UNIVERSITY OF CALIFORNIA SAN DIEGO

Improving Viral Tools for Neural Circuit Dissection: Monosynaptic Anterograde Tracing and Recombinase-Dependent AAV Expression of Sensitive Transgenes

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Neurosciences
with a Specialization in Computational Neurosciences

by

Kyle Brady Fischer

Committee in charge:

Professor Edward Callaway, Chair
Professor Brenda Bloodgood, Co-Chair
Professor Pascal Gagneux
Professor Byungkook Lim
Professor Clodagh O'Shea
Professor Samuel Pfaff

The Dissertation of Kyle Brady Fischer is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Co-chair

_____
Chair

University of California San Diego

2018

DEDICATION

This work is dedicated to my parents, for their continuing love and support, and for instilling in their children the desire to apply their talents to causes that are bigger than themselves.

EPIGRAPH

"Well, sometimes naivety can be an asset."

-

Dr. Ed Callaway,

in response to whether it was possible

to do something we knew nothing about.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Ed Callaway for his mentorship throughout my graduate research. In sharing an enthusiasm for my often harebrained ideas, he has taught me to evaluate them with reason, ground them with intent, and nurture them with patience.

I would also like to acknowledge Dr. Euiseok Kim, who, as my scientific fairy godmother, has provided the late-night discussions, critical input, and commiseration that has shaped me as a scientist. I am deeply indebted by his guidance and friendship.

Finally, I would like to acknowledge Hannah Collins, whose perseverance and intelligence have made this work possible, and whose demeanor and friendship have made in enjoyable.

Chapter 1, in full, is in preparation for submission for publication. The dissertation author will be the first author of this work. Second author will be Hannah Collins, and Dr. Edward Callaway will be the third and final author as principle investigator.

Chapter 2, in full, is in preparation for submission for publication. The dissertation author will be the first author of this work. Second author will be Hannah Collins, and Dr. Edward Callaway will be the third and final author as principle investigator.

VITA

2009 Bachelors of Science in Biology, Boston College

2009 Bachelors of Arts in Psychology, Boston College

2018 Doctor of Philosophy, University of California San Diego

FIELDS OF STUDY

Major Field: Neurosciences

Studies in the Role of Dopamine in Nucleus Accumbans and Amygdala
Dr. Jon Horvitz, Boston College

Studies in the Role of Striatum and Basal Ganglia in Stereotyped Behaviors
Dr. Ann Graybiel, Massachusetts Institute of Technology

Studies in Viral Tool Development for Neural Circuit Dissection
Dr. Ed Callaway, The Salk Institute

ABSTRACT OF THE DISSERATION

Improving Viral Tools for Neural Circuit Dissection: Monosynaptic Anterograde Tracing and Recombinase-Dependent AAV Expression of Sensitive Transgenes

By

Kyle Brady Fischer

Doctor of Philosophy in Neurosciences
with a Specialization in Computational Neurosciences

University of California San Diego, 2018

Professor Edward Callaway, Chair
Professor Brenda Bloodgood, Co-Chair

The complex tangle of synaptic connections made by the myriad of neural cell types in mammalian nervous system, each with their own genetic, chemical, and electrophysiological properties, are the basis of neural computation and thus perception, motivation, and behavior. As such, it is a fundamental task of neuroscience to map neural connectivity in a cell-type specific manner and develop robust technologies to this end. This dissertation describes two such technologies. First, targeted deletion and *in vivo* transcomplementation of the Herpes Simplex Virus 1 (HSV) gene *UL6* in the anterograde strain H129 allows for the selective labeling of postsynaptic populations from specific neuronal cell types in a cre-dependent manner. Second, the cause of transgene expressing in off-target cells from recombinase-dependent recombinant adeno-associated virus (AAV) will be described, along with efforts toward improved methods in AAV production for

"leakless" expression of sensitive transgenes. Both techniques will expand the available toolkit of the modern systems neuroscientist and provide a means by which to tackle previously intractable questions.

Chapter 1 - Monosynaptic restriction of transsynaptic neural circuit tracing with a cre-dependent version of the anterograde HSV strain H129

Abstract

Identification of synaptic partners is a fundamental task for neural circuit analysis. To date no technique exists for whole brain labelling of downstream synaptic partners in a cell-type dependent and monosynaptically restricted manner. While several efforts have been made toward this end utilizing a variety of viruses (Beier et al., 2011; Zeng et al., 2017; Zingg et al., 2017), these techniques suffer from confounds of polysynaptic spread, low reproducibility, lack of cell-type targeting, or low expression levels in the post-synaptic population. To overcome these issues we have designed a novel monosynaptic anterograde tracing system by targeting the viral capsid pore protein UL6 of a previously described cre-dependent version of the predominantly anterograde Herpes Simplex Virus 1 (HSV) strain H129 (Lo and Anderson, 2011). Functional knockout of this gene leads to restricted spread of the virus due to the inability to package viral genomes. To facilitate monosynaptic spread, a helper AAV that expresses UL6 in a cre-dependent manner is co-infected with the ΔUL6 HSV HS129 virus in target cre-expressing neuronal populations. Application of this system to three well described non-reciprocal circuits results in rapid labeling of neurons in expected projection areas. We go on to describe characteristics of the virus important to those seeking to apply this monosynaptic anterograde system, including cre-mediated AAV and HSV recombination, cell death, and transgene expression. While some caveats may preclude certain applications, this system will prove

useful for those wishing to identify postsynaptic partners and provide a basis from which further improvements can be made.

Introduction

Herpes Simplex Virus 1 (HSV) is an enveloped virus with a roughly 150kb dsDNA genome. Like other members of the alphaherpesviridae subfamily, HSV enters the sympathetic nervous system through the epithelium, lying latent in the dorsal root ganglia until entering the lytic phase of the viral life cycle, in which viral components and complete particles are trafficked back down the axon to the epithelium, from which the virus spreads (Smith, 2012). As such, the alphaherpesviridae are the few viruses that perform anterograde trafficking and spread in mammalian neurons as part of their natural life cycle, including HSV 1 & 2 (Norgren et al., 1992), varicella-zoster virus (Grigoryan et al., 2012; Rostad and Alvord, 2002), and pseudorabies virus (PRV; more accurately, but less commonly, known as suid herpesvirus 1)(Pomeranz et al., 2005). As this trafficking plays a significant role in the pathological aspects of infection, the mechanisms behind this trafficking and spread have been well investigated (Diefenbach et al., 2008; Enquist et al., 2002; Kratchmarov et al., 2012, 2013b, 2013a; Smith et al., 2001). Unsurprisingly, these characteristics have long made both HSV and PRV attractive candidates for neuronal tracing (Card and Enquist, 2014; Norgren et al., 1992; Rinaman and Schwartz, 2004; Strack et al., 1989).

In rare cases HSV1 spreads into the central nervous system and becomes encephalitic. A tropism survey of encephalitic strains identified one, denoted H129, that spreads predominantly in the anterograde direction (Dix et al., 1983). Mutations identified from the sequencing of this strain fail to fall within any known trafficking genes or otherwise account for this phenotype, suggesting polygenetic factors controlling inhibition of retrograde spread (Szpara et al., 2010). Indeed, transsynaptic retrograde spread of the

virus has been suggested to explain retina infection by this strain following inoculation of the contralateral anterior chamber of the eye (Archin and Atherton, 2002; Archin et al., 2003) and delayed retrograde spread has been found following prolonged infection in *in vitro* expriments using cultured sympathetic superior cervical ganglion neurons (Wojaczynski et al., 2014). Nevertheless, given this directional bias, H129 has found use in anatomical studies wishing to label downstream populations from specific brain regions (Barnett et al., 1995; Dum et al., 2009; Futai et al., 2013; Kelly and Strick, 2003; McGovern et al., 2012a, 2012b; Rinaman and Schwartz, 2004; Ryu et al., 2015, 2017; Song et al., 2009; Sun et al., 1996; Tang et al., 2016; Vaughan and Bartness, 2012). Recently, H129 has been modified in order to dissect neural circuits in a cell-type specific manner (Lo and Anderson, 2011). In this H129 LSL-TK-FP virus, as it will be denoted herein, H129 has been modified by replacement of the viral thymidine kinase required for replication in non-dividing cells, by a codon-modified human thymidine kinase (TK) and fluorescent protein (FP) expression cassette controlled by a lox-stop-lox site (LSL). This design allows tracing of the outputs of cre-expressing neurons because the virus is only able to replicate in those neurons and the resulting viral particles contain the irreversibly cre-activated genome expressing a fluorescent protein to label both starter cells and downstream (mono- and polysynaptic) neurons. While this virus has proven useful (McGovern et al., 2015; Padilla et al., 2016; Scott et al., 2015), the polysynaptic nature of the H129 LSL-TK-FP virus makes circuit analysis difficult.

To limit spread monosynaptically, we employed a trans-complementation strategy, much like the monosynaptic retrograde tracing system based on the glycoprotein-deleted rabies virus (Wickersham et al., 2007), in which a viral gene necessary for spread is deleted

from the neurotropic virus and selectively rescued in a targeted neuron population, referred to as the "starter cells", from which the virus can then spread monosynaptically. Due to the complexity and large peptide size of the components of the glycoprotein-mediated viral entry mechanisms and replication machinery of HSV, we focused on the six viral genes necessary for genome packaging (Higgs et al., 2008). We reasoned that this might restrict viral spread without preventing genome replication and potentially limiting the expression of reporter genes from the small number of viral particles that might spread across synapses. Of these, we selected the gene *UL6*, which forms a dodecamer pore on the surface of the viral capsid and is required for packaging of the viral genome (Albright et al., 2011; Lamberti and Weller, 1996; Newcomb et al., 1996, 2005). Besides its necessity for the spread of infectious particles, its relatively small size, low per-viral-particle copy number, and lack of any known toxicity when expressed exogenously in mammalian cells make it an attractive target for deletion and transcomplementation.

