficient *in vivo* system for further investigation. I have demonstrated its uses both for candidate-protein based approaches as well as for further study into the generalisability of interneuronal protein transfer by RGC proteome biotinylation using the BioID2 system. It may be possible that the identification of novel proteins that are transferred interneuronally as well as a greater understanding of the underlying biology

of this process will enable the development of better tools for anterograde trans-synaptic labelling and targeting of neural circuits in the future.

I envisage that the further development and optimisation of such tools will help to unravel some of the important unanswered questions in circuit neuroscience and pave the way for efficient therapeutic targeting of neural networks in disease.

References

1. Hughes, J. T. The Edwin Smith Surgical Papyrus: An analysis of the first case reports of spinal cord injuries. *Paraplegia* **26,** 71–82 (1988).
2. Celesia, G. G. Alcmaeon of Croton's Observations on Health, Brain, Mind, and Soul. *Journal of the History of the Neurosciences* **21,** 409–426. ISSN: 0964-704X (2012).
3. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* **10,** 186–198. ISSN: 1471-003X (2009).
4. Warren, J. D., Rohrer, J. D., Schott, J. M., Fox, N. C., Hardy, J. & Rossor, M. N. Molecular nexopathies: a new paradigm of neurodegenerative disease. *Trends in Neurosciences* **36,** 561–569. ISSN: 01662236 (2013).
5. Sepehri Rad, M., Choi, Y., Cohen, L. B., Baker, B. J., Zhong, S., Storace, D. A. & Braubach, O. R. Voltage and Calcium Imaging of Brain Activity. *Biophysical Journal* **113,** 2160–2167. ISSN: 0006-3495 (2017).
6. Vlasov, K., Van Dort, C. J. & Solt, K. *Methods in enzymology* 181–196 (2018). doi:10.1016/bs.mie.2018.01.022.
7. Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E. & Church, G. M. RNA-guided human genome engineering via Cas9. *Science (New York, N.Y.)* **339,** 823–6. ISSN: 1095-9203 (2013).
8. Cajal, S. y. Estructura de los centros nerviosos de las aves (1888).
9. Sherrington, C. S. Observations on the scratch-reflex in the spinal dog. *The Journal of Physiology* **34,** 1–50. ISSN: 00223751 (1906).
10. Golgi, C. Sulla sostanza grigia del cervello. *Gazetta Medica Italiana* **33,** 244–246 (1873).
11. Adrian, E. D. The impulses produced by sensory nerve endings: Part I. *The Journal of physiology* **61,** 49–72. ISSN: 0022-3751 (1926).
12. Adrian, E. D. A. *The basis of sensation, the action of the sense organs* (Christophers, 1928).

13. HUBEL, D. H. Tungsten Microelectrode for Recording from Single Units. *Science* **125,** 549–550. ISSN: 0036-8075 (1957).

14. Hubel, D. H. & Wiesel, T. N. Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology* **148,** 574–91. ISSN: 0022-3751 (1959).

15. Gross, C. G. Genealogy of the "Grandmother Cell". *The Neuroscientist* **8,** 512–518. ISSN: 1073-8584 (2002).

16. Weibell, C. J. *Principles of learning: 7 principles to guide personalized, student-centered learning in the technology-enhanced, blended learning environment.* PhD thesis (2011), Chapter 3.

17. Bain, A. *Mind and body : the theories of their relation* ISBN: 1402199740 (Henry S. King & Co., 1873).

18. McCulloch, W. S. & Pitts, W. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics* **5,** 115–133. ISSN: 00074985 (1943).

19. Rosenblatt, F. THE PERCEPTRON: A PROBABILISTIC MODEL FOR INFORMATION STORAGE AND ORGANIZATION IN THE BRAIN 1. **65,** 19–27 (1958).

20. Hebb, D. O.D. O. *The organization of behavior; a neuropsychological theory* 335. ISBN: 0471367273 (John Wiley and Sons, 1949).

21. Minsky, M. & Papert, S. *Perceptrons; an introduction to computational geometry* 258. ISBN: 9780262130431 (MIT Press, 1969).

22. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323,** 533–536. ISSN: 0028-0836 (1986).

23. Johnson, G. Artificial Brain Again Seen As a Guide To the Mind. *New York Times* (1988).

24. Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America* **79,** 2554–8. ISSN: 0027-8424 (1982).

25. Ichinose, T., Aso, Y., Yamagata, N., Abe, A., Rubin, G. M. & Tanimoto, H. Reward signal in a recurrent circuit drives appetitive long-term memory formation. *eLife* **4,** e10719. ISSN: 2050-084X (2015).

26. Eichler, K. *et al.* The complete connectome of a learning and memory centre in an insect brain. *Nature* **548,** 175–182. ISSN: 1476-4687 (2017).

27. Bedia, M. G., Corchado, J. M. & Castillo, L. F. *Bio-inspired Memory Generation by Recurrent Neural Networks* tech. rep. (2007), 55–62. doi:10.1007/978-3-540-73007-1_8.

28. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521,** 436–444. ISSN: 0028-0836 (2015).

29. Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. & Bengio, Y. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. arXiv: 1406.1078 (2014).

30. Bahdanau, D., Cho, K. & Bengio, Y. *Neural Machine Translation by Jointly Learning to Align and Translate* tech. rep. (2014). arXiv: 1409.0473.

31. Sutskever, I., Vinyals, O. & Le, Q. V. *Sequence to Sequence Learning with Neural Networks* 2014.

32. Krizhevsky, A., Sutskever, I. & E. Hinton, G. ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems* **25.** doi:10.1145/3065386 (2012).

33. He, K., Zhang, X., Ren, S. & Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. arXiv: 1502.01852 (2015).

34. Erhan, D., Bengio, Y., Courville, A. & Vincent, P. *Visualizing Higher-Layer Features of a Deep Network* tech. rep. (Département d'Informatique et Recherche Opérationnelle, 2009).

35. Zeiler, M. D. & Fergus, R. *LNCS 8689 - Visualizing and Understanding Convolutional Networks* tech. rep. (2014).

36. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A. & Torralba, A. *Object Detectors Emerge in Deep Scene CNNs* tech. rep. (2014). arXiv: 1412.6856.

37. Yamins, D. L., Hong, H., Cadieu, C. & DiCarlo, J. J. *Hierarchical Modular Optimization of Convolutional Networks Achieves Representations Similar to Macaque IT and Human Ventral Stream* 2013.

38. Cichy, R. M., Khosla, A., Pantazis, D., Torralba, A. & Oliva, A. Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. *Scientific Reports* **6,** 27755. ISSN: 2045-2322 (2016).

39. Cadieu, C. F., Hong, H., Yamins, D. L. K., Pinto, N., Ardila, D., Solomon, E. A., Majaj, N. J. & DiCarlo, J. J. Deep Neural Networks Rival the Representation of Primate IT Cortex for Core Visual Object Recognition. *PLoS Computational Biology* **10** (ed Bethge, M.) e1003963. ISSN: 1553-7358 (2014).

40. Sejnowski, T. J. & Rosenberg, C. R. *Parallel Networks that Learn to Pronounce English Text* tech. rep. (1987), 145–168.

41. Schüz, A. & Palm, G. Density of neurons and synapses in the cerebral cortex of the mouse. *Journal of Comparative Neurology* **286,** 442–455. ISSN: 00219967 (1989).

42. Braitenberg, V. & Schüz, A. *Anatomy of the Cortex* ISBN: 978-3-540-53233-0. doi:10.1007/978-3-662-02728-8 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1991).

43. *Cerebral Cortex* (eds Jones, E. G. & Peters, A.) ISBN: 978-1-4613-6706-2. doi:10.1007/978-1-4615-3824-0 (Springer US, Boston, MA, 1990).

44. Llinás, R. R. & Sugimori, M. *The Cerebellum Revisited* 167–181 (Springer US, New York, NY, 1992). doi:10.1007/978-1-4612-2840-0_8.

45. Holmgren, C., Harkany, T., Svennenfors, B. & Zilberter, Y. Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *The Journal of Physiology* **551,** 139–153. ISSN: 0022-3751 (2003).

46. Feldmeyer, D., Lübke, J. & Sakmann, B. Efficacy and connectivity of intracolumnar pairs of layer 2/3 pyramidal cells in the barrel cortex of juvenile rats. *The Journal of Physiology* **575,** 583–602. ISSN: 00223751 (2006).

47. Avermann, M., Tomm, C., Mateo, C., Gerstner, W. & Petersen, C. C. H. Microcircuits of excitatory and inhibitory neurons in layer 2/3 of mouse barrel cortex. *Journal of Neurophysiology* **107,** 3116–3134. ISSN: 0022-3077 (2012).

48. Fino, E., Packer, A. M. & Yuste, R. The Logic of Inhibitory Connectivity in the Neocortex. *The Neuroscientist* **19,** 228–237. ISSN: 1073-8584 (2013).

49. Karnani, M. M., Agetsuma, M. & Yuste, R. A blanket of inhibition: functional inferences from dense inhibitory connectivity. *Current Opinion in Neurobiology* **26,** 96–102. ISSN: 0959-4388 (2014).

50. Yuste, R. From the neuron doctrine to neural networks. *Nature Reviews Neuroscience* **16,** 487 (2015).

51. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393,** 440–442. ISSN: 0028-0836 (1998).

52. Milgram, S. The small world problem. *Psychology today* **2,** 60–67 (1967).

53. Bullmore, E. & Sporns, O. The economy of brain network organization. *Nature Reviews Neuroscience* **13,** 336–349. ISSN: 1471-003X (2012).

54. Wernicke, C. The aphasic symptom complex: a psychological study on a neurological basis. Breslau: Cohn and Weigert. Reprinted. *Boston studies in the philosophy of science* **4** (1874).

55. Pribram, K. H. & Mishkin, M. Simultaneous and successive visual discrimination by monkeys with inferotemporal lesions. *Journal of comparative and physiological psychology* **48,** 198–202. ISSN: 0021-9940 (1955).

56. Desimone, R., Albrightj, T. D., Gross, C. G. & Bruce, C. *STIMULUS-SELECTIVE PROPERTIES OF INFERIOR TEMPORAL NEURONS IN THE MACAQUE1* tech. rep. 8 (1984), 2051–2062.

57. Gross, C. G., Rocha-Miranda, C. E. & Bender, D. B. *Visual properties of neurons in inferotemporal cortex of the Macaque.* tech. rep. 1 (1972), 96–111. doi:10.1152/jn.1972.35.1.96.

58. Perrett, D., Rolls, E. & Caan, W. *Visual neurones responsive to faces in the monkey temporal cortex* tech. rep. 3 (1982), 329–342. doi:10.1007/BF00239352.

59. Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C. & Fried, I. *Invariant visual representation by single neurons in the human brain* tech. rep. 7045 (2005), 1102–1107. doi:10.1038/nature03687.

60. James, W. *The Principles of Psychology* (Dover Publications, 1890).

61. Konorski, J. Integrative activity of the brain; an interdisciplinary approach (1967).

62. Barlow, H. B. *Single units and sensation: A neuron doctrine for perceptual psychology?* tech. rep. (1972), 371–394.

63. Chang, L. & Tsao, D. Y. The Code for Facial Identity in the Primate Brain. *Cell* **169,** 1013–1028.e14. ISSN: 1097-4172 (2017).

64. Quiroga, R. Q., Kreiman, G., Koch, C. & Fried, I. Sparse but not 'Grandmother-cell' coding in the medial temporal lobe. *Trends in Cognitive Sciences* **12,** 87–91. ISSN: 1364-6613 (2008).

65. Munevar, G. *Proceedings of the XXII World Congress of Philosophy* 25–31 (Philosophy Documentation Center, 2008). doi:10.5840/wcp22200834552.

66. Sherrington, C. S. *Man on his Nature* ISBN: 9780511694196. doi:10.1017/CBO9780511694196 (Cambridge University Press, Cambridge, 1940).

67. Saxena, S. & Cunningham, J. P. Towards the neural population doctrine. *Current Opinion in Neurobiology* **55,** 103–111. ISSN: 0959-4388 (2019).

68. Briggman, K. L., Abarbanel, H. D. I. & Kristan, W. B. Optical imaging of neuronal populations during decision-making. *Science (New York, N.Y.)* **307,** 896–901. ISSN: 1095-9203 (2005).

69. Harvey, C. D., Coen, P. & Tank, D. W. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* **484,** 62–68. ISSN: 0028-0836 (2012).

70. Yu, B. M., Cunningham, J. P., Santhanam, G., Ryu, S. I., Shenoy, K. V. & Sahani, M. Gaussian-Process Factor Analysis for Low-Dimensional Single-Trial Analysis of Neural Population Activity. *Journal of Neurophysiology* **102,** 614–635. ISSN: 0022-3077 (2009).

71. Pandarinath, C. *et al.* Inferring single-trial neural population dynamics using sequential auto-encoders. *Nature Methods* **15,** 805–815. ISSN: 1548-7091 (2018).

72. Morcos, A. S. & Harvey, C. D. History-dependent variability in population dynamics during evidence accumulation in cortex. *Nature Neuroscience* **19,** 1672–1681. ISSN: 1097-6256 (2016).

73. Afshar, A., Santhanam, G., Yu, B., Ryu, S., Sahani, M. & Shenoy, K. Single-Trial Neural Correlates of Arm Movement Preparation. *Neuron* **71,** 555–564. ISSN: 0896-6273 (2011).

74. Krasne, F. B. & Lee, S. C. Response-dedicated trigger neurons as control points for behavioral actions: selective inhibition of lateral giant command neurons during feeding in crayfish. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **8,** 3703–12. ISSN: 0270-6474 (1988).

75. Esch, T. & Kristan, W. B. Decision-Making in the Leech Nervous System. *Integrative and Comparative Biology* **42,** 716–724. ISSN: 1540-7063 (2002).

76. Getting, P. A. & Dekin, M. S. *Model Neural Networks and Behavior* 3–20 (Springer US, Boston, MA, 1984). doi:10.1007/978-1-4757-5858-0_1.

77. Getting, P. A. Emerging Principles Governing the Operation of Neural Networks. *Annual Review of Neuroscience* **12,** 185–204. ISSN: 0147-006X (1989).