Starting with the H129-LSL-TK-FP version expressing tdTomato (H129 LSL-TK-TT), we successfully restricted the spread of the virus *in vitro* and *in vivo* via functional UL6 deletion. Further, transcomplementation with AAV expressing UL6 in well described non-reciprocal circuits results in rescue of viral spread from cre-expressing populations, as evidenced by viral labeling of neurons in expected output regions. In executing these experiments we identified several caveats of important consideration to those wishing to apply this system. First, we found that cre-mediated recombination across like lox sites in H129-LSL-TK-TT ΔUL6 and cre-dependent AAVs can lead to unwanted intraviral recombination and reconstitution of UL6 expression in the ΔUL6 virus. This is overcome by the use of orthogonal lox sites throughout the system. Second, due to the high

cytotoxicity of HSV, we investigated the rate of cell death in cells containing both cre-activated and inactivated versions of the H129 LSL-TK-TT ΔUL6 virus. Finally, we discuss issues resulting from low tdTomato expression from the H129 LSL-TK-TT virus.

Materials and Methods

Cells and Media

Vero cells (ATCC CCL-81) were used for propagation of UL6-intact viruses. The UL6-complementing Vero cell line 31-UL6 was kindly provided by Dr. Sandra Weller. Vero cell lines were grown in 10% Fetal Bovine Serum (FBS; HyClone SH 30070.03) DMEM media (Gibco 11995-040) and 31-UL6 cells were grown in a 5% FBS Sodium Pyruvate negative DMEM media (Gibco 11956-084) with 0.25 mg/ml Geneticin (Gibco 10131-027). For plaque picking, a 1% Methylcellulose (MC; MP Biomedicals 155496) media was made with the appropriate aforementioned medias. All cells were grown at 37C and 5% $C0_2$.

Viral Production and Titering

HSV H129 LSL TK-TT and TK-GFP were kindly provided by Dr. David Anderson (Lo and Anderson, 2011). For viral production, a single plaque was collected from a 6-well plate, freeze-thawed thrice, and applied to a 10cm plate of confluent vero or 31-UL6 cells. Once the plate had reached 100% cytopathic effect (cpe) or cells became noticeably loose, the media was collected and clarified by centrifugation. This media was then split

and applied to ten 15cm plates of the appropriate cell line. Having reached 100% cpe (typically ~3 days), the media from these plates was clarified, filtered through a 0.45um filter (Millipore SCHVU02RE) and stored at 4C for a few days or -80C if longer than a week. To concentrate, the viral supernatant was spun at 19.4K rpm for two hours in a Beckman-coulter SW-28 rotor. Viral pellets were resuspended in 100ul of HBSS (Gibco 14175-095) and aggregated atop a 2.5ml 20% w/v sucrose in HBSS solution. This was then spun for two hours at 21K rpm in a SW55 rotor. The supernatant was removed and the pellet allowed to dry by placing the centrifuge tube upside down on a paper towel before resuspension in 50-100ul of HBSS. Virus was stored at 4C overnight before final pellet dissolution and aliquoting. Viral aliquots were stored at -80C.

To titer virus, 2ul of viral aliquot as added to 1998ul of warmed 31-UL6 cell media and serially diluted 10-fold 6 times to produce a 10^{-3} to 10^{-9} serial dilution. 250ul of each dilution was added to 3 wells of confluent 31-UL6 cells on a 24 well plate. Wells were monitored for 3 days to track the growth of plaques. Once easily discernable, plaques were counted for the dilution condition containing less than 100 plaques and counts were averaged across the three wells. This average was used to calculate the plaque forming units (pfu) per ml. The titers of the HSV preparations used here were as follows: H129 LSL-TK-TT, 1.86 x 10^9 pfu/ml; H129 LSL-TK-TT ΔUL6, 1.066 x 10^7 pfu/ml; H129 LSL-TK-GFP-ΔUL6-mCherry, 1.2 x 10^7 pfu/ml.

All AAVs were prepared by the Salk Vector Core and use a UL6 coding sequence codon optimized for mouse by Genscript. Titer for the AAV8-nef-2N-DIO-GFP-p2a-coUL6 virus used for GPi, and dSub injections was 3.22 x 10^{12} GC/mL. Titer for the AAV2-nef-2N-DIO-GFP-p2a-coUL6 used for retina injections was 3.48 x 10^{11} GC/mL.

UL6-knockout production

To extract HSV DNA, media was first aspirated from a H129 LSL-TK-TT inoculated 15cm plate of vero cells showing 100% cpe. 5mls of lysis buffer (1ml of 10% SDS in 49ml TE buffer, 10ug/ml DNAse-free RNAse A) was applied to the cells, swirled, and collected in a 15ml tube. 1ul of proteinase K (NEB #P8107S) was added and the solution was incubated at 37C for 90 minutes. Phenol:chloropform DNA extraction was then performed and DNA resuspended in TE buffer following purification.

The UL6 knockout was generated using a modified two-step fluorescent protein insertion and deletion selection process, as outlined previously (Sawtell and Thompson, 2014). For initial interruption of the UL6 gene, a linearized GFP expression cassette flanked upstream and downstream, respectively, by 1500bp and 1000bp regions homologous to the target site of integration in the HSV genome was transfected with PEI into a 6 well plate of 31-UL6 cells at a 1:3 ratio with H129 LSL-TK-TT DNA. Six days later the cells were collected, resuspended in 0.4ml of cold dPBS, and aliquoted to 100ul. The volume of each aliquot was brought up to 1ml with cold dPBS and stored at -80. A single aliquot was freeze-thawed three times in liquid nitrogen and a 37C water bath before being diluted 10^{-2}, 10^{-3}, and 10^{-4} in cold dPBS. 400ul of each dilution was added to two wells of a 6 well plate of confluent 31-UL6 cells and the media was replaced with 1% MC media one hour after inoculation. Within 3 days fluorescent plaques were visible.

To pick plaques, MC media was aspirated and the cells carefully washed with warm dPBS, followed by the addition of 1 - 2 mls of warm dPBS to the well. As many as six plaques were picked and transferred to PCR tubes, where the volume was brought up to

10ul with dPBS. Three rounds of freeze/thaw were performed using liquid nitrogen and a 37C water bath, and the viral plaque solution was diluted to 10^{-1}, 10^{-2}, and 10^{-3}. Between 2 and 4ul of each dilution was applied to a single well of a 24 well plate of 31-UL6 cells, and the media was switched with 1% MC media one hour after inoculation. This plaque picking protocol was repeated every 2 to 4 days for at least 3 rounds past the point at which all plaques were fluorescent.

Once functional deletion of UL6 was confirmed, the fluorescent protein used for selection was removed similarly. In an effort to remove as much of the UL6 coding region as possible without interfering with cis-acting regions important for UL5 and UL7 regulation, six Cas9 guide RNA were designed targeting regions within the UL6 coding region, three upstream and three downstream of the inserted fluorescent protein expression cassette (Table 1). Guide sequences were cloned into CRISPR/Cas9 expression plasmid pX458 and 1 ug of each upstream and downstream plasmid combination was transfected with PEI in each well of a 12 well plate of 31-UL6 cells. Twenty-four hours after transfection, 0.1 MOI of H129 LSL-TK-TT ΔUL6-GFP or HSV H129 LSL-TK-GFP ΔUL6-mCherry was applied to each well. Once all cells displayed cytopathic effect, the cells were collected and the plaque picking of fluorescence negative plaques began as previously described.

Mice and Viral Injections

The following mouse lines were used: Som-ires-cre (Jackson Labs 013044), Slc17a6-cre (Jackson Labs 016963), Sim1-Cre KJ18 (MGI:4367070), and Ntsr1-cre

(MMRRC # 030648-UCD). Male and female mice were used and injected between the ages of 45 and 101 days.

Tissue was collected 5 days after HSV injection in experiments using Slc17a6-cre and Sim1-cre mice to compare H129 LSL-TK-TT and H129 LSL-TK-TT ΔUL6 viruses. For transcomplementation experiments, all HSV injections occurred two weeks following AAV injection, and tissue was collected seven days after HSV injection. All HSV and AAV injections in the brain were performed using pressure and delivered a volume of 200nl for each virus. Som-ires-cre mice were used to target globus pallidus internal segment using the stereotactic coordinates AP -1.1, ML 2.1, DV 4.2; Slc17a6-cre mice were used to target dorsal subiculum using the stereotactic coordinates AP -4.04, ML 2, DV 1.5; primary visual cortex injections of Ntsr1-cre and Sim1-cre mice used the sterotactic coordinates AP -3.28 ML 2.5, DV 0.5. Intravitreal retina injections in Slc17a6-cre mice were performed by creating a small incision with the tip of a 27G needle posterior to the scleral line and delivering 300nl of AAV via pressure. Retinal injections of HSV were 200nl. All injections were conducted using pulled glass pipettes (VWR 53432-706, approximate inner diameter of 0.357mm microns).

Tissue Processing and IHC

Mice were anesthetized with euthasol and intracardially perfused using 50mls of PBS and 100mls of 4% paraformaldahyde (PFA). Following perfusion, brains were dissected and placed in 5mls of 2% PFA and 15% sucrose in PBS at 4C until having sunk in this solution. Brains were then transferred to 5mls of 30% sucrose in PBS at 4C.

Following sinking, brains were cut on a microtome at 50um and slices were stored in solution of 1% NaN_3 in PBS.

Eyes were dissected along with the brain following perfusion and put in PBS. Retinas were immediately dissected and placed in 4% PFA for one hour at 4C. After being fixed, retinas were stored in 1% NaN_3 in PBS at 4C until staining.

Tissue staining started with a 5 minutes wash of 0.1% Triton X-100 in PBS followed by a 2 to 3 hour block comprised of 5% normal donkey serum in 0.5% Triton X-100 in PBS at room temperature. Primary staining occurred overnight at 4C. Primary antibodies including rabbit anti-dsRed (Takara 632496), rabbit anti-HSV (Novus NB 120-9533), and goat anti-GFP (Rockland 600-101-215) were applied at 1:1000 in block. Following primary staining, tissue was washed for 10 minutes in PBS three times. Secondary antibodies were diluted at 1:500 in PBS and applied to the tissue for 2 to 3 hours at room temperature. Secondary antibodies included Alexaflour 568 donkey anti-rabbit (Invitrogen A10042), Alexaflour 647 donkey anti-rabbit (Invitrogen A31573), and Alexaflour 488 donkey anti-goat (Invitrogen A11055). Following secondary staining, tissue was washed thrice for 10 minutes in PBS. 10um DAPI was applied for 10 minutes following by three five-minute PBS washes. Tissue was stored in 1% NaN_3 in PBS at 4C prior to mounting.

AAV/HSV recombination

AAV DIO plasmids utilizing different orthogonal *lox* pairs were transfected into vero cells 48 hours prior to H129 LSL-TK-TT ΔUL6 application at an MOI of 1. Forty-

eight hours after viral application media from these wells was collected and clarified by centrifugation at 300rpm. 1ml of this media was applied to a confluent well of a 6-well plate of vero cells. Images were taken 48 hours later using a dSLR camera mounted on a fluorescent microscope.

Flow cytometry

The shuttle plasmid used in the production of the HSV H129 LSL-TK-TT virus was kindly provided by Dr. David Anderson. 1ug of both the pHSV-CAG-LSL-TK-2a-TT plasmid and pAAV-EF1α-iCre were transfected by PEI into a single well of a 6-well plate of HEK 293 cells. As a control, 1 ug of both pAAV-nef-DIO-tdTomato and pAAV-Ef1α-iCre were transfected in a separate well in an identical fashion. Forty-eight hours after transfection cells were trypsinated, collected in 2ml Eppendorf tubes, and spun at 500 rcf for 3 minutes. The media was aspirated and the cellular pellet resuspended in 1ml of cold 1% FBS in dPBS. The cells were spun down again, the media aspirated, and cells resuspended in 1ml of cold 10um DAPI in 1% FBS in dPBS. Cells were spun down, the media aspirated, and cells resuspended in 500ul of cold 1% FBS in dPBS. This solution was moved to filter-topped tubes (Falcon 352235) on ice and run on a LSRII in the Salk Flow Cytometry Core Facility.

Results

Functional deletion of the *UL6* gene in the H129 LSL-TK-TT virus leads to restriction of viral spread

Functional knockout of the *UL6* gene from the H129 LSL-TK-TT virus required a two-step selection process (Fig. 1A) in which a GFP expression cassette was first inserted in the middle of the UL6 locus. This expression cassette was designed to split rather than fully replace the *UL6* coding sequence to avoid disrupting cis-acting regions governing the expression of the neighboring *UL5* and *UL7* genes. Following several rounds of purification by plaque picking, the virus was applied to monolayers of vero cells, vero cells constitutively expressing UL6 (31-UL6 cells), and vero cells transfected with a plasmid expressing a codon-optimized UL6 (coUL6; Fig 1B). The absence of plaques in vero cells indicates that the GFP expression cassette insertion impairs viral spread (Fig 1B, left). Further, viral spread was rescued by UL6 transcomplementation, as GFP+ plaques were found in conditions in which UL6 expression was present either by constitutive expression or plasmid transfection (Fig 1B, center & right), indicating that restriction of viral spread is a function of *UL6* disruption.

As the GFP expression cassette was only necessary as a means of selection of viruses in which *UL6* had been interrupted, the cassette was removed in a second round of selection using CRISPR/Cas9 (Fig 1A, bottom). Following exposure to Cas9 and pairs of guide RNA targeting the UL6 regions flanking the inserted GFP expression cassette, virus was diluted and applied to 31-UL6 cells. Plaques were selected and propagated based on

the absence of GFP fluorescence. Following three rounds of selection past the absence of fluorescence in all plaques, H129 LSL-TK-TT ΔUL6 virus was grown up and concentrated. Sequencing of PCR fragments targeting the UL6 locus of the H129 LSL-TK-TT ΔUL6 used herein revealed the expected frame shift after 377 amino acids in the *UL6* coding sequence, followed by 77 amino acids before a stop codon.

To test if spread of the ΔUL6 virus is also restricted *in vivo* both the original H129 LSL-TK-TT and H129 LSL-TK-TT ΔUL6 virus, they were injected into two cre mouse lines. Slc17a6-cre mice express cre in retinal ganglion cells(Martersteck et al., 2017). Retinal injection of the cre-dependent polysynaptic H129 LSL-TK-TT virus into Slc17a6-cre mice led to HSV antibody labeling throughout the visual system when tissue was collected 5 days post-injection, indicative of polysynaptic spread (Fig 1C, left column). When injected with H129 LSL-TK-TT ΔUL6, no HSV antibody labeling was found in the brain (Fig 1C, right column).

Testing in the Layer 6 mouse line Sim1-cre yielded similar results. Five days following injection of H129 LSL-TK-TT into primary visual cortex diffuse spread was observed throughout the brain, most notably in cortex, thalamus, hippocampus, and other subcortical areas (Fig 1D, left column). H129 LSL-TK-TT ΔUL6 injection resulting primarily in HSV antibody labeling near the injection site as well as some labelling in known regions providing input to V1, such as thalamus and contralateral cortex (Fig 1D, right column), likely due to expected axonal uptake and retrograde spread (Lo and Anderson, 2011). The results from these and the retinal injections of Slc17a6-cre mice indicate that the ΔUL6 virus is unable to spread from the original site of infection *in vivo*.

Transcomplementation of the H129 LSL-TK-TT ΔUL6 virus with an AAV expressing UL6 leads to anterograde monosynaptic spread in well described circuits

To test if complementation of the ΔUL6 virus in vivo lead to monosynaptic spread of the virus while ensuring that any identified spread was not due to direct or transneuronal retrograde spread, we targeted three well described circuits containing long distance projections with no known direct reciprocity: retinal ganglion cells (RGCs; cell type vglut2; mouse line Slc17a6-cre) to lateral geniculate nucleus (LGN) and superior colliculus (SC) (Martersteck et al., 2017); globus pallidus internal (GPi; cell type somatostatin; mouse line Som-ires-cre) to lateral habenula (LH) (Shabel et al., 2012; Wallace et al., 2017); and dorsal subiculum (dSub; cell line vglut2; mouse line Slc17a6-cre) to mammillary body (Kinnavane et al., 2018; Roy et al., 2017; Witter, 2006).

Two important design choices were made in the transcomplementing AAV construct. First, to better segregate starter cells from cells to which the HSV has spread, we utilized a cre-dependent AAV system and an orthogonal fluorescent protein, in this case GFP, was used to differentiate AAV and HSV infected cells. Second, due to the large portion of the original UL6 coding region still present in the H129 LSL-TK-TT ΔUL6 virus and the propensity of the HSV genome to recombine with DNA containing homologous regions we codon optimized the UL6 coding region for mouse, termed here coUL6, to decrease the sequence homology of the UL6 gene present in the complementing AAV from that in the HSV genome. This modification would impart the further benefit of improved transgene expression,

Perplexingly, early transcomplementation experiments using this AAV-nef-DIO-GFP-p2a-coUL6 construct led to frequent GFP expression in off-target neuron populations

when compared with injections of the AAV helper virus alone (Fig 2A). As this fluorescence appeared to be mediated by the HSV, we hypothesized that cre-dependent recombination across the *loxp* sites present in both the HSV loxp-stop-loxp cassette and AAV DIO construct were resulting in integration of the GFP, and potentially coUL6, coding sequences that are present between the *loxp* sites in the AAV. To test this, we designed cre-dependent pAAV plasmids expressing GFP-p2a-coUL6 using three different orthogonal pairs of *lox* elements in a DIO configuration (Fig 2B). Plasmids containing each construct were transfected in monolayers of vero cells with cre recombinase and H129 LSL-TK-TT ΔUL6 virus applied (Fig 2C). To test if the HSV had taken up the GFP and coUL6 from the transfected plasmids, media from each condition was collected and applied to a fresh well of vero cells 48 hours after the original HSV application. When imaged, conditions that had been transfected with cre recombinase and AAV-DIO plasmids containing the compatible *loxp* site resulted in GFP and tdTomato positive cells and HSV propagation, indicating that both GFP and the coUL6 genes had been incorporated into the H129 LSL-TK-TT ΔUL6 genome. As vero cells having received media from the condition containing the DIO construct using the orthogonal *lox2722* or *loxN* contained no evidence of HSV spread, and thus no evidence of recombination, we utilized this AAV-2N-DIO construct for all subsequent *in vivo* transcomplementation experiments.

Control injections without transcomplementing AAV were conducted in each circuit in which the system was applied. H129 LSL-TK-TT ΔUL6 virus was injected in the vitreous of the right eye of a Slc17a6-cre mouse and tissue collected 7 days later. Flat-mount of the injected retina revealed retinal ganglion cell (RGC) infection by anti-HSV staining (Fig 3A) but no HSV+ cells in the brain, including the known projection areas of

visual thalamus (Fig 3B) and superior colliculus (Fig 3C). In comparison, complementation of H129 LSL-TK-TT ΔUL6 with AAV2-nef-2N-DIO-GFP-p2a-coUL6 injected two weeks prior to HSV injection lead to obvious viral spread in these downstream nuclei. Retina flat mount (Fig. 3D) reveals widespread GFP and tdTomato positive RGCs, indicative of AAV and HSV infection, respectively, as well as cre-recombination of the HSV genome. RGCs that are both red and green indicate putative starter cells (Fig 3D, inset). In the brain, HSV antibody staining in visual thalamus (Fig 3E) and superior colliculus (Fig 3F) revealed HSV positive neurons amongst GFP positive axons of RGC origin. While cell death was evident in some HSV infected neurons by membrane blebbing or lack of a cell body, non-neuronal HSV labelling was rare and appeared to be associated with glial response to infected neurons when identified. Healthy neurons with discernable morphology were abundant amongst HSV labeled cells (Fig 3E & F, inset).