78. Popescu, I. R. & Frost, W. N. Highly dissimilar behaviors mediated by a multifunctional network in the marine mollusk Tritonia diomedea. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **22,** 1985–93. ISSN: 1529-2401 (2002).

79. Lieske, S. P., Thoby-Brisson, M., Telgkamp, P. & Ramirez, J. M. Reconfiguration of the neural network controlling multiple breathing patterns: eupnea, sighs and gasps. *Nature Neuroscience* **3,** 600–607. ISSN: 1097-6256 (2000).

80. Travers, J. B., DiNardo, L. A. & Karimnamazi, H. Medullary reticular formation activity during ingestion and rejection in the awake rat. *Experimental brain research* **130,** 78–92. ISSN: 0014-4819 (2000).

81. Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M. & Harris, K. D. Spontaneous behaviors drive multidimensional, brainwide activity. *Science (New York, N.Y.)* **364,** 255. ISSN: 1095-9203 (2019).

82. Golde, T. E., Borchelt, D. R., Giasson, B. I. & Lewis, J. Thinking laterally about neurodegenerative proteinopathies. *The Journal of Clinical Investigation* **123,** 1847–1855. ISSN: 0021-9738 (2013).

83. Rubinov, M. & Bullmore, E. Fledgling pathoconnectomics of psychiatric disorders. *Trends in Cognitive Sciences* **17,** 641–647. ISSN: 1364-6613 (2013).

84. John Martyn Harlow. Recovery from the passage of an iron bar through the head. *Publications of the Massachusetts Medical Society* **2,** 327–347 (1868).

85. Anonymous. Report of Prof. Bigelow. *Stethoscope Va Med Gaz 1* **1,** 162–163 (1851).

86. Catani, M. & Ffytche, D. H. The rises and falls of disconnection syndromes. *Brain* **128,** 2224–2239. ISSN: 1460-2156 (2005).

87. WHO. *The world health report 2001 - Mental Health: New Understanding, New Hope* tech. rep. (2013).

88. Association, A. P. *et al. Diagnostic and statistical manual of mental disorders (DSM-5®)* (American Psychiatric Pub, 2013).

89. Uhlhaas, P. & Singer, W. Neuronal Dynamics and Neuropsychiatric Disorders: Toward a Translational Paradigm for Dysfunctional Large-Scale Networks. *Neuron* **75,** 963–980. ISSN: 08966273 (2012).

90. Bullmore, E. & Sporns, O. The economy of brain network organization. *Nature Reviews Neuroscience* **13,** 336–349. ISSN: 1471-003X (2012).

91. Kubicki, M. & Shenton, M. E. Diffusion Tensor Imaging findings and their implications in schizophrenia. *Current Opinion in Psychiatry* **27,** 179–184. ISSN: 0951-7367 (2014).

92. Rotarska-Jagiela, A., Schönmeyer, R., Oertel, V., Haenschel, C., Vogeley, K. & Linden, D. E. The corpus callosum in schizophrenia-volume and connectivity changes affect specific regions. *NeuroImage* **39,** 1522–1532. ISSN: 10538119 (2008).

93. Henze, R., Brunner, R., Thiemann, U., Parzer, P., Klein, J., Resch, F. & Stieltjes, B. White matter alterations in the corpus callosum of adolescents with first-admission schizophrenia. *Neuroscience Letters* **513,** 178–182. ISSN: 03043940 (2012).

94. Alexander-Bloch, A. F., Vértes, P. E., Stidd, R., Lalonde, F., Clasen, L., Rapoport, J., Giedd, J., Bullmore, E. T. & Gogtay, N. The Anatomical Distance of Functional Connections Predicts Brain Network Topology in Health and Schizophrenia. *Cerebral Cortex* **23,** 127–138. ISSN: 1460-2199 (2013).

95. Van den Heuvel, M. P., Mandl, R. C. W., Stam, C. J., Kahn, R. S. & Hulshoff Pol, H. E. Aberrant Frontal and Temporal Complex Network Structure in Schizophrenia: A Graph Theoretical Analysis. *Journal of Neuroscience* **30,** 15915–15926. ISSN: 0270-6474 (2010).

96. Liu, X., Lai, Y., Wang, X., Hao, C., Chen, L., Zhou, Z., Yu, X. & Hong, N. Reduced white matter integrity and cognitive deficit in never-medicated chronic schizophrenia: A diffusion tensor study using TBSS. *Behavioural Brain Research* **252,** 157–163. ISSN: 01664328 (2013).

97. Guo, W., Liu, F., Liu, Z., Gao, K., Xiao, C., Chen, H. & Zhao, J. Right lateralized white matter abnormalities in first-episode, drug-naive paranoid schizophrenia. *Neuroscience Letters* **531,** 5–9. ISSN: 03043940 (2012).

98. Hadjikhani, N., Joseph, R. M., Snyder, J. & Tager-Flusberg, H. Anatomical Differences in the Mirror Neuron System and Social Cognition Network in Autism. *Cerebral Cortex* **16,** 1276–1282. ISSN: 1460-2199 (2006).

99. Gotts, S. J., Simmons, W. K., Milbury, L. A., Wallace, G. L., Cox, R. W. & Martin, A. Fractionation of social brain circuits in autism spectrum disorders. *A JOURNAL OF NEUROLOGY* **135,** 2711–2725 (2012).

100. Liao, Y. *et al.* Is depression a disconnection syndrome? Meta- analysis of diffusion tensor imaging studies in patients with MDD. *Journal of Psychiatry & Neuroscience* **38,** 49–56. ISSN: 11804882 (2013).

101. Sacchet, M. D., Prasad, G., Foland-Ross, L. C., Joshi, S. H., Hamilton, J, Thompson, P. M. & Gotlib, I. H. Structural abnormality of the corticospinal tract in major depressive disorder. *Biology of Mood & Anxiety Disorders* **4,** 8. ISSN: 2045-5380 (2014).

102. Rutter, L., Carver, F. W., Holroyd, T., Nadar, S. R., Mitchell-Francis, J., Apud, J., Weinberger, D. R. & Coppola, R. Magnetoencephalographic gamma power reduction in patients with schizophrenia during resting condition. *Human brain mapping* **30,** 3254–64. ISSN: 1097-0193 (2009).

103. Uhlhaas, P. J. & Singer, W. Abnormal neural oscillations and synchrony in schizophrenia. *Nature Reviews Neuroscience* **11,** 100–113. ISSN: 1471-003X (2010).

104. Boutros, N. N., Arfken, C., Galderisi, S., Warrick, J., Pratt, G. & Iacono, W. The status of spectral EEG abnormality as a diagnostic test for schizophrenia. *Schizophrenia Research* **99,** 225–237. ISSN: 0920-9964 (2008).

105. Bernier, R, Dawson, G, Webb, S & Murias, M. EEG mu rhythm and imitation impairments in individuals with autism spectrum disorder. *Brain and cognition* **64,** 228–37. ISSN: 0278-2626 (2007).

106. Oberman, L. M., Hubbard, E. M., McCleery, J. P., Altschuler, E. L., Ramachandran, V. S. & Pineda, J. A. EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Cognitive Brain Research* **24,** 190–198. ISSN: 0926-6410 (2005).

107. Braak, H & Braak, E. Neuropathological stageing of Alzheimer-related changes. *Acta neuropathologica* **82,** 239–59. ISSN: 0001-6322 (1991).

108. Braak, H., Ghebremedhin, E., Rüb, U., Bratzke, H. & Del Tredici, K. Stages in the development of Parkinson's disease-related pathology. *Cell and Tissue Research* **318,** 121–134. ISSN: 0302-766X (2004).

109. Brettschneider, J. *et al.* Stages of pTDP-43 pathology in amyotrophic lateral sclerosis. *Annals of Neurology* **74,** 20–38. ISSN: 03645134 (2013).

110. Kassubek, J., Müller, H.-P., Del Tredici, K., Brettschneider, J., Pinkhardt, E. H., Lulé, D., Böhm, S., Braak, H. & Ludolph, A. C. Diffusion tensor imaging analysis of

sequential spreading of disease in amyotrophic lateral sclerosis confirms patterns of TDP-43 pathology. *Brain* **137** (2014).

111. Sapp, E, Penney, J, Young, A, Aronin, N, Vonsattel, J. P. & DiFiglia, M. Axonal transport of N-terminal huntingtin suggests early pathology of corticostriatal projections in Huntington disease. *Journal of neuropathology and experimental neurology* **58,** 165–73. ISSN: 0022-3069 (1999).

112. Gutekunst, C.-A., Li, S.-H., Yi, H., Mulroy, J. S., Kuemmerle, S., Jones, R., Rye, D., Ferrante, R. J., Hersch, S. M. & Li, X.-J. Nuclear and Neuropil Aggregates in Huntington's Disease: Relationship to Neuropathology. *Journal of Neuroscience* **19** (1999).

113. Rosas, H. D., Salat, D. H., Lee, S. Y., Zaleta, A. K., Pappu, V., Fischl, B., Greve, D., Hevelone, N. & Hersch, S. M. Cerebral cortex and the clinical expression of Huntington's disease: complexity and heterogeneity. *Brain : a journal of neurology* **131,** 1057–68. ISSN: 1460-2156 (2008).

114. G. Vonsattel, J. P. & DiFiglia, M. Huntington Disease. *Journal of Neuropathology and Experimental Neurology* **57,** 369–384. ISSN: 0022-3069 (1998).

115. Brettschneider, J. *et al.* Sequential distribution of pTDP-43 pathology in behavioral variant frontotemporal dementia (bvFTD). *Acta Neuropathologica* **127,** 423–439. ISSN: 0001-6322 (2014).

116. Bero, A. W., Yan, P., Roh, J. H., Cirrito, J. R., Stewart, F. R., Raichle, M. E., Lee, J.-M. & Holtzman, D. M. Neuronal activity regulates the regional vulnerability to amyloid-β deposition. *Nature Neuroscience* **14,** 750–756. ISSN: 1097-6256 (2011).

117. Pooler, A. M. *et al.* Physiological release of endogenous tau is stimulated by neuronal activity. *EMBO reports* **14,** 389–94. ISSN: 1469-3178 (2013).

118. Pecho-Vrieseling, E. *et al.* Transneuronal propagation of mutant huntingtin contributes to non–cell autonomous pathology in neurons. *Nature Neuroscience* **17,** 1064–1072. ISSN: 1097-6256 (2014).

119. de Calignon, A. *et al.* Propagation of Tau Pathology in a Model of Early Alzheimer's Disease. *Neuron* **73,** 685–697. ISSN: 08966273 (2012).

120. Palop, J. J., Chin, J. & Mucke, L. A network dysfunction perspective on neurodegenerative diseases. *Nature* **443,** 768–773. ISSN: 0028-0836 (2006).

121. British Medical Journal. *BMJ Best Practices* (BMJ Publ. Group, 2019).

122. Lomen-Hoerth, C, Murphy, J, Langmore, S, Kramer, J. H., Olney, R. K. & Miller, B. Are amyotrophic lateral sclerosis patients cognitively normal? *Neurology* **60,** 1094–7. ISSN: 1526-632X (2003).

123. Ringholz, G. M., Appel, S. H., Bradshaw, M., Cooke, N. A., Mosnik, D. M. & Schulz, P. E. Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology* **65,** 586–590. ISSN: 0028-3878 (2005).

124. Strong, M. J., Grace, G. M., Orange, J. B., Leeper, H. A., Menon, R. S. & Aere, C. A prospective study of cognitive impairment in ALS. *Neurology* **53,** 1665–70. ISSN: 0028-3878 (1999).

125. Bowen, B., Pattany, P., Bradley, W., Seres, P., Lynch, M., Sekhon, R., Johnston, W. & Kalra, S. MR Imaging and Localized Proton Spectroscopy of the Precentral Gyrus in Amyotrophic Lateral Sclerosis. *American Journal of Neuroradiology* **21,** 647–658. ISSN: 0195-6108 (2011).

126. Wilson, C. M., Grace, G. M., Munoz, D. G., He, B. P. & Strong, M. J. Cognitive impairment in sporadic ALS: a pathologic continuum underlying a multisystem disorder. *Neurology* **57,** 651–7. ISSN: 0028-3878 (2001).

127. Walker, M. P., Ayre, G. A., Cummings, J. L., Wesnes, K., McKeith, I. G., O'Brien, J. T. & Ballard, C. G. Quantifying fluctuation in dementia with Lewy bodies, Alzheimer's disease, and vascular dementia. *Neurology* **54,** 1616–1625. ISSN: 0028-3878 (2000).

128. Bradshaw, J. Fluctuating cognition in dementia with Lewy bodies and Alzheimer's disease is qualitatively distinct. *Journal of Neurology, Neurosurgery & Psychiatry* **75,** 382–387. ISSN: 0022-3050 (2004).

129. Wei, W., Nguyen, L. N., Kessels, H. W., Hagiwara, H., Sisodia, S. & Malinow, R. Amyloid beta from axons and dendrites reduces local spine number and plasticity. *Nature Neuroscience* **13,** 190–196. ISSN: 1097-6256 (2010).

130. Imlach, W., Beck, E., Choi, B., Lotti, F., Pellizzoni, L. & McCabe, B. SMN Is Required for Sensory-Motor Circuit Function in Drosophila. *Cell* **151,** 427–439. ISSN: 00928674 (2012).

131. Roselli, F. & Caroni, P. A Circuit Mechanism for Neurodegeneration. *Cell* **151,** 250–252. ISSN: 0092-8674 (2012).

132. Katsuno, M., Sahashi, K., Iguchi, Y. & Hashizume, A. Preclinical progression of neurodegenerative diseases. *Nagoya journal of medical science* **80,** 289–298. ISSN: 2186-3326 (2018).

133. Smith, M. L., Glass, G. V. & Miller, T. I. *The benefits of psychotherapy.* 269 (Johns Hopkins University Press, Baltimore, 1980).

134. Wampold, B. E. The Great Psychotherapy Debate: Models, Methods, and Findings (2001).

135. Chorpita, B. F. *et al.* Evidence-Based Treatments for Children and Adolescents: An Updated Review of Indicators of Efficacy and Effectiveness. *Clinical Psychology: Science and Practice* **18,** 154–172. ISSN: 09695893 (2011).