The same injection schedule was performed on Som-ires-cre mice targeting the internal segment of the globus pallidus (GPi) in which SOM+ neurons send projections to the lateral habenula(Shabel et al., 2012; Wallace et al., 2017). Tissue collected 7 days after control injections of H129 LSL-TK-TT ΔUL6 into GPi (without helper AAV injections) showed HSV positive neurons in the target population as well as some labeling in neighboring nuclei (Fig 4A). No HSV labeling was found in the habenula (Fig 4B). The injection site in experiments complemented by AAV8 nef-2N-DIO-GFP-p2a-coUL6 show both GFP and HSV positive neurons in the target region. GFP positive axons originating from GPi were found in lateral habenula (LH) amongst which many HSV positive neurons were also found (Fig 4D). While some infection of both the AAV and HSV was found in regions surrounding GPi , particularly by AAV in the central amygdala nucleus, we have

no reason believe this expression could lead to a false-positive result in interpreting the noted HSV labeling in LH.

Finally, injections were performed targeting vglut2 cells of the dorsal subiculum (dSub) using a Slc17a6-cre mouse line. This neuron population sends long projections via the columns of fornix to the lateral mammillary bodies (LM) from which they receive no known direct reciprocal connection(Kinnavane et al., 2018; Roy et al., 2017; Witter, 2006). Tissue collected 7 days after control injections of H129 LSL-TK-TT ΔUL6 injection into the dSub of a Slc17a6-cre mouse (without helper AAV injections) resulted in HSV staining in the dSub, postsubiculum, and retrosplenial cortex (Fig 5A). This labeling pattern is expected due to direct infection of neurons at the injections site or which project axons to the injection site and directly take up HSV via their terminals. No labelling was found in the ipsilateral mammillary body (Fig 5B). In tissue collected 7 days following H129 LSL-TK-TT ΔUL6 injection that had been complemented with AAV8 nef-2N-DIO-GFP-p2a-coUL6, obvious GFP and HSV positive cells were found in dSub (Fig 5C). Some HSV+ cells can also be seen amongst GFP+ axons projecting from this population to ventral retrosplenial cortex (RSPv), but due to the possibility of locally reciprocal connectivity between RSPv and dSub these cells may result from retrograde spread from the injection site and are not submitted here as evidence of monosynaptic anterograde spread. At the terminus of the columns of the fornix, of which the projections from dSub to the LM comprise, HSV+ neurons are found amongst GFP+ axons in the LH (Fig 5D), indicating the expected anterograde transynaptic spread from dSub neurons.

These data collectively indicate monosynaptic restriction of the H129 LSL-TK-TT ΔUL6 virus from the targeted cre-expressing populations when complemented by an AAV

helper virus expressing UL6 as HSV+ neurons were found amongst GFP+ axons in known projection areas while not evidence of polysynaptic spread (ie, HSV+ neurons in areas outside of those regions receiving direct invervation from the target population) were identified..

HSV infection leads to rapid cell death in cre-activated cells

Cell death is an obvious hallmark of replication competent neurotropic viruses, HSV being no exception. While loose tissue and abscesses are common in experiments using the polysynaptic virus H129 LSL-TK-TT (see Fig 1D), dense DAPI staining indicative of gliosis was commonly seen at the injection sites of the ΔUL6 version of the virus (Fig 6A). Cell death was evident in postsynaptic neurons labeled by HSV as well. LGN labeled neurons resulting from transcomplementation of H129 LSL-TK-TT ΔUL6 in the retina of Slc17a6 displayed both healthy morphologies and cytopathic effect (Fig 6B) evident both in gross morphology and increased local DAPI density.

To assess the rate of cell death in both cre-expressing and non-cre-expressing neural populations we injected a H129 LSL-TK-GFP ΔUL6-mCherry virus into the primary visual cortex of Som-ires-cre and Ntsr1-cre mice. All cells infected with this virus should express mCherry, while cells that express cre and recombine the virus will express both TK and GFP. Somatostatin cortical interneurons were chosen to assess longevity of neighboring cre-negative neurons infected with the virus because sparse somatostatin interneuron distribution ensured that any collateral damage associated with their cell death resulting from cre-activated TK expression and viral replication would be minimal. Tissue

collected 4, 8, and 16 days post injections showed robust mCherry expression, suggesting a latent-like infection in non-cre-expressing neurons (Fig 6C). Due to the sparse nature of cortical somatostatin neurons very few GFP+ neurons were found in these experiments, and the experiment was repeated using the cortical Ntsr1-cre layer 5 cell line, a far denser neuronal population. Well labeled morphologically healthy GFP+ neurons were found 2 days after injection, but very few were found at 7 days post-injections, and none at 14 (Fig 6D). These data show that thymidine kinase expression rapidly activates the virus in cre-expressing populations, so rapidly that, in transcomplementation experiments, the time required to visualize postsynaptic populations is likely longer than the survival of the starter cell population.

Unreliable tdTomato expression from the cre-dependent expression cassette confounds analysis of transcomplementation experiments

The H129 LSL-TK-TT version of the virus (opposed to the H129 LSL-TK-GFP version) was initially used for ΔUL6 virus production due to the known high fluorescence of tdTomato and its use in previous publications. In our hands, tdTomato signal from injections of the H129 LSL-TK-TT virus and its derivatives proved unreliable, resulting in the use of an HSV polyclonal antibody to identify infected cells. Typically, HSV infected neurons in the target cre-expressing population or in the postsynaptic population, such as those found in LH following transcomplementation in somatostatin neurons of the GPi (Fig 7A, top), showed very little tdTomato expression. This low expression could not be sufficiently amplified by a battery of immunohistological techniques (data not shown). In contrast, in infrequent cases, such as in the transcomplementation experiment targeting

vglut2 neurons in dorsal subiculum, tdTomato expression levels were strong enough to identify neurites without counterstaining (Fig 7A, bottom). Virus aliquots used in both these experiments were from the same production batch.

We hypothesized that the weak expression levels were due to some unknown defect in the LSL-TK-TT expression cassette used to make the original cre-dependent version of the virus. To test this, we transfected HEK 293 cells with plasmids containing the HSV expression cassette, pHSV-LSL-TK-2a-tdTomato, and a plasmid expressing cre recombinase. As a control, the cre-dependent AAV plasmid pAAV-nef-DIO-tdTomato and was transfected along with a plsmid expressing cre. Flow cytometry conducted 48hrs after transfections resulted in no discernable difference in tdTomato expression in either conditions (Fig 7B). These results, along with the seemingly stochastic nature of the tdTomato expression and the promiscuity of the HSV genome to recombination, has lead us to hypothesize that the duplicative nature of the tdTomato sequence leads to frequent recombination events that greatly reduce or abolish tdTomato expression or fluorescence. As such, we have developed a H129 LSL-TK-GFP ΔUL6 virus, which has shown much more reliable expression levels (eg, Fig 6D).

Discussion

Herein we show that the H129 LSL-TK-TT ΔUL6 virus can be used for cre-dependent monosynaptic anterograde spread when complemented with an AAV expressing UL6 by application of the system to well described non-reciprocal circuits (Fig 3 – 5). While this tool fills a hole in the viral toolkit available to the modern systems neuroscientist, potential users of this system should be aware of its limitations and caveats.

As a replication competent virus, at least after cre-activation, HSV causes rapid cell death. Our experiments showed cell blebbing as soon as two days after injection, and nearly total disappearance of infected cre-positive neurons by one week post-injection. As such, depending on the circuit, starter cell populations are likely to have died by the time transneuronal spread is evident. Indeed, clearly overlapping HSV and AAV neurons were scarce at the injection site of experiments in which anterograde spread was evident (Fig 3 – 5), suggesting that the cells responsible for spread had died by time of tissue collection. Depending on the circuit, however, the incubation time may be shortened: while all data shown here came from experiments in which tissue was collected 7 days following HSV injection, transneuronal spread was evident in as few as 5 days in experiments targeting RGCs and 3 days in those targeting the Ntsr1 cortical layer 5 cell population (data not shown). Despite this, the rate of cell death is likely to make manipulations that require viable pre- and postsynaptic partners difficult.

We utilized a polyclonal HSV antibody for labelling of HSV infected cells in most experiments due to sporadic and unreliable tdTomato expression from the H129 LSL-TK-

TT virus (Fig 7A), which was surprising due to the typically high fluorescence of tdTomato. This made segregation of HSV-infected neurons that contained inactivated or cre-activated versions of the virus difficult. There are several possible causes of this weak expression, including HSV interference in the CMV-derived region of the CAG promoter present in the introduced LSL-TK-TT cassette and incomplete cleavage at the 2a element between the thymidine kinase and tdTomato. However, since infrequently, injections of H129 LSL-TK-TT ΔUL6 resulted in strong tdTomato flourescence despite being from the same viral batch as those that did not, and the fact that H129 LSL-TK-GFP ΔUL6 brightly labeled infected cells despite being in the same expression cassette, both possibilities seem unlikely. Further, comparisons of the LSL-TK-TT expression cassette with a cre-dependent tdTomato AAV plasmid yielded similar expression profiles (Fig 7B). It is our hypothesis that the repetitive nature of the tdTomato coding sequence makes it susceptible to recombination or mutation, and as such potential users may choose to use the GFP version of the virus until a version containing a monomeric red fluorescent protein is developed.

Assessing the efficiency of transsynaptic tracers is difficult, as the true number of connected neurons and synapses comprising those connections is typically unknown. Further, the rapid rate of cell death of starter neurons and our inability to reliably differentiate cre-activated starter cells from local cre-negative neurons infected with HSV made quantifying the number of labelled pre to post-synaptic neurons impossible. It is our hope that iterations of the virus currently being tested will allow for identification of both starter cells and their postsynaptic partners so that obvious biases in spread may be assessed.