136. DeRubeis, R. J., Siegle, G. J. & Hollon, S. D. Cognitive therapy versus medication for depression: treatment outcomes and neural mechanisms. *Nature reviews. Neuroscience* **9,** 788–96. ISSN: 1471-0048 (2008).

137. Brooks, S. J. & Stein, D. J. A systematic review of the neural bases of psychotherapy for anxiety and related disorders. *Dialogues in clinical neuroscience* **17,** 261–79. ISSN: 1958-5969 (2015).

138. Shapiro, F. Eye movement desensitization: A new treatment for post-traumatic stress disorder. *Journal of Behavior Therapy and Experimental Psychiatry* **20,** 211–217. ISSN: 0005-7916 (1989).

139. Shapiro, F. *Eye movement desensitization and reprocessing (EMDR) therapy : basic principles, protocols, and procedures* 568. ISBN: 1462532764 (2001).

140. Foa, E. B., Keane, T. M., Friedman, M. J. & Cohen, J. A. *Effective treatments for PTSD: practice guidelines from the International Society for Traumatic Stress Studies* (Guilford Press, 2008).

141. WHO *et al. Guidelines for the management of conditions that are specifically related to stress* (World Health Organization, 2013).

142. Batelaan, N. M., Bosman, R. C., Muntingh, A., Scholten, W. D., Huijbregts, K. M. & van Balkom, A. J.L. M. Risk of relapse after antidepressant discontinuation in anxiety disorders, obsessive-compulsive disorder, and post-traumatic stress disorder: systematic review and meta-analysis of relapse prevention trials. *BMJ (Clinical research ed.)* **358,** j3927. ISSN: 1756-1833 (2017).

143. Wurtz, H., El-Khoury-Malhame, M., Wilhelm, F., Michael, T., Beetz, E., Roques, J., Reynaud, E., Courtin, J., Khalfa, S. & Herry, C. Preventing long-lasting fear recovery using bilateral alternating sensory stimulation: A translational study. *Neuroscience* **321,** 222–235. ISSN: 0306-4522 (2016).

144. Baek, J. *et al.* Neural circuits underlying a psychotherapeutic regimen for fear disorders. *Nature* **566,** 339. ISSN: 0028-0836 (2019).

145. Mingozzi, F. & High, K. A. Immune responses to AAV vectors: overcoming barriers to successful gene therapy. *Blood* **122,** 23–36. ISSN: 1528-0020 (2013).

146. Verma, I. M. Gene Therapy That Works. *Science* **341,** 853–855. ISSN: 0036-8075 (2013).

147. Costantini, L. C., Bakowska, J. C., Breakefield, X. O. & Isacson, O. Gene therapy in the CNS. *Gene Therapy* **7,** 93–109. ISSN: 0969-7128 (2000).

148. Hwu, W.-L., Muramatsu, S.-i., Tseng, S.-H., Tzen, K.-Y., Lee, N.-C., Chien, Y.-H., Snyder, R. O., Byrne, B. J., Tai, C.-H. & Wu, R.-M. Gene Therapy for Aromatic l-Amino Acid Decarboxylase Deficiency. *Science Translational Medicine* **4,** 134ra61–134ra61. ISSN: 1946-6234 (2012).

149. Yang, B. *et al.* Global CNS transduction of adult mice by intravenously delivered rAAVrh.8 and rAAVrh.10 and nonhuman primates by rAAVrh.10. *Molecular therapy : the journal of the American Society of Gene Therapy* **22,** 1299–1309. ISSN: 1525-0024 (2014).

150. Chan, K. Y. *et al.* Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. *Nature neuroscience* **20,** 1172–1179. ISSN: 1546-1726 (2017).

151. Biffi, A. *et al.* Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science (New York, N.Y.)* **341,** 1233158. ISSN: 1095-9203 (2013).

152. Serlin, Y., Shelef, I., Knyazer, B. & Friedman, A. Anatomy and physiology of the blood-brain barrier. *Seminars in cell & developmental biology* **38,** 2–6. ISSN: 1096-3634 (2015).

153. Demeule, M., Régina, A., Jodoin, J., Laplante, A., Dagenais, C., Berthelet, F., Moghrabi, A. & Béliveau, R. Drug transport to the brain: key roles for the efflux pump P-glycoprotein in the blood-brain barrier. *Vascular pharmacology* **38,** 339–48. ISSN: 1537-1891 (2002).

154. Loscher, W. & Potschka, H. Role of Multidrug Transporters in Pharmacoresistance to Antiepileptic Drugs. *Journal of Pharmacology and Experimental Therapeutics* **301,** 7–14. ISSN: 00223565 (2002).

155. Salunkhe, S. S., Bhatia, N. M., Kawade, V. S. & Bhatia, M. S. Development of Lipid Based Nanoparticulate Drug Delivery Systems and Drug Carrier Complexes for Delivery to Brain. *Journal of Applied Pharmaceutical Science* **5,** 110–129. ISSN: 2231-3354 (2015).

156. Kantor, B., McCown, T., Leone, P. & Gray, S. J. Clinical applications involving CNS gene transfer. *Advances in genetics* **87,** 71–124. ISSN: 0065-2660 (2014).

157. Dencker, D., Thomsen, M., Wörtwein, G., Weikop, P., Cui, Y., Jeon, J., Wess, J. & Fink-Jensen, A. Muscarinic Acetylcholine Receptor Subtypes as Potential Drug Targets for the Treatment of Schizophrenia, Drug Abuse and Parkinson's Disease. *ACS chemical neuroscience* **3,** 80–89. ISSN: 1948-7193 (2012).

158. Ito, H. T. & Schuman, E. M. Frequency-dependent signal transmission and modulation by neuromodulators. *Frontiers in neuroscience* **2,** 138–44. ISSN: 1662-453X (2008).

159. Kowiański, P., Lietzau, G., Czuba, E., Waśkow, M., Steliga, A. & Moryś, J. BDNF: A Key Factor with Multipotent Impact on Brain Signaling and Synaptic Plasticity. *Cellular and molecular neurobiology* **38,** 579–593. ISSN: 1573-6830 (2018).

160. Chivet, M., Hemming, F., Pernet-Gallay, K., Fraboulet, S. & Sadoul, R. Emerging role of neuronal exosomes in the central nervous system. *Frontiers in physiology* **3,** 145. ISSN: 1664-042X (2012).

161. Zappulli, V., Friis, K. P., Fitzpatrick, Z., Maguire, C. A. & Breakefield, X. O. Extracellular vesicles and intercellular communication within the nervous system. *Journal of Clinical Investigation* **126,** 1198–1207. ISSN: 0021-9738 (2016).

162. Horowitz, L. F., Montmayeur, J. P., Echelard, Y & Buck, L. B. A genetic approach to trace neural circuits. *Proceedings of the National Academy of Sciences of the United States of America* **96,** 3194–9. ISSN: 0027-8424 (1999).

163. Itaya, S. K. & Van Hoesen, G. W. *WGA-HRP as a transneuronal marker in the visual pathways of monkey and rat* 1982.

164. Baker, H. & Spencer, R. F. Transneuronal transport of peroxidase-conjugated wheat germ agglutinin (WGA-HRP) from the olfactory epithelium to the brain of the adult rat. *Experimental Brain Research* **63,** 461–473. ISSN: 0014-4819 (1986).

165. Margolis, F. L. & Getchell, T. V. *Molecular neurobiology of the olfactory system : molecular, membranous, and cytological studies* 379. ISBN: 9781461309895 (Plenum Press, 1988).

166. Carpenter, M. B. & Cowie, R. J. Transneuronal transport in the vestibular and auditory systems of the squirrel monkey and the arctic ground squirrel. I. Vestibular system. *Brain research* **358,** 249–63. ISSN: 0006-8993 (1985).

167. Schwab, M. E., Suda, K & Thoenen, H. Selective retrograde transsynaptic transfer of a protein, tetanus toxin, subsequent to its retrograde axonal transport. *The Journal of cell biology* **82,** 798–810. ISSN: 0021-9525 (1979).

168. Lai, B.-Q., Qiu, X.-C., Zhang, K., Zhang, R.-Y., Jin, H., Li, G., Shen, H.-Y., Wu, J.-L., Ling, E.-A. & Zeng, Y.-S. Cholera Toxin B Subunit Shows Transneuronal Tracing after Injection in an Injured Sciatic Nerve. *PloS one* **10,** e0144030. ISSN: 1932-6203 (2015).

169. Sugiyama, S., Di Nardo, A. A., Aizawa, S., Matsuo, I., Volovitch, M., Prochiantz, A. & Hensch, T. K. Experience-dependent transfer of Otx2 homeoprotein into the visual cortex activates postnatal plasticity. *Cell* **134,** 508–20. ISSN: 1097-4172 (2008).

170. Ashley, J., Cordy, B., Lucia, D., Fradkin, L. G., Budnik, V. & Thomson, T. Retrovirus-like Gag Protein Arc1 Binds RNA and Traffics across Synaptic Boutons. *Cell* **172,** 262–274.e11. ISSN: 00928674 (2018).

171. Guidolin, D., Marcoli, M., Maura, G. & Agnati, L. F. New dimensions of connectomics and network plasticity in the central nervous system. *Reviews in the Neurosciences* **0.** ISSN: 2191-0200. doi:10.1515/revneuro-2016-0051 (2016).

172. Liu, L., Drouet, V., Wu, J. W., Witter, M. P., Small, S. A., Clelland, C. & Duff, K. Trans-synaptic spread of tau pathology in vivo. *PloS one* **7,** e31302. ISSN: 1932-6203 (2012).

173. Dujardin, S. *et al.* Neuron-to-neuron wild-type Tau protein transfer through a trans-synaptic mechanism: relevance to sporadic tauopathies. *Acta Neuropathologica Communications* **2,** 14. ISSN: 2051-5960 (2014).

174. Grafstein, B. Transneuronal transfer of radioactivity in the central nervous system. *Science (New York, N.Y.)* **172,** 177–9. ISSN: 0036-8075 (1971).

175. Webster, M. J., Ungerleider, L. G. & Bachevalier, J. Connections of inferior temporal areas TE and TEO with medial temporal-lobe structures in infant and adult monkeys. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **11,** 1095–116. ISSN: 0270-6474 (1991).

176. Brunet, I., Di Nardo, A. A., Sonnier, L., Beurdeley, M. & Prochiantz, A. The topological role of homeoproteins in the developing central nervous system. *Trends in Neurosciences* **30,** 260–267. ISSN: 01662236 (2007).

177. Simon, H. H., Saueressig, H, Wurst, W, Goulding, M. D. & O'Leary, D. D. Fate of midbrain dopaminergic neurons controlled by the engrailed genes. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **21,** 3126–34. ISSN: 1529-2401 (2001).

178. Joliot, A *et al.* Identification of a signal sequence necessary for the unconventional secretion of Engrailed homeoprotein. *Current biology : CB* **8,** 856–63. ISSN: 0960-9822 (1998).

179. Prochiantz, A. & Joliot, A. Can transcription factors function as cell–cell signalling molecules? *Nature Reviews Molecular Cell Biology* **4,** 814–819. ISSN: 1471-0072 (2003).

180. Niessing, D., Driever, W., Sprenger, F., Taubert, H., Jäckle, H. & Rivera-Pomar, R. Homeodomain Position 54 Specifies Transcriptional versus Translational Control by Bicoid. *Molecular Cell* **5,** 395–401. ISSN: 10972765 (2000).

181. Joliot, A, Pernellet, C, Deagostini-Bazint, H, Prochiantz & A. Antennapedia homeobox peptide regulates neural morphogenesis. *Neurobiology* **88,** 1864–1868 (1991).

182. Marie, B., Cruz-Orengo, L. & Blagburn, J. M. Persistent engrailed Expression Is Required to Determine Sensory Axon Trajectory, Branching, and Target Choice. *The Journal of Neuroscience* **22,** 832–841. ISSN: 0270-6474 (2018).

183. Gallo, F. T., Katche, C., Morici, J. F., Medina, J. H. & Weisstaub, N. V. Immediate Early Genes, Memory and Psychiatric Disorders: Focus on c-Fos, Egr1 and Arc. *Frontiers in Behavioral Neuroscience* **12,** 79. ISSN: 1662-5153 (2018).

184. Vazdarjanova, A., McNaughton, B. L., Barnes, C. A., Worley, P. F. & Guzowski, J. F. Experience-dependent coincident expression of the effector immediate-early genes arc and Homer 1a in hippocampal and neocortical neuronal networks. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **22,** 10067–71. ISSN: 1529-2401 (2002).

185. Alhowikan, A. M. Activity-Regulated Cytoskeleton-Associated Protein Dysfunction May Contribute to Memory Disorder and Earlier Detection of Autism Spectrum Disorders. *Medical Principles and Practice* **25,** 350–354. ISSN: 1011-7571 (2016).

186. Fromer, M. *et al.* De novo mutations in schizophrenia implicate synaptic networks. *Nature* **506,** 179–184. ISSN: 0028-0836 (2014).

187. Pastuzyn, E. D. *et al.* The Neuronal Gene Arc Encodes a Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer. *Cell* **172,** 275–288.e18. ISSN: 00928674 (2018).

188. Kubala, M. H., Kovtun, O., Alexandrov, K. & Collins, B. M. Structural and thermodynamic analysis of the GFP:GFP-nanobody complex. *Protein science : a publication of the Protein Society* **19,** 2389–401. ISSN: 1469-896X (2010).

189. Wickersham, I. R., Sullivan, H. A. & Seung, H. S. Production of glycoprotein-deleted rabies viruses for monosynaptic tracing and high-level gene expression in neurons. *Nature Protocols* **5,** 595–606. ISSN: 1754-2189 (2010).

190. Reardon, T. R., Murray, A. J., Turi, G. F., Wirblich, C., Croce, K. R., Schnell, M. J., Jessell, T. M. & Losonczy, A. Rabies Virus CVS-N2cΔG Strain Enhances Retrograde Synaptic Transfer and Neuronal Viability. *Neuron* **89,** 711–724. ISSN: 0896-6273 (2016).

191. Ciabatti, E., González-Rueda, A., Mariotti, L., Morgese, F. & Tripodi, M. Life-Long Genetic and Functional Access to Neural Circuits Using Self-Inactivating Rabies Virus. *Cell* **170,** 382–392.e14. ISSN: 0092-8674 (2017).

192. Matsuyama, M., Jin, L., Lavin, T. K., Sullivan, H. A., Hou, Y., Lea, N. E., Pruner, M. T., Ferdínez, M. L. D. & Wickersham, I. R. "Self-inactivating" rabies viruses are just first-generation, ΔG rabies viruses. *bioRxiv,* 550640 (2019).

193. Wang, Z. W., Sarmento, L., Wang, Y., Li, X.-q., Dhingra, V., Tseggai, T., Jiang, B. & Fu, Z. F. Attenuated rabies virus activates, while pathogenic rabies virus evades, the host innate immune responses in the central nervous system. *Journal of virology* **79,** 12554–65. ISSN: 0022-538X (2005).

194. Miao, F.-m., Zhang, S.-f., Wang, S.-c., Liu, Y., Zhang, F. & Hu, R.-l. Comparison of immune responses to attenuated rabies virus and street virus in mouse brain. *Archives of Virology* **162,** 247–257. ISSN: 0304-8608 (2017).

195. Zhang, D. *et al.* Genome-Wide Transcriptional Profiling Reveals Two Distinct Outcomes in Central Nervous System Infections of Rabies Virus. *Frontiers in Microbiology* **7,** 751. ISSN: 1664-302X (2016).

196. Wang, C. & Mei, L. In utero electroporation in mice. *Methods in molecular biology (Clifton, N.J.)* **1018,** 151–63. ISSN: 1940-6029 (2013).

197. Mi, H., Muruganujan, A., Huang, X., Ebert, D., Mills, C., Guo, X. & Thomas, P. D. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nature Protocols* **14,** 703–721. ISSN: 1754-2189 (2019).

198. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: A tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10,** 48. ISSN: 14712105 (2009).

199. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* **26,** 589–95. ISSN: 1367-4811 (2010).

200. Crick, F. H. Thinking about the brain. *Scientific American* **241,** 219–32. ISSN: 0036-8733 (1979).

201. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. *The Structure of the Nervous System of the Nematode Caenorhabditis elegans* tech. rep. 1165 (1986), 1–340. doi:10.1098/rstb.1986.0056.

202. Zheng, Z., Lauritzen, J. S., Perlman, E., Saalfeld, S., Fetter, R. D. & Bock Correspondence, D. D. *A Complete Electron Microscopy Volume of the Brain of Adult Drosophila melanogaster* tech. rep. (2018), 730–743. doi:10.1016/j.cell.2018.06.019.

203. Stepien, A. E., Tripodi, M. & Arber, S. Monosynaptic rabies virus reveals premotor network organization and synaptic specificity of cholinergic partition cells. *Neuron* **68,** 456–72. ISSN: 1097-4199 (2010).

204. Tripodi, M., Stepien, A. E. & Arber, S. Motor antagonism exposed by spatial segregation and timing of neurogenesis. *Nature* **479,** 61–66. ISSN: 0028-0836 (2011).

205. Augustine, V., Gokce, S. K., Lee, S., Wang, B., Davidson, T. J., Reimann, F., Gribble, F., Deisseroth, K., Lois, C. & Oka, Y. Hierarchical neural architecture underlying thirst regulation. *Nature* **555,** 204–209. ISSN: 1476-4687 (2018).

206. Kaelberer, M. M., Buchanan, K. L., Klein, M. E., Barth, B. B., Montoya, M. M., Shen, X. & Bohórquez, D. V. A gut-brain neural circuit for nutrient sensory transduction. *Science (New York, N.Y.)* **361,** eaat5236. ISSN: 1095-9203 (2018).

207. Evans, D. A., Stempel, A. V., Vale, R., Ruehle, S., Lefler, Y. & Branco, T. A synaptic threshold mechanism for computing escape decisions. *Nature* **558,** 590–594. ISSN: 1476-4687 (2018).

208. Alivisatos, A., Chun, M., Church, G., Greenspan, R., Roukes, M. & Yuste, R. The Brain Activity Map Project and the Challenge of Functional Connectomics. *Neuron* **74,** 970–974. ISSN: 0896-6273 (2012).

209. Alivisatos, A. P. *et al.* Nanotools for neuroscience and brain activity mapping. *ACS nano* **7,** 1850–66. ISSN: 1936-086X (2013).

210. Finke, S. & Conzelmann, K.-K. Replication strategies of rabies virus. *Virus Research* **111,** 120–131. ISSN: 0168-1702 (2005).

211. Mebatsion, T., König, M. & Conzelmann, K.-K. Budding of Rabies Virus Particles in the Absence of the Spike Glycoprotein. *Cell* **84,** 941–951. ISSN: 0092-8674 (1996).

212. Fadai-Ghotbi, B., Natelson, B., Tsiang, H., Etessami, R., Ceccaldi, P.-E. & Conzelmann, K.-K. Spread and pathogenic characteristics of a G-deficient rabies virus recombinant: an in vitro and in vivo study. *Journal of General Virology* **81,** 2147–2153. ISSN: 0022-1317 (2000).

213. Schnell, M. J., Mebatsion, T. & Conzelmann1, K.-K. *Infectious rabies viruses from cloned cDNA* tech. rep. 18 (1994), 4195–4203.

214. Wickersham, I. R., Lyon, D. C., Barnard, R. J. O., Mori, T., Finke, S., Conzelmann, K.-K., Young, J. A. T. & Callaway, E. M. Monosynaptic restriction of transsynaptic tracing from single, genetically targeted neurons. *Neuron* **53,** 639–47. ISSN: 0896-6273 (2007).

215. Miyamichi, K. *et al.* Cortical representations of olfactory input by trans-synaptic tracing. *Nature* **472,** 191–6. ISSN: 1476-4687 (2011).

216. Hornung, V. *et al.* 5'-Triphosphate RNA is the ligand for RIG-I. *Science (New York, N.Y.)* **314,** 994–7. ISSN: 1095-9203 (2006).

217. Prehaud, C., Megret, F., Lafage, M. & Lafon, M. Virus Infection Switches TLR-3-Positive Human Neurons To Become Strong Producers of Beta Interferon. *Journal of Virology* **79,** 12893–12904. ISSN: 0022-538X (2005).

218. Jackson, A. C., Rossiter, J. P. & Lafon, M. Expression of Toll-like receptor 3 in the human cerebellar cortex in rabies, herpes simplex encephalitis, and other neurological diseases. *Journal of Neurovirology* **12,** 229–234. ISSN: 1355-0284 (2006).

219. Ito, N., Moseley, G. W. & Sugiyama, M. The importance of immune evasion in the pathogenesis of rabies virus. *The Journal of veterinary medical science* **78,** 1089–98. ISSN: 1347-7439 (2016).

220. Suja, M. S., Mahadevan, A., Madhusudana, S. N. & Shankar, S. K. Role of apoptosis in rabies viral encephalitis: a comparative study in mice, canine, and human brain with a review of literature. *Pathology research international* **2011,** 374286. ISSN: 2042-003X (2011).

221. Jackson, A. C., Randle, E., Lawrance, G. & Rossiter, J. P. Neuronal apoptosis does not play an important role in human rabies encephalitis. *Journal of Neurovirology* **14,** 368–375. ISSN: 1355-0284 (2008).

222. Yan, X, Prosniak, M, Curtis, M. T., Weiss, M. L., Faber, M, Dietzschold, B & Fu, Z. F. Silver-haired bat rabies virus variant does not induce apoptosis in the brain of

experimentally infected mice. *Journal of Neurovirology* **7,** 518–527. ISSN: 1355-0284 (2001).

223. Sarmento, L., Tseggai, T., Dhingra, V. & Fu, Z. F. Rabies virus-induced apoptosis involves caspase-dependent and caspase-independent pathways. *Virus Research* **121,** 144–151. ISSN: 01681702 (2006).

224. Roy, A., Phares, T. W., Koprowski, H. & Hooper, D. C. Failure to open the blood-brain barrier and deliver immune effectors to central nervous system tissues leads to the lethal outcome of silver-haired bat rabies virus infection. *Journal of virology* **81,** 1110–8. ISSN: 0022-538X (2007).

225. Gnanadurai, C. W. & Fu, Z. F. CXCL10 and blood-brain barrier modulation in rabies virus infection. *Oncotarget* **7,** 10694–5. ISSN: 1949-2553 (2016).

226. Shimizu, K., Ito, N., Mita, T., Yamada, K., Hosokawa-Muto, J., Sugiyama, M. & Minamoto, N. Involvement of nucleoprotein, phosphoprotein, and matrix protein genes of rabies virus in virulence for adult mice. *Virus Research* **123,** 154–160. ISSN: 0168-1702 (2007).

227. Masatani, T., Ito, N., Shimizu, K., Ito, Y., Nakagawa, K., Sawaki, Y., Koyama, H. & Sugiyama, M. Rabies virus nucleoprotein functions to evade activation of the RIG-I-mediated antiviral response. *Journal of virology* **84,** 4002–12. ISSN: 1098-5514 (2010).

228. Masatani, T., Ito, N., Shimizu, K., Ito, Y., Nakagawa, K., Abe, M., Yamaoka, S. & Sugiyama, M. Amino acids at positions 273 and 394 in rabies virus nucleoprotein are important for both evasion of host RIG-I-mediated antiviral response and pathogenicity. *Virus Research* **155,** 168–174. ISSN: 0168-1702 (2011).

229. Brzózka, K., Finke, S. & Conzelmann, K.-K. Identification of the rabies virus alpha/beta interferon antagonist: phosphoprotein P interferes with phosphorylation of interferon regulatory factor 3. *Journal of virology* **79,** 7673–81. ISSN: 0022-538X (2005).

230. Vidy, A., Chelbi-Alix, M. & Blondel, D. Rabies virus P protein interacts with STAT1 and inhibits interferon signal transduction pathways. *Journal of virology* **79,** 14411–20. ISSN: 0022-538X (2005).

231. Brzózka, K., Finke, S. & Conzelmann, K.-K. Inhibition of interferon signaling by rabies virus phosphoprotein P: activation-dependent binding of STAT1 and STAT2. *Journal of virology* **80,** 2675–83. ISSN: 0022-538X (2006).

232. Vidy, A., El Bougrini, J., Chelbi-Alix, M. K. & Blondel, D. The nucleocytoplasmic rabies virus P protein counteracts interferon signaling by inhibiting both nuclear accumulation and DNA binding of STAT1. *Journal of virology* **81,** 4255–63. ISSN: 0022-538X (2007).

233. Luco, S., Delmas, O., Vidalain, P. O., Tangy, F., Weil, R. & Bourhy, H. RelAp43, a Member of the NF-κB Family Involved in Innate Immune Response against Lyssavirus Infection. *PLoS Pathogens* **8** (ed Schnell, M. J.) e1003060. ISSN: 15537366 (2012).

234. Zhang, G., Wang, H., Mahmood, F. & Fu, Z. F. Rabies virus glycoprotein is an important determinant for the induction of innate immune responses and the pathogenic mechanisms. *Veterinary microbiology* **162,** 601–13. ISSN: 1873-2542 (2013).

235. Morimoto, K, Hooper, D. C., Spitsin, S, Koprowski, H & Dietzschold, B. Pathogenicity of different rabies virus variants inversely correlates with apoptosis and rabies virus glycoprotein expression in infected primary neuron cultures. *Journal of virology* **73,** 510–8. ISSN: 0022-538X (1999).

236. Li, C. *et al.* Deficient Incorporation of Rabies Virus Glycoprotein into Virions Enhances Virus-Induced Immune Evasion and Viral Pathogenicity. *Viruses* **11,** 218. ISSN: 1999-4915 (2019).

237. Morimoto, K, Hooper, D. C., Carbaugh, H, Fu, Z. F., Koprowski, H & Dietzschold, B. Rabies virus quasispecies: implications for pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **95,** 3152–6. ISSN: 0027-8424 (1998).

238. Combe, M. & Sanjuán, R. Variation in RNA Virus Mutation Rates across Host Cells. *PLoS Pathogens* **10** (ed Lauring, A.) e1003855. ISSN: 1553-7374 (2014).

239. Holmes, E. C., Woelk, C. H., Kassis, R. & Bourhy, H. Genetic Constraints and the Adaptive Evolution of Rabies Virus in Nature. doi:10.1006/viro.2001.1271 (2002).

240. Elena, S. F. & Sanjuán, R. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. *Journal of virology* **79,** 11555–8. ISSN: 0022-538X (2005).

241. Duffy, S. Why are RNA virus mutation rates so damn high? *PLOS Biology* **16,** e3000003. ISSN: 1545-7885 (2018).

242. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. arXiv: 1207.3907 (2012).

243. Osakada, F. & Callaway, E. M. Design and generation of recombinant rabies virus vectors. *Nature Protocols* **8,** 1583–1601. ISSN: 1754-2189 (2013).

244. Kim, E., Jacobs, M., Ito-Cole, T. & Callaway, E. Improved Monosynaptic Neural Circuit Tracing Using Engineered Rabies Virus Glycoproteins. *Cell Reports* **15,** 692–699. ISSN: 2211-1247 (2016).

245. Nazmi, A., Dutta, K. & Basu, A. RIG-I Mediates Innate Immune Response in Mouse Neurons Following Japanese Encephalitis Virus Infection. *PLoS ONE* **6** (ed Schneider, B. S.) e21761. ISSN: 1932-6203 (2011).

246. Faul, E. J., Wanjalla, C. N., Suthar, M. S., Gale, M., Wirblich, C. & Schnell, M. J. Rabies Virus Infection Induces Type I Interferon Production in an IPS-1 Dependent Manner While Dendritic Cell Activation Relies on IFNAR Signaling. *PLoS Pathogens* **6** (ed Garcia-Sastre, A.) e1001016. ISSN: 1553-7374 (2010).