An important caveat for potential users of this tool is HSV's strong propensity for axonal uptake from the site of injection and thus retrograde infection . In systems in which cre-expressing populations exist presynaptic to the target region, HSV LSL-TK-FP ΔUL6 uptake at those axon terminals would lead to cre-activation and potential falsely identified postsynaptic partners. As such, great care should be taken in avoiding such instances and results should always be compared to expression patterns seen in control injections of the H129 LSL-TK-FP ΔUL6 virus alone.

Two major experiments remain to be executed for proper completion of this project. The first is in regards to the potential transneuronal retrograde spread of the H129 virus in this system, based on previous reports of delayed retrograde spread (Archin and Atherton, 2002b; Archin et al., 2003; Wojaczynski et al., 2014). In our experiments we have noted no obvious transneuronal retrograde spread of the virus in our transcomplementation experiments, but due to the lack of reliable fluorescent protein expression from the virus, direct retrograde infection from the injection site is impossible to differentiate from transneuronal spread. To get around this, we are currently performing AAV8 nef-2N-DIO-GFP-p2a-coUL6 injections into the LGN of the thalamus-specific crh-cre mouse line, followed by H129 LSL-TK-TT ΔUL6 injection in V1. Axonal uptake by CRH+ neurons projecting from LGN to V1 will result in transcomplementation, and any HSV labeling present in the retina, impossible to infect directly from V1, would be evidence of retrograde spread. Results from these experiments will be included in the final publication of this work.

Previous studies using the H129 strain have relied upon the assumption that the virus travels in a synapse specific manner. This assumption is not without evidence,

however, as the oft repeated tracing studies from retina to visual cortex (Fig 1C; Lo and Anderson, 2011; Sun et al., 1996), would result in obvious polysynaptic spread to non-visual areas if spread were not synapse specific. Yet the difficulty in assessing synapse specificity of the polysynaptic virus, its propensity to induce rapid cell death, and cell-to-cell fusion all encourage direct evidence of synapse specificity. Unfortunately, due to the rapid death of the starter population, using channelrhodopisin to activate starter cells while recoding from labelled putative postsynaptic cells in the output region is impossible. As has been previously carried out for rabies tracing (Marshel et al., 2010), replicating this system in *in vitro* organotypic slices using biolistics, where viral spread can be easily monitored for rapid channelrhodopsin activation and recording of postsynaptic cells, has proven difficult due to the high ph sensitivity of herpes' fusogenic glycoproteins (data not shown).

As such, we are currently using an anatomical approach to investigate the synapse specificity of the H129 ΔUL6 system by targeting pontine vglut1 neurons that send highly specific projections to granule cells in the cerebellum (Huang et al., 2013). The specificity of this connection and the numerous anatomically and morphologically distinct populations in the vicinity should provide for a readout of synapse specificity. Moreover, the large nature of the pontine-granular synapse should make synapse specificity obvious.

The H129 LSL-TK-FP ΔUL6 tracing system is useful for the identification of postsynaptic partners of specific neuronal cell types. While the technique is far from an ideal transsynaptic tracer, it is our hope that this system will serve as a chassis upon which further improvements can be made. Beyond thymidine kinase, the latent-lytic life cycle of HSV provides for many genetic levers upon which further manipulation may allow for

decreased cell death by induction of latency or temporally control of spread (Knipe and Cliffe, 2008; Mador et al., 1998; Umbach et al., 2008). Further, ΔUL6 HSV mutants have been shown to produce virions with empty capsids (Nellissery et al., 2007) which could be manipulated to carry a proteinaceous payload. Indeed, non-infectious empty viral-like particles termed "L-particles" (for "light"; in comparison to "heavy" genome-bearing bands found in viral preparations ultracentrifuged in gradient solutions) are known to play important roles in the herpes life cycle (Szilágyi and Cunningham, 1991), including intercellular transmission of viral proteins (Heilingloh et al., 2015), and have been found to travel down axons similarly to infectious particles (Ibiricu et al., 2013). Recent work on engineered retroviral virus-like particles have shown that these empty viral structures can be used to carry functional proteins to foreign cells, including cre (Kaczmarczyk et al., 2011). Similarly, fusion of a fluorescent protein to the common viral tegument protein VP22 has been shown to lead to efficient packaging of the fusion protein into HSV particles (Elliott and O'Hare, 1999). Taken together, it may be possible to further modify the H129 LSL-TK-FP ΔUL6 virus to carry flp recombinase in empty particles transsynaptically, removing the need for AAV transcomplementation and potentially relieving the postsynaptic cell of the toxic pressures associated with viral infection.

Acknowledgements

Chapter 1, in full, is in preparation for submission for publication. The dissertation author will be the first author of this work. Second author will be Hannah Collins, and Dr. Edward Callaway will be the third and final author as principle investigator.

References

Albright, B.S., Nellissery, J., Szczepaniak, R., and Weller, S.K. (2011). Disulfide Bond Formation in the Herpes Simplex Virus 1 UL6 Protein Is Required for Portal Ring Formation and Genome Encapsidation. J. Virol. *85*, 8616–8624.

Archin, N.M., and Atherton, S.S. (2002a). Infiltration of T-lymphocytes in the brain after anterior chamber inoculation of a neurovirulent and neuroinvasive strain of HSV-1. J. Neuroimmunol. *130*, 117–127.

Archin, N.M., and Atherton, S.S. (2002b). Rapid spread of a neurovirulent strain of HSV-1 through the CNS of BALB/c mice following anterior chamber inoculation. J. Neurovirol. *8*, 122–135.

Archin, N.M., Van den Boom, L., Perelygina, L., Hilliard, J.M., and Atherton, S.S. (2003). Delayed spread and reduction in virus titer after anterior chamber inoculation of a recombinant of HSV-1 expressing IL-16. Investig. Ophthalmol. Vis. Sci. *44*, 3066–3076.

Barnett, E.M., Evans, G.D., Sun, N., Perlman, S., and Cassell, M.D. (1995). Anterograde tracing of trigeminal afferent pathways from the murine tooth pulp to cortex using herpes simplex virus type 1. J. Neurosci. *15*, 2972–2984.

Beier, K.T., Saunders, A., Oldenburg, I. a, Miyamichi, K., Akhtar, N., Luo, L., Whelan, S.P.J., Sabatini, B., and Cepko, C.L. (2011). Anterograde or retrograde transsynaptic labeling of CNS neurons with vesicular stomatitis virus vectors. Proc. Natl. Acad. Sci. U. S. A. *108*, 15414–15419.

Card, J.P., and Enquist, L.W. (2014). Transneuronal circuit analysis with pseudorabies viruses. Curr. Protoc. Neurosci. 1–39.

Diefenbach, R.J., Miranda-Saksena, M., Douglas, M.W., and Cunningham, A.L. (2008). Transport and egress of herpes simplex virus in neurons. Rev. Med. Virol. *18*, 35–51.

Dix, R.D., McKendall, R.R., and Baringer, J.R. (1983). Comparative neurovirulence of herpes simplex virus type 1 strains after peripheral or intracerebral inoculation of BALB/c mice. Infect. Immun. *40*, 103–112.

Dum, R.P., Levinthal, D.J., and Strick, P.L. (2009). The Spinothalamic System Targets Motor and Sensory Areas in the Cerebral Cortex of Monkeys. J. Neurosci. *29*, 14223–14235.

Elliott, G., and O'Hare, P. (1999). Live-Cell Analysis of a Green Fluorescent Protein-Tagged Herpes Simplex Virus Infection. J. Virol. *73*, 4110–4119.

Enquist, L.W., Tomishima, M.J., Gross, S., and Smith, G. a (2002). Directional spread of an alpha-herpesvirus in the nervous system. Vet. Microbiol. *86*, 5–16.

Futai, K., Doty, C.D., Baek, B., Ryu, J., and Sheng, M. (2013). Specific trans-synaptic interaction with inhibitory interneuronal neurexin underlies differential ability of neuroligins to induce functional inhibitory synapses. J. Neurosci. *33*, 3612–3623.

Grigoryan, S., Kinchington, P.R., Yang, I.H., Selariu, A., Zhu, H., Yee, M., and Goldstein,

R.S. (2012). Retrograde axonal transport of VZV: Kinetic studies in hESC-derived neurons. J. Neurovirol. *18*, 462–470.

Heilingloh, C.S., Kummer, M., Mühl-Zürbes, P., Drassner, C., Daniel, C., Klewer, M., and Steinkasserer, A. (2015). L Particles Transmit Viral Proteins from Herpes Simplex Virus 1-Infected Mature Dendritic Cells to Uninfected Bystander Cells, Inducing CD83 Downmodulation. J. Virol. *89*, 11046–11055.

Higgs, M.R., Preston, V.G., and Stow, N.D. (2008). The UL15 protein of herpes simplex virus type 1 is necessary for the localization of the UL28 and UL33 proteins to viral DNA replication centres. J. Gen. Virol. *89*, 1709–1715.

Huang, C.C., Sugino, K., Shima, Y., Guo, C., Bai, S., Mensh, B.D., Nelson, S.B., and Hantman, A.W. (2013). Convergence of pontine and proprioceptive streams onto multimodal cerebellar granule cells. Elife *2013*, 1–17.

Ibiricu, I., Maurer, U.E., and Grünewald, K. (2013). Characterization of herpes simplex virus type 1L-particle assembly and egress in hippocampal neurones by electron cryo-tomography. Cell. Microbiol. *15*, 285–291.

Kaczmarczyk, S.J., Sitaraman, K., Young, H.A., Hughes, S.H., and Chatterjee, D.K. (2011). Protein delivery using engineered virus-like particles.

Kelly, R.M., and Strick, P.L. (2003). Cerebellar Loops with Motor Cortex and Prefrontal Cortex of a Nonhuman Primate. J. Neurosci. *23*, 8432–8444.

Kinnavane, L., Vann, S.D., Nelson, A.J.D., O'Mara, S.M., and Aggleton, J.P. (2018). Collateral Projections Innervate the Mammillary Bodies and Retrosplenial Cortex: A New Category of Hippocampal Cells. Eneuro *5*, ENEURO.0383-17.2018.