247. Park, K. M. & Bowers, W. J. Tumor necrosis factor-alpha mediated signaling in neuronal homeostasis and dysfunction. *Cellular Signalling* **22,** 977–983. ISSN: 08986568 (2010).

248. Kuniyoshi, K., Takeuchi, O., Pandey, S., Satoh, T., Iwasaki, H., Akira, S. & Kawai, T. Pivotal role of RNA-binding E3 ubiquitin ligase MEX3C in RIG-I-mediated antiviral innate immunity. *Proceedings of the National Academy of Sciences of the United States of America* **111,** 5646–51. ISSN: 1091-6490 (2014).

249. Langereis, M. A., Feng, Q. & van Kuppeveld, F. J. MDA5 localizes to stress granules, but this localization is not required for the induction of type I interferon. *Journal of virology* **87,** 6314–25. ISSN: 1098-5514 (2013).

250. Onomoto, K. *et al.* Critical role of an antiviral stress granule containing RIG-I and PKR in viral detection and innate immunity. *PloS one* **7,** e43031. ISSN: 1932-6203 (2012).

251. Río-Hortega, P. El tercer elemento de los centros nerviosos I-IV. *Boletín Soc. Española Biol.* **8,** 67–166 (1919).

252. Ray, N. B., Power, C, Lynch, W. P., Ewalt, L. C. & Lodmell, D. L. Rabies viruses infect primary cultures of murine, feline, and human microglia and astrocytes. *Archives of virology* **142,** 1011–9. ISSN: 0304-8608 (1997).

253. Rock, R. B., Gekker, G., Hu, S., Sheng, W. S., Cheeran, M., Lokensgard, J. R. & Peterson, P. K. Role of microglia in central nervous system infections. *Clinical microbiology reviews* **17,** 942–64, table of contents. ISSN: 0893-8512 (2004).

254. Fernández-Arjona, M. d. M., Grondona, J. M., Granados-Durán, P., Fernández-Llebrez, P. & López-Ávalos, M. D. Microglia Morphological Categorization in a Rat Model of Neuroinflammation by Hierarchical Cluster and Principal Components Analysis. *Frontiers in Cellular Neuroscience* **11,** 235. ISSN: 1662-5102 (2017).

255. Eng, L. F., Ghirnikar, R. S. & Lee, Y. L. Glial fibrillary acidic protein: GFAP-thirty-one years (1969-2000). *Neurochemical research* **25,** 1439–51. ISSN: 0364-3190 (2000).

256. Zhang, D., Hu, X., Qian, L., O'Callaghan, J. P. & Hong, J.-S. Astrogliosis in CNS pathologies: is there a role for microglia? *Molecular neurobiology* **41,** 232–41. ISSN: 1559-1182 (2010).

257. Pozner, R. G., Collado, S., de Giusti, C. J., Ure, A. E., Biedma, M. E., Romanowski, V., Schattner, M. & Gómez, R. M. Astrocyte response to Junín virus infection. *Neuroscience Letters* **445,** 31–35. ISSN: 0304-3940 (2008).

258. Gardner, J., Borgmann, K., Deshpande, M. S., Dhar, A., Wu, L., Persidsky, R. & Ghorpade, A. Potential mechanisms for astrocyte-TIMP-1 downregulation in chronic inflammatory diseases. *Journal of Neuroscience Research* **83,** 1281–1292. ISSN: 0360-4012 (2006).

259. Shrikant, P & Benveniste, E. N. The central nervous system as an immunocompetent organ: role of glial cells in antigen presentation. *Journal of immunology (Baltimore, Md. : 1950)* **157,** 1819–22. ISSN: 0022-1767 (1996).

260. Pfefferkorn, C., Kallfass, C., Lienenklaus, S., Spanier, J., Kalinke, U., Rieder, M., Conzelmann, K.-K., Michiels, T. & Staeheli, P. Abortively Infected Astrocytes Appear To Represent the Main Source of Interferon Beta in the Virus-Infected Brain. *Journal of virology* **90,** 2031–8. ISSN: 1098-5514 (2016).

261. Tian, B. *et al.* Lab-Attenuated Rabies Virus Causes Abortive Infection and Induces Cytokine Expression in Astrocytes by Activating Mitochondrial Antiviral-Signaling Protein Signaling Pathway. *Frontiers in Immunology* **8,** 2011. ISSN: 1664-3224 (2018).

262. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15,** 550. ISSN: 1474-760X (2014).

263. Ning, S, Pagano, J. S. & Barber, G. N. IRF7: activation, regulation, modification and function. *Genes & Immunity* **12,** 399–414. ISSN: 1466-4879 (2011).

264. Kuang, Y., Lackay, S. N., Zhao, L. & Fu, Z. F. Role of chemokines in the enhancement of BBB permeability and inflammatory infiltration after rabies virus infection. doi:10.1016/j.virusres.2009.03.014.

265. Müller, M., Carter, S., Hofer, M. J. & Campbell, I. L. Review: The chemokine receptor CXCR3 and its ligands CXCL9, CXCL10 and CXCL11 in neuroimmunity - a tale of conflict and conundrum. *Neuropathology and Applied Neurobiology* **36,** 368–387. ISSN: 03051846 (2010).

266. Chai, Q., He, W. Q., Zhou, M., Lu, H. & Fu, Z. F. Enhancement of Blood-Brain Barrier Permeability and Reduction of Tight Junction Protein Expression Are Modulated by Chemokines/Cytokines Induced by Rabies Virus Infection. *Journal of Virology* **88,** 4698–4710. ISSN: 0022-538X (2014).

267. Huppert, J. *et al.* Cellular mechanisms of IL-17-induced blood-brain barrier disruption. *The FASEB Journal* **24,** 1023–1034. ISSN: 0892-6638 (2010).

268. Jongsma, M. L., Guarda, G. & Spaapen, R. M. The regulatory network behind MHC class I expression. *Molecular Immunology*. ISSN: 0161-5890. doi:10.1016/J.MOLIMM.2017.12.005 (2017).

269. Saha, S., Murthy, S. & Rangarajan, P. N. Identification and characterization of a virus-inducible non-coding RNA in mouse brain. *Journal of General Virology* **87,** 1991–1995. ISSN: 0022-1317 (2006).

270. Ouyang, J. *et al.* NRAV, a long noncoding RNA, modulates antiviral responses through suppression of interferon-stimulated gene transcription. *Cell host & microbe* **16,** 616–26. ISSN: 1934-6069 (2014).

271. Menegas, W., Akiti, K., Amo, R., Uchida, N. & Watabe-Uchida, M. Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. *Nature Neuroscience* **21,** 1421–1430. ISSN: 1097-6256 (2018).

272. Van der Oost, J. New Tool for Genome Surgery. *Science* **339,** 768–770. ISSN: 0036-8075 (2013).

273. Baker Monya. Method of the Year 2011: Gene-editing nucleases. *Nature Video* **9,** 2012. ISSN: 1548-7091 (2011).

274. Van Soolingen, D, de Haas, P. E., Hermans, P. W., Groenen, P. M. & van Embden, J. D. Comparison of various repetitive DNA elements as genetic markers for strain differentiation and epidemiology of Mycobacterium tuberculosis. *Journal of clinical microbiology* **31,** 1987–95. ISSN: 0095-1137 (1993).

275. Mojica, F. J. & Montoliu, L. On the Origin of CRISPR-Cas Technology: From Prokaryotes to Mammals. *Trends in Microbiology* **24,** 811–820. ISSN: 0966-842X (2016).

276. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *Journal of Molecular Evolution* **60,** 174–182. ISSN: 1432-1432 (2005).

277. Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A. & Horvath, P. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science* **315,** 1709–1712. ISSN: 0036-8075 (2007).

278. Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., Dickman, M. J., Makarova, K. S., Koonin, E. V. & van der Oost, J. Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science* **321,** 960–964. ISSN: 0036-8075 (2008).

279. Marraffini, L. A. & Sontheimer, E. J. CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. *Science* **322,** 1843–1845. ISSN: 0036-8075 (2008).

280. Garneau, J. E., Dupuis, M.-È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A. H. & Moineau, S. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468,** 67–71. ISSN: 0028-0836 (2010).

281. Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P. & Siksnys, V. The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic acids research* **39,** 9275–82. ISSN: 1362-4962 (2011).

282. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences* **109,** E2579–E2586. ISSN: 0027-8424 (2012).

283. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339,** 819–823. ISSN: 0036-8075 (2013).

284. Jinek, M., East, A., Cheng, A., Lin, S., Ma, E. & Doudna, J. RNA-programmed genome editing in human cells. *eLife* **2.** ISSN: 2050-084X. doi:10.7554/eLife.00471 (2013).

285. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337,** 816–821. ISSN: 0036-8075 (2012).

286. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339,** 819–823. ISSN: 0036-8075 (2013).

287. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533,** 420–424. ISSN: 0028-0836 (2016).

288. Cullot, G. *et al.* CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. *Nature Communications* **10,** 1136. ISSN: 2041-1723 (2019).

289. Mullen, R. J., Buck, C. R. & Smith, A. M. NeuN, a neuronal specific nuclear protein in vertebrates. *Development* **116,** 201–211. ISSN: 0950-1991 (1992).

290. Haeussler, M. *et al.* Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biology* **17,** 148. ISSN: 1474-760X (2016).

291. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology* **34,** 184–191. ISSN: 1087-0156 (2016).

292. Yeung, A. T., Hattangadi, D., Blakesley, L. & Nicolas, E. Enzymatic mutation detection technologies. *BioTechniques* **38,** 749–758. ISSN: 0736-6205 (2005).

293. Conzelmann, K. K., Cox, J. H., Schneider, L. G. & Thiel, H. J. Molecular cloning and complete nucleotide sequence of the attenuated rabies virus SAD B19. *Virology* **175,** 485–99. ISSN: 0042-6822 (1990).

294. Osakada, F., Mori, T., Cetin, A. H., Marshel, J. H., Virgen, B. & Callaway, E. M. New rabies virus variants for monitoring and manipulating activity and gene expression in defined neural circuits. *Neuron* **71,** 617–31. ISSN: 1097-4199 (2011).

295. Briner, A., Donohoue, P., Gomaa, A., Selle, K., Slorach, E., Nye, C., Haurwitz, R., Beisel, C., May, A. & Barrangou, R. Guide RNA Functional Modules Direct Cas9 Activity and Orthogonality. *Molecular Cell* **56,** 333–339. ISSN: 1097-2765 (2014).

296. Finke, S, Cox, J. H. & Conzelmann, K. K. Differential transcription attenuation of rabies virus genes by intergenic regions: generation of recombinant viruses overexpressing the polymerase gene. *Journal of virology* **74,** 7261–9. ISSN: 0022-538X (2000).

297. Luo, J., Zhao, J., Tian, Q., Mo, W., Wang, Y., Chen, H. & Guo, X. A recombinant rabies virus carrying GFP between N and P affects viral transcription in vitro. *Virus Genes* **52,** 379–387. ISSN: 0920-8569 (2016).

298. Abraham, G & Banerjee, A. K. Sequential transcription of the genes of vesicular stomatitis virus. *Proceedings of the National Academy of Sciences of the United States of America* **73,** 1504–8. ISSN: 0027-8424 (1976).

299. Iverson, L. E. & Rose, J. K. Localized attenuation and discontinuous synthesis during vesicular stomatitis virus transcription. *Cell* **23,** 477–484. ISSN: 0092-8674 (1981).

300. Pringle, C. R. & Easton, A. J. Monopartite Negative Strand RNA Genomes. *Seminars in Virology* **8,** 49–57. ISSN: 1044-5773 (1997).

301. Conzelmann, K.-K. NONSEGMENTED NEGATIVE-STRAND RNA VIRUSES: Genetics and Manipulation of Viral Genomes. *Annual Review of Genetics* **32,** 123–162. ISSN: 0066-4197 (1998).

302. Leid, M., Ishmael, J. E., Avram, D., Shepherd, D., Fraulob, V. & Dollé, P. CTIP1 and CTIP2 are differentially expressed during mouse embryogenesis. *Gene expression patterns : GEP* **4,** 733–9. ISSN: 1567-133X (2004).

303. Weinberg, B. H., Pham, N. T. H., Caraballo, L. D., Lozanoski, T., Engel, A., Bhatia, S. & Wong, W. W. Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. *Nature Biotechnology* **35,** 453–462. ISSN: 1087-0156 (2017).

304. Moll, J. R., Ruvinov, S. B., Pastan, I & Vinson, C. Designed heterodimerizing leucine zippers with a ranger of pIs and stabilities up to 10(-15) M. *Protein science : a publication of the Protein Society* **10,** 649–55. ISSN: 0961-8368 (2001).

305. Guschin, D. Y., Waite, A. J., Katibah, G. E., Miller, J. C., Holmes, M. C. & Rebar, E. J., 247–256 (Humana Press, Totowa, NJ, 2010). doi:10.1007/978-1-60761-753-2_15.

306. Xie, H. *et al.* SaCas9 Requires 5-NNGRRT-3 PAM for Sufficient Cleavage and Possesses Higher Cleavage Activity than SpCas9 or FnCpf1 in Human Cells. *Biotechnology Journal* **13,** 1700561. ISSN: 18606768 (2018).

307. Itaya, S. K. & Van Hoesen, G. W. WGA-HRP as a transneuronal marker in the visual pathways of monkey and rat. *Brain Research* **236,** 199–204. ISSN: 0006-8993 (1982).

308. Martersteck, E. M. *et al.* Diverse Central Projection Patterns of Retinal Ganglion Cells. *CellReports* **18,** 2058–2072 (2017).

309. Ruggeri, M., Tsechpenakis, G., Jiao, S., Jockovich, M. E., Cebulla, C., Hernandez, E., Murray, T. G. & Puliafito, C. A. Retinal tumor imaging and volume quantification in mouse model using spectral-domain optical coherence tomography. *Optics express* **17,** 4074–83. ISSN: 1094-4087 (2009).

310. Seabrook, T. A., Burbridge, T. J., Crair, M. C. & Huberman, A. D. Architecture, Function, and Assembly of the Mouse Visual System. *Annual Review of Neuroscience* **40,** 499–538. ISSN: 0147-006X (2017).

311. Martin, K. R. G., Quigley, H. A., Zack, D. J., Levkovitch-Verbin, H., Kielczewski, J., Valenta, D., Baumrind, L., Pease, M. E., Klein, R. L. & Hauswirth, W. W. Gene Therapy with Brain-Derived Neurotrophic Factor As a Protection: Retinal Ganglion Cells in a Rat Glaucoma Model. *Investigative Opthalmology & Visual Science* **44,** 4357. ISSN: 1552-5783 (2003).

312. Harvey, A. R., Kamphuis, W, Eggers, R, Symons, N. A., Blits, B, Niclou, S, Boer, G. J. & Verhaagen, J. Intravitreal injection of adeno-associated viral vectors results in the transduction of different types of retinal neurons in neonatal and adult rats: a comparison with lentiviral vectors. *Molecular and cellular neurosciences* **21,** 141–57. ISSN: 1044-7431 (2002).

313. Smith, C. A. & Chauhan, B. C. In vivo imaging of adeno-associated viral vector labelled retinal ganglion cells. *Scientific reports* **8,** 1490. ISSN: 2045-2322 (2018).

314. Tervo, D. G. R. *et al.* A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons. *Neuron* **92,** 372–382. ISSN: 1097-4199 (2016).

315. Kim, D. I., Jensen, S. C., Noble, K. A., KC, B., Roux, K. H., Motamedchaboki, K. & Roux, K. J. An improved smaller biotin ligase for BioID proximity labeling. *Molecular Biology of the Cell* **27** (ed Zheng, Y.) 1188–1196. ISSN: 1059-1524 (2016).

316. Uezu, A., Kanak, D. J., Bradshaw, T. W., Soderblom, E. J., Catavero, C. M., Burette, A. C., Weinberg, R. J. & Soderling, S. H. Identification of an elaborate complex mediating postsynaptic inhibition. *Science* **353,** 1123–1129. ISSN: 10959203 (2016).

317. Mahdavi, A., Hamblin, G. D., Jindal, G. A., Bagert, J. D., Dong, C., Sweredoski, M. J., Hess, S., Schuman, E. M. & Tirrell, D. A. Engineered Aminoacyl-tRNA Synthetase for Cell-Selective Analysis of Mammalian Protein Synthesis. *Journal of the American Chemical Society* **138,** 4278–81. ISSN: 1520-5126 (2016).

318. Alvarez-Castelao, B. *et al.* Cell-type-specific metabolic labeling of nascent proteomes in vivo. doi:10.1038/nbt.4016 (2017).

319. Yang, N., Ng, Y. H., Pang, Z. P., Südhof, T. C. & Wernig, M. Induced neuronal cells: how to make and define a neuron. *Cell stem cell* **9,** 517–25. ISSN: 1875-9777 (2011).

320. JoVE. Primary Neuronal Cultures. *JoVE Science Education Database. Neuroscience.* (2019).

321. Xie, Y., Schutte, R. J., Ng, N. N., Ess, K. C., Schwartz, P. H. & O'Dowd, D. K. Reproducible and efficient generation of functionally active neurons from human hiPSCs for preclinical disease modeling. *Stem cell research* **26,** 84–94. ISSN: 1876-7753 (2018).

322. Tremblay, R. G., Sikorska, M., Sandhu, J. K., Lanthier, P., Ribecco-Lutkiewicz, M. & Bani-Yaghoub, M. Differentiation of mouse Neuro 2A cells into dopamine neurons. *Journal of Neuroscience Methods* **186,** 60–67. ISSN: 0165-0270 (2010).

323. Kumar, M. & Katyal, A. Data on retinoic acid and reduced serum concentration induced differentiation of Neuro-2a neuroblastoma cells. *Data in Brief* **21,** 2435–2440. ISSN: 2352-3409 (2018).

324. Wasilewska-Sampaio, A. P., Silveira, M. S., Holub, O., Goecking, R., Gomes, F. C. A., Neto, V. M., Linden, R., Ferreira, S. T. & De Felice, F. G. Neuritogenesis and neuronal differentiation promoted by 2,4-dinitrophenol, a novel anti-amyloidogenic compound. *The FASEB Journal* **19,** 1627–1636. ISSN: 0892-6638 (2005).

325. Lee, M. K. & Nikodem, V. M. Differential role of ERK in cAMP-induced Nurr1 expression in N2A and C6 cells. *Neuroreport* **15,** 99–102. ISSN: 0959-4965 (2004).

326. Evangelopoulos, M. E., Weis, J. & Krüttgen, A. Signalling pathways leading to neuroblastoma differentiation after serum withdrawal: HDL blocks neuroblastoma differentiation by inhibition of EGFR. *Oncogene* **24,** 3309–3318. ISSN: 0950-9232 (2005).

327. Mao, A. J., Bechberger, J, Lidington, D, Galipeau, J, Laird, D. W. & Naus, C. C. Neuronal differentiation and growth control of neuro-2a cells after retroviral gene delivery of connexin43. *The Journal of biological chemistry* **275,** 34407–14. ISSN: 0021-9258 (2000).

328. Tropepe, V., Sibilia, M., Ciruna, B. G., Rossant, J., Wagner, E. F. & van der Kooy, D. Distinct Neural Stem Cells Proliferate in Response to EGF and FGF in the Developing Mouse Telencephalon. *Developmental Biology* **208,** 166–188. ISSN: 00121606 (1999).

329. Bani-Yaghoub, M., Tremblay, R., Voicu, R., Mealing, G., Monette, R., Py, C., Faid, K. & Sikorska, M. Neurogenesis and neuronal communication on micropatterned neurochips. *Biotechnology and Bioengineering* **92,** 336–345. ISSN: 0006-3592 (2005).

330. Xu, X., Sun, Y., Holmes, T. C. & López, A. J. Noncanonical connections between the subiculum and hippocampal CA1. *The Journal of comparative neurology* **524,** 3666–3673. ISSN: 1096-9861 (2016).

331. Jullien, N., Sampieri, F., Enjalbert, A. & Herman, J. Regulation of Cre recombinase by ligand-induced complementation of inactive fragments. *Nucleic Acids Research* **31,** 131e–131. ISSN: 1362-4962 (2003).

332. Burmakina, S., Geng, Y., Chen, Y. & Fan, Q. R. Heterodimeric coiled-coil interactions of human GABAB receptor. *Proceedings of the National Academy of Sciences of the United States of America* **111,** 6958–63. ISSN: 1091-6490 (2014).

333. Hodel, A. E., Harreman, M. T., Pulliam, K. F., Harben, M. E., Holmes, J. S., Hodel, M. R., Berland, K. M. & Corbett, A. H. Nuclear localization signal receptor affinity correlates with in vivo localization in Saccharomyces cerevisiae. *The Journal of biological chemistry* **281,** 23545–56. ISSN: 0021-9258 (2006).

334. Jérôme, V & Müller, R. A synthetic leucine zipper-based dimerization system for combining multiple promoter specificities. *Gene Therapy* **8,** 725–729. ISSN: 0969-7128 (2001).

335. Van Duyne, G. D. A Structural View of Cre-loxP Site-Specific Recombination. *Annual Review of Biophysics and Biomolecular Structure* **30,** 87–104 (2001).

336. Mibayashi, M., Martínez-Sobrido, L., Loo, Y.-M., Cárdenas, W. B., Gale, M., García-Sastre, A. & García-Sastre, A. Inhibition of retinoic acid-inducible gene I-mediated induction of beta interferon by the NS1 protein of influenza A virus. *Journal of virology* **81,** 514–24. ISSN: 0022-538X (2007).

337. Sternson, S. M. & Eiselt, A.-K. Three Pillars for the Neural Control of Appetite. *Annual Review of Physiology* **79,** 401–423. ISSN: 0066-4278 (2017).

338. Shi, Y.-C. *et al.* Arcuate NPY Controls Sympathetic Output and BAT Function via a Relay of Tyrosine Hydroxylase Neurons in the PVN. *Cell Metabolism* **17,** 236–248. ISSN: 15504131 (2013).

339. Oldfield, B. J., Giles, M. E., Watson, A, Anderson, C, Colvill, L. M. & McKinley, M. J. The neurochemical characterisation of hypothalamic pathways projecting polysynaptically to brown adipose tissue in the rat. *Neuroscience* **110,** 515–26. ISSN: 0306-4522 (2002).

340. Stefanidis, A., Wiedmann, N., Adler, E. & Oldfield, B. Hypothalamic control of adipose tissue. *Best Practice & Research Clinical Endocrinology & Metabolism* **28,** 685–701 (2014).

Appendix A

Sequencing data for individual SiR viral particles *in vitro*

Batch A					
	Frequency	Sequence	Position	Mutation	Effect on CDS
Leader sequence	1/50	GAT > GAC	-54	Substitution	-
	1/50	AAA > AAG	-18	Substitution	-
N gene	1/50	GCC > GCT	+186	Substitution	Synonymous A62
	1/50	TTT > TTTT	+243	Insertion	Frameshift
	1/50	AAG > A-G	+485	Deletion	Frameshift
	1/50	ATG > CTG	+562	Substitution	M188L
	1/50	GTG > G--	+677/8	Deletion	Frameshift
	1/50	ACG > ACCG	+983	Insertion	Frameshift
	1/50	GAA > AAA	+1093	Substitution	E365K
	1/50	TCA > CCA	+1276	Substitution	S426P
TEVs-PEST	-	-	-	-	-
Intergenic N/P	4/50	AAA > AAAA	+1571	Insertion	-
	1/50	CCC > CCA	+1581	Substitution	-
P gene	-	-	-	-	-
Batch B					
	Frequency	Sequence	Position	Mutation	Effect on CDS
Leader sequence	1/50	AAC > A-C	-63	Deletion	-
	1/50	CAA > CA-	-60	Deletion	
	1/50	CTA > CTG	-3	Substitution	-
N gene	1/50	TTT > TTTT	+243	Insertion	Frameshift
	1/50	GAC > GAA	+501	Substitution	D167E
	1/50	AAT > AAC	+588	Substitution	Synonymous N196
	1/50	GCT > GCC	+1002	Substitution	Synonymous A334
	1/50	AAA > AAAA	+1056	Insertion	Frameshift
TEVs-PEST	1/50	TCC > TGC	+1385	Substitution	S462C in GSG linker after TEVs
Intergenic N/P	1/50	TAT > TAA	+1554	Substitution	-
	2/50	AAA > AAAA	+1571	Insertion	-
P gene	1/50	GAA > GAG	+1671	Substitution	Synonymous E23
	1/50	CTG > CCG	+1775	Substitution	L58P
	1/50	GGA > TGA	+2014	Deletion	Nonsense G138>STOP

Table A.1 Sequences of individual SiR B19 particles Continued ...

Batch C					
	Frequency	Sequence	Position	Mutation	Effect on CDS
Leader sequence	2/50	AAA > AAAA	-43	Insertion	-
N gene	1/50	TGT > TTT	+212	Substitution	C71F
	1/50	AGA > AGG	+1074	Substitution	Synonymous R358
	1/50	GGT > GAT	+1190	Substitution	G397D
TEVs-PEST	-	-	-	-	-
Intergenic N/P	1/50	AAA > AAG	+1569	Substitution	-
	3/50	AAA > AAAA	+1571	Insertion	-
	1/50	AAA > AA-	+1571	Deletion	
P gene	1/50	CAA > AAA	+1720	Substitution	Q40K
Batch D					
	Frequency	Sequence	Position	Mutation	Effect on CDS
Leader sequence	-	-	-	-	-
N gene	1/50	AAG > AGG	+113	Substitution	K38R
	1/50	AAA > CAA	+295	Substitution	K99Q
	1/50	CAT > AAT	+655	Substitution	H219N
	1/50	TCA > TCC	+873	Substitution	Synonymous S291
	1/50	ACC > AAC	+1196	Substitution	T399N
TEVs-PEST	-	-	-	-	-
Intergenic N/P	3/50	AAA > AAAA	+1571	Insertion	-
	1/50	ATC > ATT	+1596	Substitution	-
P gene	1/50	AAA > AAAA	+1671	Insertion	Frameshift
	1/50	CGT > CTA	+1878	Substitution	Synonymous L92
	1/50	AGA > AGT	+1941	Substitution	R113S
	1/50	GGA > GGG	+2016	Substitution	Synonymous G138
	1/50	ACT > ACA	+2046	Substitution	Synonymous T148
Batch E					
	Frequency	Sequence	Position	Mutation	Effect on CDS
Leader sequence	1/50	CCA > CC-	-57	Deletion	-
N gene	1/50	CCT >CAT	+200	Substitution	P67H
	1/50	TTT > TTTT	+243	Insertion	Frameshift
	1/50	GGA > GAA	+371	Substitution	G124E
	1/50	ACA > ACG	+387	Substitution	Synonymous T129
	2/50	GAC > GAT	+393	Substitution	Synonymous D131
	1/50	CAC > C--	+551/2	Deletion	Frameshift
	1/50	ACT > AAT	+557	Substitution	T186N
	1/50	TTT > TTTT	+779	Insertion	Frameshift
TEVs-PEST	-	-	-	-	-
Intergenic N/P	1/50	CAT > CAC	+1560	Substitution	-
	1/50	AAA > AAC	+1570	Substitution	
	4/50	AAA > AAAA	+1571	Insertion	
	1/50	ATC > ATT	+1596	Substitution	-
P gene	1/50	GAA > GGA	+1667	Substitution	E22G

Table A.1 Sequences of individual SiR B19 particles Continued ...