Knipe, D.M., and Cliffe, A. (2008). Chromatin control of herpes simplex virus lytic and latent infection. Nat. Rev. Microbiol. *6*, 211–221.

Kratchmarov, R., Taylor, M.P., and Enquist, L.W. (2012). Making the case : Married versus Separate models of alphaherpes virus anterograde transport in axons. Rev. Med. Virol. 378–391.

Kratchmarov, R., Kramer, T., Greco, T.M., Taylor, M.P., Ch'ng, T.H., Cristea, I.M., and Enquist, L.W. (2013b). Glycoproteins gE and gI are required for efficient KIF1A-dependent anterograde axonal transport of alphaherpesvirus particles in neurons. J. Virol. *87*, 9431–9440.

Kratchmarov, R., Taylor, M.P., and Enquist, L.W. (2013a). Role of Us9 phosphorylation in axonal sorting and anterograde transport of pseudorabies virus. PLoS One *8*, e58776.

Lamberti, C., and Weller, S.K. (1996). The Herpes simplex virus type 1 UL6 protein is essential for cleavage and packaging but not for genomic inversion. Virology *226*, 403–407.

Lo, L., and Anderson, D.J. (2011). A Cre-dependent, anterograde transsynaptic viral tracer for mapping output pathways of genetically marked neurons. Neuron *72*, 938–950.

Mador, N., Goldenberg, D., Cohen, O., Panet, A., and Steiner, I. (1998). Herpes simplex virus type 1 latency-associated transcripts suppress viral replication and reduce immediate-early gene mRNA levels in a neuronal cell line. J. Virol. *72*, 5067–5075.

Marshel, J.H., Mori, T., Nielsen, K.J., and Callaway, E.M. (2010). Targeting single neuronal networks for gene expression and cell labeling in vivo. Neuron *67*, 562–574.

Martersteck, E.M., Hirokawa, K.E., Evarts, M., Bernard, A., Duan, X., Li, Y., Ng, L., Oh, S.W., Ouellette, B., Royall, J.J., et al. (2017). Diverse Central Projection Patterns of Retinal Ganglion Cells. Cell Rep. *18*, 2058–2072.

McGovern, A.E., Davis-Poynter, N., Rakoczy, J., Phipps, S., Simmons, D.G., and Mazzone, S.B. (2012a). Anterograde neuronal circuit tracing using a genetically modified herpes simplex virus expressing EGFP. J. Neurosci. Methods *209*, 158–167.

McGovern, A.E., Davis-Poynter, N., Farrell, M.J., and Mazzone, S.B. (2012b). Transneuronal tracing of airways-related sensory circuitry using herpes simplex virus 1, strain H129. Neuroscience *207*, 148–166.

McGovern, A.E., Driessen, A.K., Simmons, D.G., Powell, J., Davis-Poynter, N., Farrell, M.J., and Mazzone, S.B. (2015). Distinct Brainstem and Forebrain Circuits Receiving Tracheal Sensory Neuron Inputs Revealed Using a Novel Conditional Anterograde Transsynaptic Viral Tracing System. J. Neurosci. *35*, 7041–7055.

Nellissery, J.K., Szczepaniak, R., Lamberti, C., and Weller, S.K. (2007). A Putative Leucine Zipper within the Herpes Simplex Virus Type 1 UL6 Protein Is Required for Portal Ring Formation. J. Virol. *81*, 8868–8877.

Newcomb, W.W., Homa, F.L., Thomsen, D.R., Booy, F.P., Trus, B.L., Steven, a C., Spencer, J. V, and Brown, J.C. (1996). Assembly of the herpes simplex virus capsid: characterization of intermediates observed during cell-free capsid formation. J. Mol. Biol. *263*, 432–446.

Newcomb, W.W., Homa, F.L., and Brown, J.C. (2005). Involvement of the Portal at an Early Step in Herpes Simplex Virus Capsid Assembly Involvement of the Portal at an Early Step in Herpes Simplex Virus Capsid Assembly. *79*, 10540–10546.

Norgren, R.B., McLean, J.H., Curt Bubel, H., Wander, A., Bernstein, D.I., and Lehman, M.N. (1992). Anterograde transport of HSV-1 and HSV-2 in the visual system. Brain Res. Bull. *28*, 393–399.

Padilla, S.L., Qiu, J., Soden, M.E., Sanz, E., Nestor, C.C., Barker, F.D., Quintana, A., Zweifel, L.S., Rønnekleiv, O.K., Kelly, M.J., et al. (2016). Agouti-related peptide neural circuits mediate adaptive behaviors in the starved state. Nat. Neurosci. *19*, 734–741.

Pomeranz, L.E., Reynolds, A.E., Christoph, J., and Hengartner, C.J. (2005). Molecular Biology of Pseudorabies Virus : Impact on Neurovirology and Veterinary Medicine Molecular Biology of Pseudorabies Virus : Impact on Neurovirology and Veterinary Medicine. Society *69*, 462–500.

Rinaman, L., and Schwartz, G. (2004). Anterograde transneuronal viral tracing of central viscerosensory pathways in rats. J. Neurosci. *24*, 2782–2786.

Rostad, W., and Alvord, C. (2002). Transsynaptic Spread of Varicella Zoster Virus Through the Visual System. 1–6.

Roy, D.S., Kitamura, T., Okuyama, T., Ogawa, S.K., Sun, C., Obata, Y., Yoshiki, A., and Tonegawa, S. (2017). Distinct Neural Circuits for the Formation and Retrieval of Episodic Memories. Cell *170*, 1000–1012.e19.

Ryu, V., Garretson, J.T., Liu, Y., Vaughan, C.H., and Bartness, T.J. (2015). Brown Adipose Tissue Has Sympathetic-Sensory Feedback Circuits. J. Neurosci. *35*, 2181–2190.

Ryu, V., Watts, A.G., Xue, B., and Bartness, T.J. (2017). Bidirectional crosstalk between the sensory and sympathetic motor systems innervating brown and white adipose tissue in male Siberian hamsters. Am. J. Physiol. - Regul. Integr. Comp. Physiol. *312*, R324–R337.

Sawtell, N.M., and Thompson, R.L. (2014). Herpes Simplex Virus Mutant Generation and Dual-Detection Methods for Gaining Insight into Latent/Lytic Cycles in Vivo. *1144*, 129–147.

Scott, N., Prigge, M., Yizhar, O., and Kimchi, T. (2015). A sexually dimorphic hypothalamic circuit controls maternal care and oxytocin secretion. Nature *525*, 519–522.

Shabel, S.J., Proulx, C.D., Trias, A., Murphy, R.T., and Malinow, R. (2012). Input to the Lateral Habenula from the Basal Ganglia Is Excitatory, Aversive, and Suppressed by Serotonin. Neuron *74*, 475–481.

Smith, G. (2012). Herpesvirus transport to the nervous system and back again. Annu. Rev. Microbiol. *66*, 153–176.

Smith, G. a, Gross, S.P., and Enquist, L.W. (2001). Herpesviruses use bidirectional fast-axonal transport to spread in sensory neurons. Proc. Natl. Acad. Sci. U. S. A. *98*, 3466–3470.

Song, C.K., Schwartz, G.J., and Bartness, T.J. (2009). Anterograde transneuronal viral tract tracing reveals central sensory circuits from white adipose tissue. Am J Physiol Regul Integr Comp Physiol *296*, 501–511.

Strack, A.M., Sawyer, W.B., Hughes, J.H., Platt, K.B., and Loewy, A.D. (1989). A general pattern of CNS innervation of the sympathetic outflow demonstrated by transneuronal pseudorabies viral infections. Brain Res. *491*, 156–162.

Sun, N., Cassell, M.D., and Perlman, S. (1996). Anterograde , transneuronal transport of herpes simplex virus type 1 strain H129 in the murine visual system . Anterograde , Transneuronal Transport of Herpes Simplex Virus Type 1 Strain H129 in the Murine Visual System. *70*.

Szilágyi, J.F., and Cunningham, C. (1991). Identification and characterization of a novel non-infectious herpes simplex virus-related particle. J. Gen. Virol. *72 (Pt 3)*, 661–668.

Szpara, M.L., Parsons, L., and Enquist, L.W. (2010). Sequence variability in clinical and laboratory isolates of herpes simplex virus 1 reveals new mutations. J. Virol. *84*, 5303–5313.

Tang, H., Wu, G.S., Xie, J., He, X., Deng, K., Wang, H., Xu, F., and Luo, H.R. (2016). Brain-wide map of projections from mice ventral subiculum. Neurosci. Lett. *629*, 171–179.

Umbach, J.L., Kramer, M.F., Jurak, I., Karnowski, H.W., Coen, D.M., and Cullen, B.R. (2008). MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. Nature *454*, 780–783.

Vaughan, C.H., and Bartness, T.J. (2012). Anterograde transneuronal viral tract tracing reveals central sensory circuits from brown fat and sensory denervation alters its thermogenic responses. Am. J. Physiol. Integr. Comp. Physiol. *302*, R1049–R1058.

Wallace, M.L., Saunders, A., Huang, K.W., Philson, A.C., Goldman, M., Macosko, E.Z., McCarroll, S.A., and Sabatini, B.L. (2017). Genetically Distinct Parallel Pathways in the Entopeduncular Nucleus for Limbic and Sensorimotor Output of the Basal Ganglia. Neuron *94*, 138–152.e5.

Wickersham, I.R., Lyon, D.C., Barnard, R.J.O., Mori, T., Finke, S., Conzelmann, K.K., Young, J.A.T., and Callaway, E.M. (2007). Monosynaptic Restriction of Transsynaptic Tracing from Single, Genetically Targeted Neurons. Neuron.

Witter, M.P. (2006). Connections of the subiculum of the rat: Topography in relation to columnar and laminar organization. Behav. Brain Res. *174*, 251–264.

Wojaczynski, G.J., Engel, E. a, Steren, K.E., Enquist, L.W., and Patrick Card, J. (2014). The neuroinvasive profiles of H129 (herpes simplex virus type 1) recombinants with putative anterograde-only transneuronal spread properties. Brain Struct. Funct. *129*.

Zeng, W.B., Jiang, H.F., Gang, Y.D., Song, Y.G., Shen, Z.Z., Yang, H., Dong, X., Tian, Y.L., Ni, R.J., Liu, Y., et al. (2017). Anterograde monosynaptic transneuronal tracers derived from herpes simplex virus 1 strain H129. Mol. Neurodegener. *12*, 1–17.

Zingg, B., Chou, X. lin, Zhang, Z. gang, Mesik, L., Liang, F., Tao, H.W., and Zhang, L.I. (2017). AAV-Mediated Anterograde Transsynaptic Tagging: Mapping Corticocollicular Input-Defined Neural Pathways for Defense Behaviors. Neuron *93*, 33–47.

Table 1.1: Guide RNA sequences used for removal of the GFP expression cassette from the *UL6* locus for functional deletion of UL6

Targeting the 3' region of UL6	Targeting the 5' region of UL6
TAATCACTATTAGACGGCCG	ACCCTCCCGGGATTCGGCAA
CACCTTGTTAACGACCCGCG	GCGCCTGCGCGAGACCAACG
AGACGGTCCCCGGTCCACAA	GACATACGTGTCGTCGTCCA

Figure 1.1: Functional deletion of HSV UL6 in H129 LSL-TK-TT prohibits viral spread *in vitro* and *in vivo*

(A) UL6 deletion scheme. The UL6 locus (top) was targeted with a GFP expression cassette (middle) flanked by regions homologous to the UL6 gene, resulting in the virus HSV H129 LSL-TK-TT ΔUL6-GFP. Following purification, targeted removal of the GFP expression cassette by CRISPR/Cas9 resulted in a frameshift after 377 amino acids of the UL6 coding sequence and premature UL6 termination (bottom). (B) Insertion of the GFP expression cassette prohibits spread of the virus. Application of H129 LSL-TK-TT ΔUL6-GFP to monolayers of vero cells, vero cells constitutively expressing UL6 (31-UL6), and vero cells transfected prior in inoculation with the plasmid pNef-coUL6. (C & D) Comparative injections of H129 LSL-TK-TT and H129 LSL-TK-TT ΔUL6 reveal restriction of spread in the ΔUL6 virus. (C) Brain tissue collected 5-days after intravitreal injection of H129 LSL-TK-TT (left column) and HSV H129 LSL-TK-TT ΔUL6 (right column) in the right eye of Slc17a6-cre (vglut2) mice. (D) Brain tissue collected 5-days after injection of HSV H129 LSL-TK-TT (left column) and HSV H129 LSL-TK-TT ΔUL6 (right column) in primary visual cortex of Sim1-cre mice. Anti-HSV antibody staining in magenta, DAPI in blue.

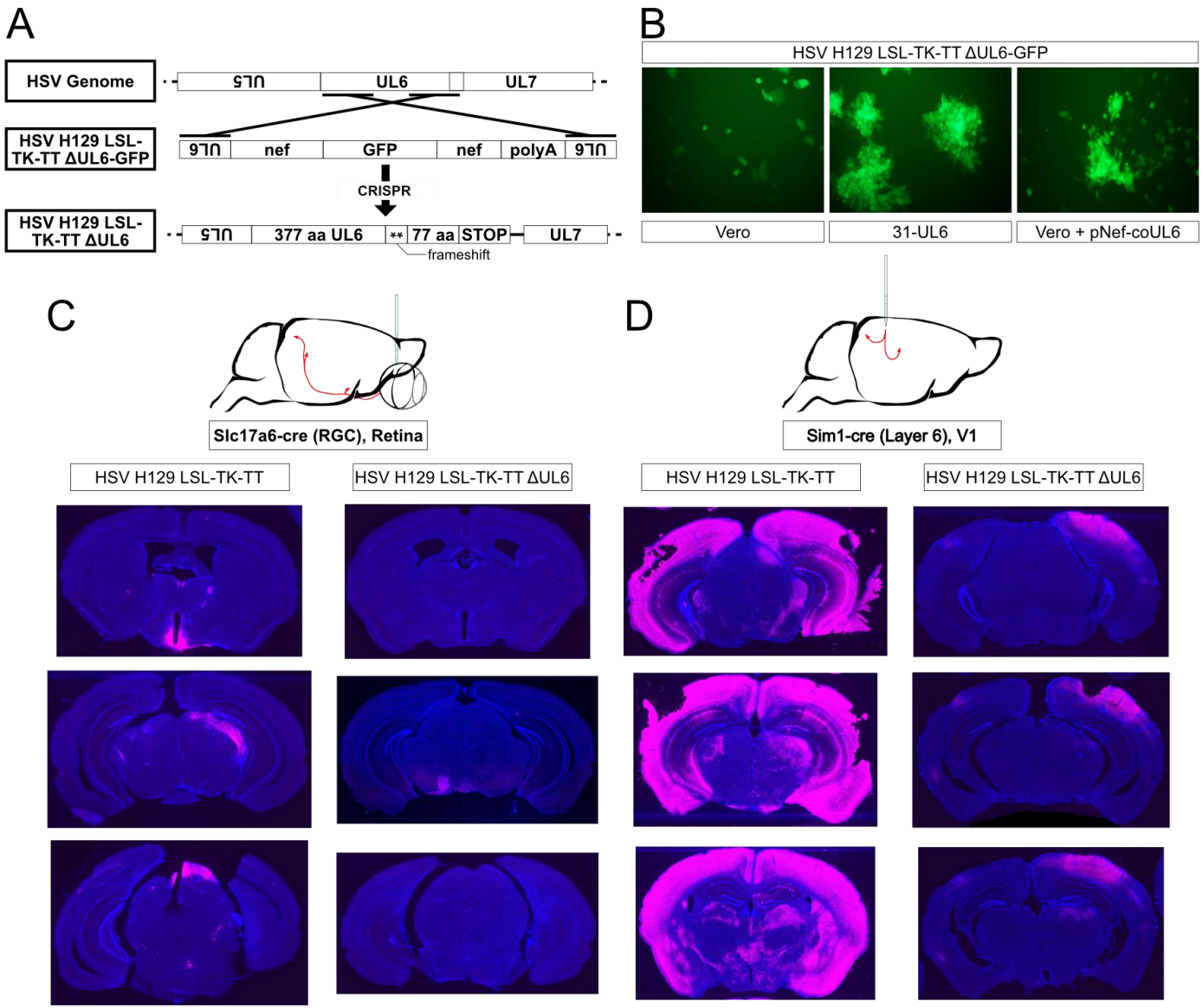

A
HSV Genome
UL5
UL6
UL7
HSV H129 LSL-TK-TT ΔUL6-GFP
UL5
nef
GFP
nef
polyA
UL5
CRISPR
HSV H129 LSL-TK-TT ΔUL6
UL5
377 aa UL6
**
77 aa
STOP
UL7
frameshift
B
HSV H129 LSL-TK-TT ΔUL6-GFP
Vero
31-UL6
Vero + pNef-coUL6
C
Slc17a6-cre (RGC), Retina
HSV H129 LSL-TK-TT
HSV H129 LSL-TK-TT ΔUL6
D
Sim1-cre (Layer 6), V1
HSV H129 LSL-TK-TT
HSV H129 LSL-TK-TT ΔUL6

Figure 1.2: Recombination across *loxp* sites present in in HSV and AAV leads to cre-dependent integration of the GFP-p2a-coUL6 expression cassette from AAV in the H129 LSL-TK-TT ΔUL6 virus

(A) Injections of H129 LSL-TK-TT ΔUL6 following AAV8-nef-DIO-GFP-p2a-coUL6 in primary visual cortex of Ntsr1-cre mice leads to GFP+ cells outside of layer 5. Injection of AAV8-nef-DIO-GFP-p2a-coUL6 alone (left) results in restricted expression of GFP in the target layer 5 population. A similar injection followed two weeks later (right) shows tdTomato+ and GFP+ cells outside of layer 5. (B) pAAV-nef-DIO-GFP-p2a-coUL6 design (top) and table of tested orthogonal *lox* combinations. "2P" is the standard DIO configuration. (C) Vero cell monolayers 48hrs after application of H129 LSL-TK-TT ΔUL6 at a MOI 1. Transfection parameters are shown above each column. pAAV test conditions in the right 3 columns. (D) Vero cell monolayers 48hrs after application of 1 ml of filtered media from each condition in (C).