Batch F					
	Frequency	Sequence	Position	Mutation	Effect on CDS
Leader sequence	1/50	ACC > AC-	-58	Deletion	-
	1/50	CAG > CA-	-56	Deletion	-
	1/50	TCA > TCG	-52	Substitution	-
	1/50	AAA > AAAA	-43	Insertion	-
	1/50	AAG > AA-	-22	Deletion	-
N gene	1/50	TTT >TTTTT	+243/4	Insertion	Frameshift
	1/50	TTG > TCG	+434	Substitution	L145S
	1/50	TTT > TT-	+534	Deletion	Frameshift
	1/50	GCA > GTA	+767	Substitution	A256V
	1/50	ACA > ATA	+836	Substitution	T279I
	1/50	AAA > AAAA	+908	Insertion	Frameshift
	1/50	321 bp	+1041 - 1362	Deletion	Deletion of C-terminal of N, upstream (E347) is in frame with PEST domain
	1/50	GGA > GGG	+1038	Substitution	Synonymous G346
TEVs-PEST	-	-	-	-	-
Intergenic N/P	4/50	AAA > AAAA	+1571	Insertion	-
P gene	1/50	CCT > CCC	+1626	Substitution	Synonymous P8
	1/50	GAA > GGA	+1727	Substitution	E42G
	1/50	TTT > TTC	+1845	Substitution	Synonymous F81
Batch G					
	Frequency	Sequence	Position	Mutation	Effect on CDS
Leader sequence	1/50	CCA > CC-	-57	Deletion	-
	1/50	AAA > AA-	-16	Deletion	-
N gene	1/50	GCA > GTA	+290	Substitution	A97V
	1/50	CAT > GAT	+409	Substitution	H137D
	1/50	TTT > TT-	+534	Deletion	Frameshift
	1/50	TAT > TGT	+1271	Substitution	Y424C
	1/50	GCC > GTC	+1316	Substitution	A439V
TEVs-PEST	-	-	-	-	-
Intergenic N/P	4/50	AAA > AAAA	+1571	Insertion	-
P gene	1/50	AAA > CAA	+1786	Substitution	K62Q
	1/50	GAA > GGA	+1823	Substitution	E74G
	1/50	CGA > CAA	+1834	Substitution	R78Q

Table A.1 Sequences of individual SiR B19 particles Continued ...

Batch H					
	Frequency	Sequence	Position	Mutation	Effect on CDS
Leader sequence	1/50	AAA > AAAA	-43	Insertion	-
	1/50	AAC > AA-	-42	Deletion	
N gene	1/50	TTA > CTA	+145	Substitution	Synonymous L49
	1/50	ATG > ATA	+234	Substitution	M78I
	1/50	TTT > TTTT	+243	Insertion	Frameshift
	1/50	AAA > CAA	+295	Substitution	K99Q
	1/50	GAT > AAT	+301	Substitution	D101N
	1/50	GGA > AGA	+622	Substitution	G208R
	1/50	GCT > TCT	+838	Substitution	A280S
	1/50	GGC > G-C	+1028	Deletion	Frameshift
	1/50	GAC > AAC	+1132	Substitution	D378N
TEVs-PEST	1/50	CTG > CTA	+1437	Substitution	Synonymous L16 in PEST domain
Intergenic N/P	3/50	AAA > AAAA	+1571	Insertion	-
	1/50	AAC > AAA	+1592	Substitution	-
P gene	1/50	AAA > AAAA	+1788	Insertion	Frameshift

Table A.1 Sequences of individual SiR B19 particles
List of detected mutations divided by batch (50 individual clones per batch). The position of the mutations is calculated referring to +1 as the first base of the nucleoprotein N coding sequence.

Batch A					
	Frequency	Sequence	Position	Mutation	Effect on CDS
N gene	1/47	TTG > -- G	+616/7	Deletion	Frameshift
	1/47	ATC > AAC	+1226	Substitution	I409N
	2/47		-48 to +1638	Deletion	Block Deletion of N and P gene
TEVs-PEST	-	-	-	-	-
Intergenic N/P	2/47	AAA > AAAA	+1571	Insertion	-
P gene	-	-	-	-	-

Batch B					
	Frequency	Sequence	Position	Mutation	Effect on CDS
N gene	1/47	AAG > GAG	+13	Substitution	K5E
	1/47	AAC > GAC	+1307	Substitution	N436S
	1/47		-48 to +1638	Deletion	Block Deletion of N and P gene
TEVs-PEST	1/47	GGT > TGT	+1387	Substitution	G463C in GSG linker after TEVs
Intergenic N/P	2/47	AAA > AAAA	+1571	Insertion	-
P gene	-	-	-	-	-

Table A.2 Sequences of individual SiR N2C particles
List of detected mutations divided by batch (47 individual clones per batch). The position of the mutations is calculated referring to +1 as the first base of the nucleoprotein N coding sequence.

Appendix B

Sequencing data for SiR viral populations *in vivo*

	Position	Reference	Mutation	Mutation Type	Frequency (out of 3 samples)	Total Read Depth
Leader Sequence	30	GACAGCGTCAATTGCAAAGCA	AGACAGCGTCAATTGCAAAGC	complex	2	20
	55	TGTAACACCCCTACAATGGATGCCGACA AGATTGTGTTCAAAGTCAATAATCAGGT GGTCTCTTTGAAGCCTGAGATTATCGTG GATCAATATGAGTACAAG	ATGTAACACCCCTACAATGGATGCCGACA AGATTGTGTTCAAAGTCAATAATCAGGTG GTCTCTTTGAAGCCTGAGATTATCGTGGA TCAATATGAGTACAA	complex	1	5
	55	TGTAACACCCCTACAATGGATGCCGACA AGATTGTGTTCAAAGTCAATAATCAGGT GGTCTCTTTGAAGCCTGAGATTATCG	ATGTAACACCCCTACAATGGATGCCGACA AGATTGTGTTCAAAGTCAATAATCAGGTG GTCTCTTTGAAGCCTGAGATTATC	complex	1	5
N	183	GCCTTGTATCACCCTAGGGAAAGC	AGCCTTGTATCACCCTAGGGAAAG	complex	3	31
	211	GACTTGAACAAAGCATACAAATCAGTTT TATCAGGCATGAATGCCGCCAAACTTGA TCCGGATGATGTATGCTCCTACTTGGCA GCAGCAATGCAGTTCTTTGAG	CGACTTGAACAAAGCATACAAATCAGTTT TATCAGGCATGAATGCCGCCAAACTTGAT CCGGATGATGTATGCTCCTACTTGGCAGC AGCAATGCAGTTCTTTGA	complex	2	22
	211	GACTTGAACAAAGCATACAAATCAGTTT TATCAGGCATGAATGCCGCCAAACTTGA TCCGGATGATGTATGCTCCTACTTGGCA GCAGCAATGCAGTTCTTTGAGGGGACA TGTCCGGAAGACTGGAC	CGACTTGAACAAAGCATACAAATCAGTTT TATCAGGCATGAATGCCGCCAAACTTGAT CCGGATGATGTATGCTCCTACTTGGCAGC AGCAATGCAGTTCTTTGAGGGGACATGT CCGGAAGACTGGA	complex	1	5
	340	AGCTATGGAATCCTGATTGCACGAAAAG GAGATAGGATCACCCCAAACTCTCTAGT GGAGATAAAGCGTACTGATGTAGAAGG GAATTGGGCTCTGACAGGAGGCATGGA ATTGACAAG	CAGCTATGGAATCCTGATTGCACGAAAAG GAGATAGGATCACCCCAAACTCTCTAGTG GAGATAAAGCGTACTGATGTAGAAGGGAA TTGGGCTCTGACAGGAGGCATGGAATTG ACAA	complex	1	7
	460	GACCCCACTGTCTCTGAACATGCATCTT TAGTCGGTCTTCTCCTGAGTCTGTACAG GTTGAGCAAAATATCAGGACAGAACACT GGTAACTATAAGACAAACATTGCAGATA GGATAGAGCAGATTTTCGAGA	GGACCCCACTGTCTCTGAACATGCATCTT TAGTCGGTCTTCTCCTGAGTCTGTACAGG TTGAGCAAAATATCAGGACAGAACACTGG TAACTATAAGACAAACATTGCAGATAGGAT AGAGCAGATTTTCGAG	complex	1	6
	466	A	C	snp	2	150
	467	CTGTCTCTGAACATGCATCTTTAGTCGG TCTTCTCCTGAGTCTGTACAGGTTGAGC AAAATATCAGGACAGAACACTGGTAACT ATAAGACAAACATTGCAGATAGGATAGA GCA	ACTGTCTCTGAACATGCATCTTTAGTCGG TCTTCTCCTGAGTCTGTACAGGTTGAGCA AAATATCAGGACAGAACACTGGTAACTATA AGACAAACATTGCAGATAGGATAGAGC	complex	2	34
	582	GATTTTCGAGA	AGATTTTCGAG	complex	2	120
	593	CAGCACCTTTTGTTAAGATCGTGGAACA CCATACCCTAATGACAACTCACAAGATG TGTGCTAATTGGAGTACTATACCGAACT TCAGAT	ACAGCACCTTTTGTTAAGATCGTGGAACA CCATACCCTAATGACAACTCACAAGATGT GTGCTAATTGGAGTACTATACCGAACTTC AGA	complex	2	50
	593	C	A	snp	1	15
	687	GGCCGGAACCTACGACATGT	TGGCCGGAACCTACGACATG	complex	3	117
	711	CTCACGGATTGAGCATCTATATTCGGCA ATCAGAGTGGGCACAGTCGTCACCGCT TATGAAGACTGCTCAGGACTGGTATCGT TTACAGGGTTCATAAAGCAGATCAATCT CACCGCAAGGGAAGCA	TCTCACGGATTGAGCATCTATATTCGGCA ATCAGAGTGGGCACAGTCGTCACCGCTT ATGAAGACTGCTCAGGACTGGTATCGTTT ACAGGGTTCATAAAGCAGATCAATCTCAC CGCAAGGGAAGC	complex	1	5
	839	TACTATATTTCTTCCACAAGAACTTTGAG GAAGAGATAAGAAGAATGTTCGAGCCA GGGCAAGAGACAGCTGTTCCTCACTCT TATTTCATCCACTTCCGTTCACT	ATACTATATTTCTTCCACAAGAACTTTGAG GAAGAGATAAGAAGAATGTTCGAGCCAG GGCAAGAGACAGCTGTTCCTCACTCTTAT TTCATCCACTTCCGTTCAC	complex	2	24
	869	A	G	snp	1	16
	870	AGAGATAAGAAGAATGTTCGAGCCAGG GCAAGAGACAGCTGTTCCTCACTCTTAT TTCATCCACT	AAGAGATAAGAAGAATGTTCGAGCCAGG GCAAGAGACAGCTGTTCCTCACTCTTATT TCATCCAC	complex	1	6
	1067	GTGCCCCTCATGAGATGTCTGTTCTAG	TGTGCCCCTCATGAGATGTCTGTTCTA	complex	2	132
	1227	CGTCAACTCTGATGACGAGGACTATTTC TCTGGTGAAACCAGAAGTCCAGAAGCT GTCTATACTCGAATCATGATGAATGGAG GTCGACTGAAGA	CCGTCAACTCTGATGACGAGGACTATTTC TCTGGTGAAACCAGAAGTCCAGAAGCTG TCTATACTCGAATCATGATGAATGGAGGT CGACTGAAG	complex	2	24
	1232	ACTCTGATGACGAGGACTATTTCTCTGG TGAAACCAGAAGTCCAGAAGCTGTCTAT ACTCGAATCATGATGAATGGAGGTCGAC TGAAGA	AACTCTGATGACGAGGACTATTTCTCTGG TGAAACCAGAAGTCCAGAAGCTGTCTATA CTCGAATCATGATGAATGGAGGTCGACTG AAG	complex	1	8
	1322	G	A	snp	1	22
	1323	ATCTCATATACGGAGATATGTCTCAGTCA GTTCCAATCATCAAGCCCGTCCAAACTC ATTCGCCGAAT	GATCTCATATACGGAGATATGTCTCAGTCA GTTCCAATCATCAAGCCCGTCCAAACTCA TTCGCCGAA	complex	1	4
	1395	AAACAAGACGTATTCGAATGACTCAGGT TCCGGAGAGAACCTCTACTTCCAATCG GGATCCGGTAGCCATGGCTTCCCGCCG	TAAACAAGACGTATTCGAATGACTCAGGT TCCTGAGAGAACCTCTACTTCCAATCGGG ATCCGGTAGCCATGGCTTCCCGCC	complex	2	28
TEVs-PEST		no mutations identified (1421-1582)				
N/P intergenic	1583	GGAGTTGATTGACAGGGTGCCAGAAAT CTATAGATTGTATATATCCATCATGA	AGGAGTTGATTGACAGGGTGCCAGAAAT CTATAGATTGTATATATCCATCATG	complex	2	42
	1641	CTAACACTCCTCCTTTCAAACCATCCCA AATATGAGCAAGATCTTTGTTAATCCGA GTGCAATCAGAGCCGGTCTGGCCGATC TTGAGATGGCCGA	ACTAACACTCCTCCTTTCAAACCAGCCCA AATATGAGCAAGATCTTTGTTAATCCGAGT GCAATCAGAGCCGGTCTGGCCGATCTTG AGATGGCCG	complex	1	8
	1641	CT	AC	mnp	1	5
	1643	A	T	snp	1	3
	1645	C	A	snp	1	3
	1646	A	C	snp	1	3

Table B.1 SiR N2C mutations detected by NGS 1 week post injection *in vivo* Continued ...

	1651	T	C	snp	1	3
	1652	C	T	snp	1	3
	1654	T	C	snp	1	3
	1657	C	T	snp	1	3
P	1737	AGA	AAG	mnp	2	102
	1911	C	A	snp	1	14
	2236	ACCAATGAAGAAGATGATCTATCAGTAG AGGCTGAGATCGCTCATCAGATTGCTG AAAGCTTTTCCAAGAAGTACAAGTTTC	CACCAATGAAGAAGATGATCTATCAGTAG AGGCTGAGATCGCTCATCAGATTGCTGAA AGCTTTTCCAAGAAGTACAAGTTT	complex	2	44
	2318	CCTCCCGATCTTCAGGAATATTCTTGTAT AATTTTGAGCAACTGAAGATGAACCTTG ATG	CCCTCCCGATCTTCAGGAATATTCTTGTAT AATTTTGAGCAACTGAAGATGAACCTTGA T	complex	2	42
	2378	ACATAGTTAAAGAGGCA	GACATAGTTAAAGAGGC	complex	2	44
	2400	TGTACCGGGCGTGACCCGTCTGGCCCA TGATGGATCCAAAATC	ATGCACCGGGCGTGACCCGTCTGGCCCA TGATGGATCCAAAAT,ATGTACCGGGCGT GACCCGTCTGGCCCATGATGGATCCAAA AT	complex	2	34
	2400	TGTACCGGGCGTGACCCGTCTGGCCCA TGATGGATCCAAAATC	ATGTACCGGGCGTGACCCGTCTGGCCCA TGATGGATCCAAAAT	complex	1	14
P/M intergenic	2609	T	A	snp	1	11
	2610	GTTATGGAGTAGACATGA	TGTTATGGAGTAGACATG	complex	1	7
	2620	A	T	snp	2	28
	2621	GACATGA	AGACATG	mnp	2	20
M	2634	CAGGCAACACCACTGATAAAATGAACGT TCTACGCAAGATAGTGA	ACAGGCAACACCACTGATAAAATGAACGT TCTACGCAAGATAGTG	complex	3	22
	3182	GCATGTCAACTATGGTCTGACATGTCTC TTCAGACACAAAGGTCTGAAGAGGACA AAGACTCTTCTCTGCTTCTAGAATAATCA GATTATATCCTGCAAGTGTATCACTTGTT TACCTCTGGAGGAGAGAGCATACAGGC TTGACTCCGA	GGCATGTCAACTATGGTCTGACATGTCTC TTCAGACACAAAGGTCTGAAGAGGACAA AGACTCTTCTCTGCTTCTAGAATAATCAGA TTATATCCTGCAAGTGTATCACTTGTTTAC CTCTGGAGGAGAGAGCATACAGGCTTGA CTCCG	complex	1	5
	3183	C	G	snp	1	16
M/CRE mCHERRY intergenic	3340	A	G	snp	2	58
	3341	GCAATAGAACA	AGCAATAGAAC	complex	2	52
	3358	CACACGTTATGGTGCCGTTAAATCGCTG CATTTTATCAAAGTCAAGTTGATAACCTT TACATTTTGA	ACACACGTTATGGTGCCGTTAAATCGCTG CATTTTATCAAAGTCAAGTTGATAACCTTT ACATTTTG	complex	1	6
	3358	CACACGTTATGGTGCCGTTAAATCGCTG CATTTTATCAAAGTCAAGT	ACACACGTTATGGTGCCGTTAAATCGCTG CATTTTATCAAAGTCAAG	complex	1	4
	3425	G	A	snp	1	7
	3426	CCTCTTGGATGTGA	GCCTCTTGGATGTG	complex	1	5
CRE mCHERRY	3498	A	G	snp	1	3
	3728	CCTCCTGTACCTGCAAGCCAGAGGCCT GGCTGTGAAGACCATCCAACAGCACCT GGGCCAGCTCAACATGCTGCACAGGAG ATCTGGC	ACCTCCTGTACCTGCAAGCCAGAGGCCT GGCTGTGAAGACCATCCAACAGCACCTG GGCCAGCTCAACATGCTGCACAGGAGAT CTGG	complex	1	5
	4057	ACATCTCCCGCACCGATGGTGGGAGAA TGCTGATCCACATTGGCAGGACCAAGA CCCTGGTGTCCACAGCTGGTGTGGAGA AGGCCCTGTCCCTG	GACATCTCCCGCACCGATGGTGGGAGAA TGCTGATCCACATTGGCAGGACCAAGAC CCTGGTGTCCACAGCTGGTGTGGAGAAG GCCCTGTCCCT	complex	2	34
	4720	GCGAGGGCGAGGGCCGCCCCTACGAG GGCACCCAGACCGCCAAGCTGAAGGT GACCAAGGGTGGC	GGCGAGGGCGAGGGCCGCCCCTACGAG GGCACCCAGACCGCCAAGCTGAAGGTG ACCAAGGGTGG	complex	1	5
	4789	T	C	snp	1	6
	4790	G	T	snp	1	6
	4791	C	G	snp	1	6
	4794	T	C	snp	1	6
	4796	C	T	snp	1	6
	4797	G	C	snp	1	6
	4798	C	G	snp	1	6
	4945	C	A	snp	1	13
	5187	G	C	snp	1	4
	5188	T	G	snp	1	3
mCHERRY /L intergenic		no mutations identified (5455-5984)				
L	7282	C	A	snp	1	4
	8407	A	G	snp	1	4
	8889	T	A	snp	1	16
	8984	G	A	snp	1	13
	9132	C	T	snp	2	70
	10483	A	G	snp	1	19
	10966	C	T	snp	1	11
	10967	T	C	snp	1	11
	10968	C	T	snp	1	11
	10969	T	C	snp	1	11
	10974	G	A	snp	1	11
	10994	T	C	snp	1	10
	10995	C	T	snp	1	10
	11079	T	C	snp	2	110
	11127	T	C	snp	2	178
	11128	C	T	snp	2	175

Table B.1 SiR N2C mutations detected by NGS 1 week post injection *in vivo* Continued ...

	11132	A	C	snp	2	178
	11135	T	G	snp	2	178
	11142	A	G	snp	2	178
	11146	T	C	snp	2	178
Trailer Sequence		no mutations identified (12369-12499)				

Table B.1 SiR N2C mutations detected by NGS 1 week post injection *in vivo* Mice were injected in CA1 with SiR N2C and hippocampal total RNA extracted 1 week p.i. Sequenced reads from NGS of extracted RNA were aligned to the SiR N2C reference genome and a list of mutations identified using freeBayes software. Table lists the genome position, genomic change, type of mutation (snp = single-nucleotide polymorphism, indel = insertions and deletions, mnp = multi-nucleotide polymorphism, complex = composite insertion and substitution events), frequency of mutation identification from triplicate samples and total read depth.

Appendix C

Gene list of upregulated transcripts following SiR N2C injection *in vivo*

Gene Name	Gene ID	Gene Description	Log2 Fold Change	P-value
Gm4951	ENSMUSG00000073555	predicted gene 4951 [Source:MGI Symbol;Acc:MGI:3644953]	6.78	7.44E-06
Iigp1	ENSMUSG00000054072	interferon inducible GTPase 1 [Source:MGI Symbol;Acc:MGI:1926259]	6.72	5.30E-05
Mx1	ENSMUSG00000000386	MX dynamin-like GTPase 1 [Source:MGI Symbol;Acc:MGI:97243]	6.07	8.75E-07
H2-Q6	ENSMUSG00000073409	histocompatibility 2, Q region locus 6 [Source:MGI Symbol;Acc:MGI:95935]	6.02	1.78E-06
Cxcl10	ENSMUSG00000034855	chemokine (C-X-C motif) ligand 10 [Source:MGI Symbol;Acc:MGI:1352450]	5.84	1.50E-06
Zbp1	ENSMUSG00000027514	Z-DNA binding protein 1 [Source:MGI Symbol;Acc:MGI:1927449]	5.76	8.85E-06
Ifi44	ENSMUSG00000028037	interferon-induced protein 44 [Source:MGI Symbol;Acc:MGI:2443016]	5.65	2.03E-06
Oasl2	ENSMUSG00000029561	2'-5' oligoadenylate synthetase-like 2 [Source:MGI Symbol;Acc:MGI:1344390]	5.45	8.15E-05
Ifi209	ENSMUSG00000043263	interferon activated gene 209 [Source:MGI Symbol;Acc:MGI:2138243]	5.40	6.63E-05
F830016 B08Rik	ENSMUSG00000090942	RIKEN cDNA F830016B08 gene [Source:MGI Symbol;Acc:MGI:3588218]	5.36	6.25E-05
Igtp		No annotated features. Position chromosome 11, 58199556-58222782	5.33	8.37E-05
Epsti1	ENSMUSG00000022014	epithelial stromal interaction 1 (breast) [Source:MGI Symbol;Acc:MGI:1915168]	5.28	1.70E-06
Irf7	ENSMUSG00000025498	interferon regulatory factor 7 [Source:MGI Symbol;Acc:MGI:1859212]	5.25	9.04E-08
H2-K1	ENSMUSG00000061232	histocompatibility 2, K1, K region [Source:MGI Symbol;Acc:MGI:95904]	5.14	1.78E-05
Ciita	ENSMUSG00000022504	class II transactivator [Source:MGI Symbol;Acc:MGI:108445]	5.13	3.72E-05
Atp8b4	ENSMUSG00000060131	ATPase, class I, type 8B, member 4 [Source:MGI Symbol;Acc:MGI:1859664]	5.12	9.68E-05
Psmb8	ENSMUSG00000024338	proteasome (prosome, macropain) subunit, beta type 8 (large multifunctional peptidase 7) [Source:MGI Symbol;Acc:MGI:1346527]	5.11	4.59E-06
Ifi213	ENSMUSG00000073491	interferon activated gene 213 [Source:MGI Symbol;Acc:MGI:3695276]	5.11	4.63E-06
Usp18	ENSMUSG00000030107	ubiquitin specific peptidase 18 [Source:MGI Symbol;Acc:MGI:1344364]	4.99	1.61E-06
H2-Q7	ENSMUSG00000060550	histocompatibility 2, Q region locus 7 [Source:MGI Symbol;Acc:MGI:95936]	4.97	4.65E-05
Ifit1	ENSMUSG00000034459	interferon-induced protein with tetratricopeptide repeats 1 [Source:MGI Symbol;Acc:MGI:99450]	4.92	8.90E-10

Table C.1 Table of upregulated transcripts following *in vivo* injection of SiR N2C Continued ...

Gene Name	Gene ID	Gene Description	Log2 Fold Change	P-value
Irgm1	ENSMUSG00000046879	immunity-related GTPase family M member 1 [Source:MGI Symbol;Acc:MGI:107567]	4.92	3.34E-05
Bst2	ENSMUSG00000046718	bone marrow stromal cell antigen 2 [Source:MGI Symbol;Acc:MGI:1916800]	4.67	9.93E-06
Rnf213	ENSMUSG00000070327	ring finger protein 213 [Source:MGI Symbol;Acc:MGI:1289196]	4.67	7.37E-08
Oas2	ENSMUSG00000032690	2'-5' oligoadenylate synthetase 2 [Source:MGI Symbol;Acc:MGI:2180852]	4.64	8.42E-05
Stat4			4.32	1.72E-06
Apobec1	ENSMUSG00000040613	apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1 [Source:MGI Symbol;Acc:MGI:103298]	4.29	9.27E-05
Mx2	ENSMUSG00000023341	MX dynamin-like GTPase 2 [Source:MGI Symbol;Acc:MGI:97244]	4.10	2.97E-05
Samd9l	ENSMUSG00000047735	sterile alpha motif domain containing 9-like [Source:MGI Symbol;Acc:MGI:1343184]	4.01	1.55E-05
Ddx60	ENSMUSG00000037921	DEAD (Asp-Glu-Ala-Asp) box polypeptide 60 [Source:MGI Symbol;Acc:MGI:2384570]	3.91	1.20E-07
AU020206	ENSMUSG00000097415	expressed sequence AU020206 [Source:MGI Symbol;Acc:MGI:2142134]	3.91	7.62E-06
Il10ra	ENSMUSG00000032089	interleukin 10 receptor, alpha [Source:MGI Symbol;Acc:MGI:96538]	3.85	1.05E-04
Parp9	ENSMUSG00000022906	poly (ADP-ribose) polymerase family, member 9 [Source:MGI Symbol;Acc:MGI:1933117]	3.77	4.53E-05
Xaf1	ENSMUSG00000040483	XIAP associated factor 1 [Source:MGI Symbol;Acc:MGI:3772572]	3.72	1.15E-06
Mpeg1	ENSMUSG00000046805	macrophage expressed gene 1 [Source:MGI Symbol;Acc:MGI:1333743]	3.66	3.34E-06
Ifit3	ENSMUSG00000074896	interferon-induced protein with tetratricopeptide repeats 3 [Source:MGI Symbol;Acc:MGI:1101055]	3.50	5.29E-07
Zc3hav1	ENSMUSG00000029826	zinc finger CCCH type, antiviral 1 [Source:MGI Symbol;Acc:MGI:1926031]	3.48	9.61E-05
ddx	ENSMUSG00000060802	beta-2 microglobulin [Source:MGI Symbol;Acc:MGI:88127]	3.36	2.78E-06
Eif2ak2	ENSMUSG00000024079	eukaryotic translation initiation factor 2- alpha kinase 2 [Source:MGI Symbol;Acc:MGI:1353449]	3.18	3.10E-06
Lgals3bp	ENSMUSG00000033880	lectin, galactoside-binding, soluble, 3 binding protein [Source:MGI Symbol;Acc:MGI:99554]	2.99	2.00E-06
Pik3ap1	ENSMUSG00000025017	phosphoinositide-3-kinase adaptor protein 1 [Source:MGI Symbol;Acc:MGI:1933177]	2.98	1.22E-04

Table C.1 Table of upregulated transcripts following *in vivo* injection of SiR N2C Continued ...

Gene Name	Gene ID	Gene Description	Log2 Fold Change	P-value
C1qa	ENSMUSG00000036887	complement component 1, q subcomponent, alpha polypeptide [Source:MGI Symbol;Acc:MGI:88223]	2.65	1.34E-05
Inpp5d	ENSMUSG00000026288	inositol polyphosphate-5-phosphatase D [Source:MGI Symbol;Acc:MGI:107357]	2.59	7.77E-05
Stat2	ENSMUSG00000040033	signal transducer and activator of transcription 2 [Source:MGI Symbol;Acc:MGI:103039]	2.49	4.32E-06
Vim	ENSMUSG00000026728	vimentin [Source:MGI Symbol;Acc:MGI:98932]	2.37	9.53E-05
Grn	ENSMUSG00000034708	granulin [Source:MGI Symbol;Acc:MGI:95832]	2.16	9.70E-05
Etv6	ENSMUSG00000030199	ets variant 6 [Source:MGI Symbol;Acc:MGI:109336]	1.28	3.05E-05

Table C.1 Table of upregulated transcripts following *in vivo* injection of SiR N2C List of upregulated transcripts following SiR N2C infection in CA1 region of the hippocampus *in vivo* as identified by DESeq2 analysis. Genes listed in order of fold change differences in transcript levels compared to PBS injected control mice.