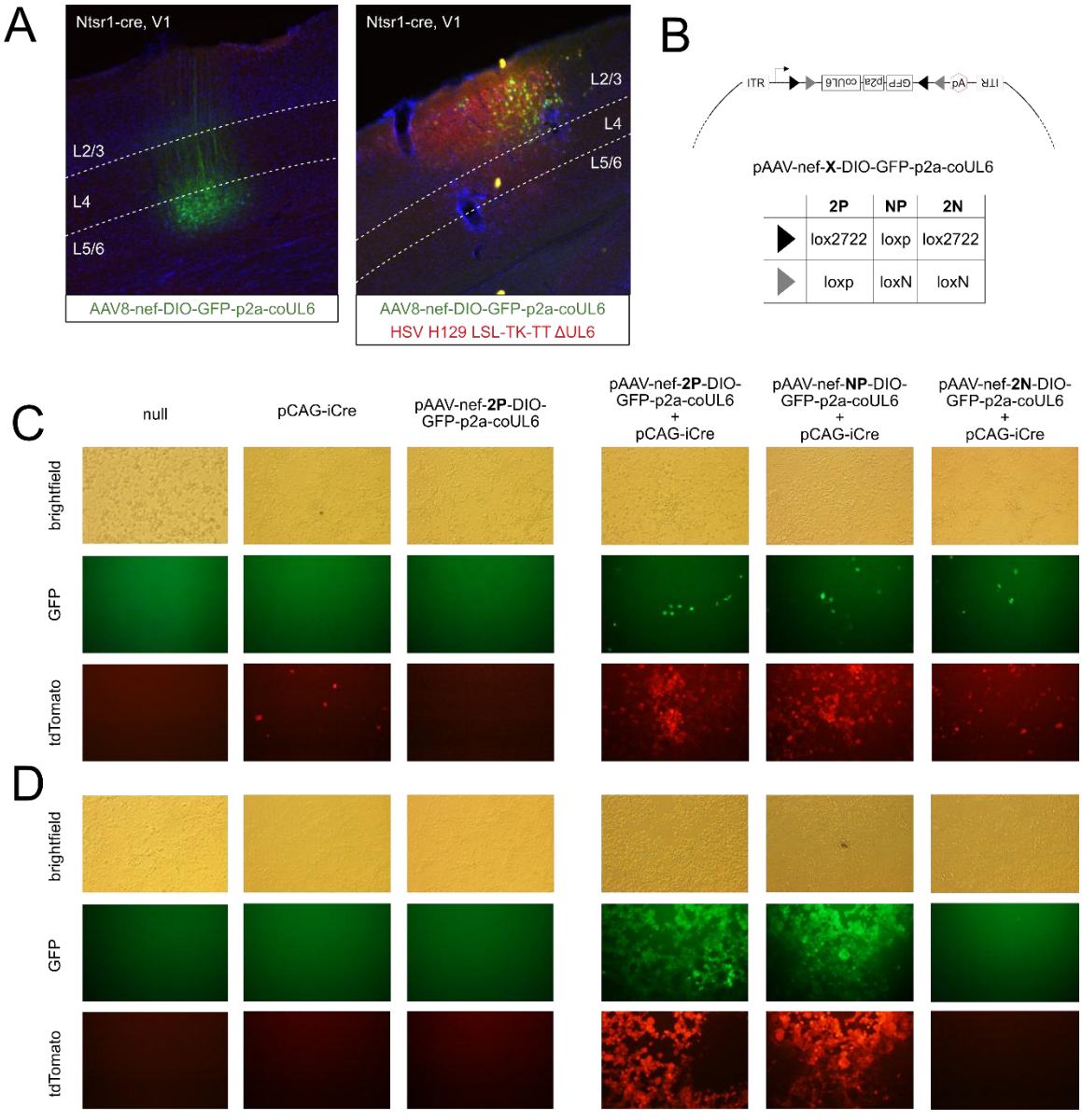

A
Ntsr1-cre, V1
L2/3
L4
L5/6
AAV8-nef-DIO-GFP-p2a-coUL6
Ntsr1-cre, V1
L2/3
L4
L5/6
AAV8-nef-DIO-GFP-p2a-coUL6
HSV H129 LSL-TK-TT ΔUL6
B
ITR
coUL6
p2a
GFP
pA
ITR
pAAV-nef-X-DIO-GFP-p2a-coUL6
2P
NP
2N
lox2722
loxp
lox2722
loxp
loxN
loxN
C
null
pCAG-iCre
pAAV-nef-2P-DIO-GFP-p2a-coUL6
pAAV-nef-2P-DIO-GFP-p2a-coUL6 + pCAG-iCre
pAAV-nef-NP-DIO-GFP-p2a-coUL6 + pCAG-iCre
pAAV-nef-2N-DIO-GFP-p2a-coUL6 + pCAG-iCre
brightfield
GFP
tdTomato
D
brightfield
GFP
tdTomato

Figure 1.3: Transcomplementation of H129 LSL-TK-TT ΔUL6 with AAV in retinal ganglion cells results in HSV labeled neurons in lateral geniculate nucleus and superior colliculus

(A) Flat-mount retina in the right eye of a Slc17a6-cre mouse after intravitreal injection of H129 LSL-TK-TT ΔUL6. Inset shows dotted-box region of injection site. Anti-HSV antibody in magenta. Brain slices from the same mouse showing visual thalamic nuclei (B) and superior colliculus (C). DAPI in blue, anti-HSV antibody in magenta. Each channel separated below. (D) Flat-mount retina following injection of AAV8-nef-2N-DIO-GFP-p2a-coUL6 and H129 LSL-TK-TT ΔUL6 in GPi of Som-ires-cre mouse. GFP in green, tdTomato in red. Inset show dotted-box region contain double labeled neurons. Brain slices from the same mouse show HSV positive neurons amongst GFP positive axons originating from retina in lateral geniculate nucleus (E) and superior colliculus (F). GFP in green, anti-HSV antibody in magenta, DAPI in blue. Center column shows split channels. Inset shows example HSV infected neuron within the dotted box. All tissue collected 7 days after HSV injection. dLG, dorsal lateral geniculate nucleus; vLG, ventral geniculate nucleus; LP, lateral posterior nucleus of the thalamus; SCsg, superior colliculus superficial gray layer; SCm, superior colliculus, motor related.

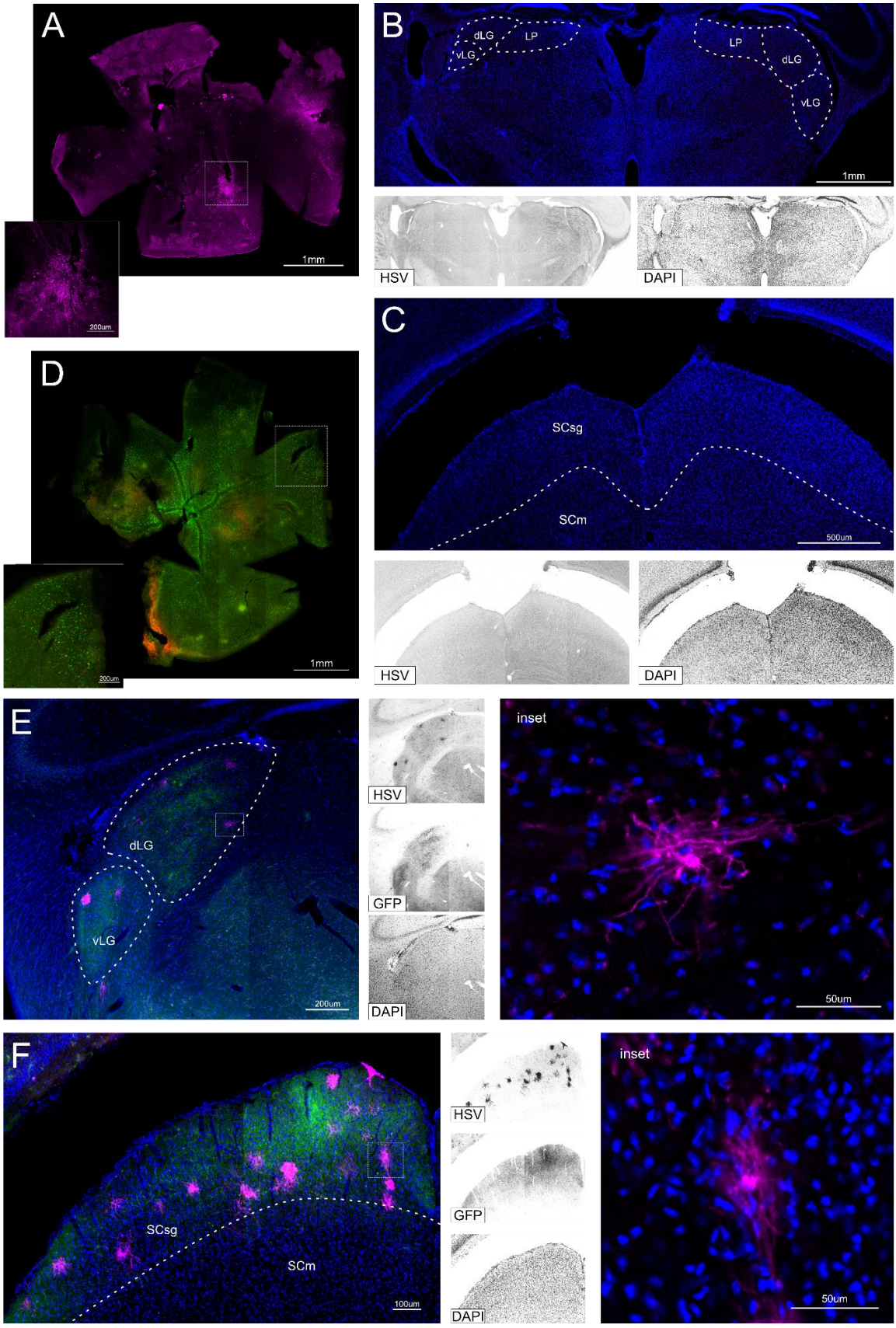

A
1mm
200um
B
dLG
LP
vLG
LP
dLG
vLG
1mm
HSV
DAPI
C
SCsg
SCm
500um
HSV
DAPI
D
1mm
200um
E
dLG
vLG
200um
HSV
GFP
DAPI
inset
50um
F
SCsg
SCm
100um
HSV
GFP
DAPI
inset
50um

Figure 1.4: Transcomplementation of H129 LSL-TK-TT ΔUL6 with AAV in somatostatin neurons of the internal segment of the globus pallidus results in HSV labeled neurons in lateral habenula

(A) Injection site of H129 LSL-TK-TT ΔUL6 in the internal segment of globus pallidus in an Som-ires-cre mouse. (B) Habenula from the same mouse. Channels split at right. (C) Injection site of AAV8-nef-2N-DIO-GFP-p2a-coUL6 and H129 LSL-TK-TT ΔUL6 in the internal segment of globus pallidus in a Som-ires-cre mouse. (D) HSV positive neurons amongst GFP positive axons originating from globus pallidus internal segment in lateral habenula. Channels split at right. GFP in green, anti-HSV antibody in magenta, DAPI in blue. All tissue collected 7 days following HSV injection. GPe, globus pallidus external segment; GPi, globus pallidus internal segment; int, internal capsule; RT, reticular thalamus; AMY, amygdalar nuclei; MH, medial habenula; LH, lateral habenula.

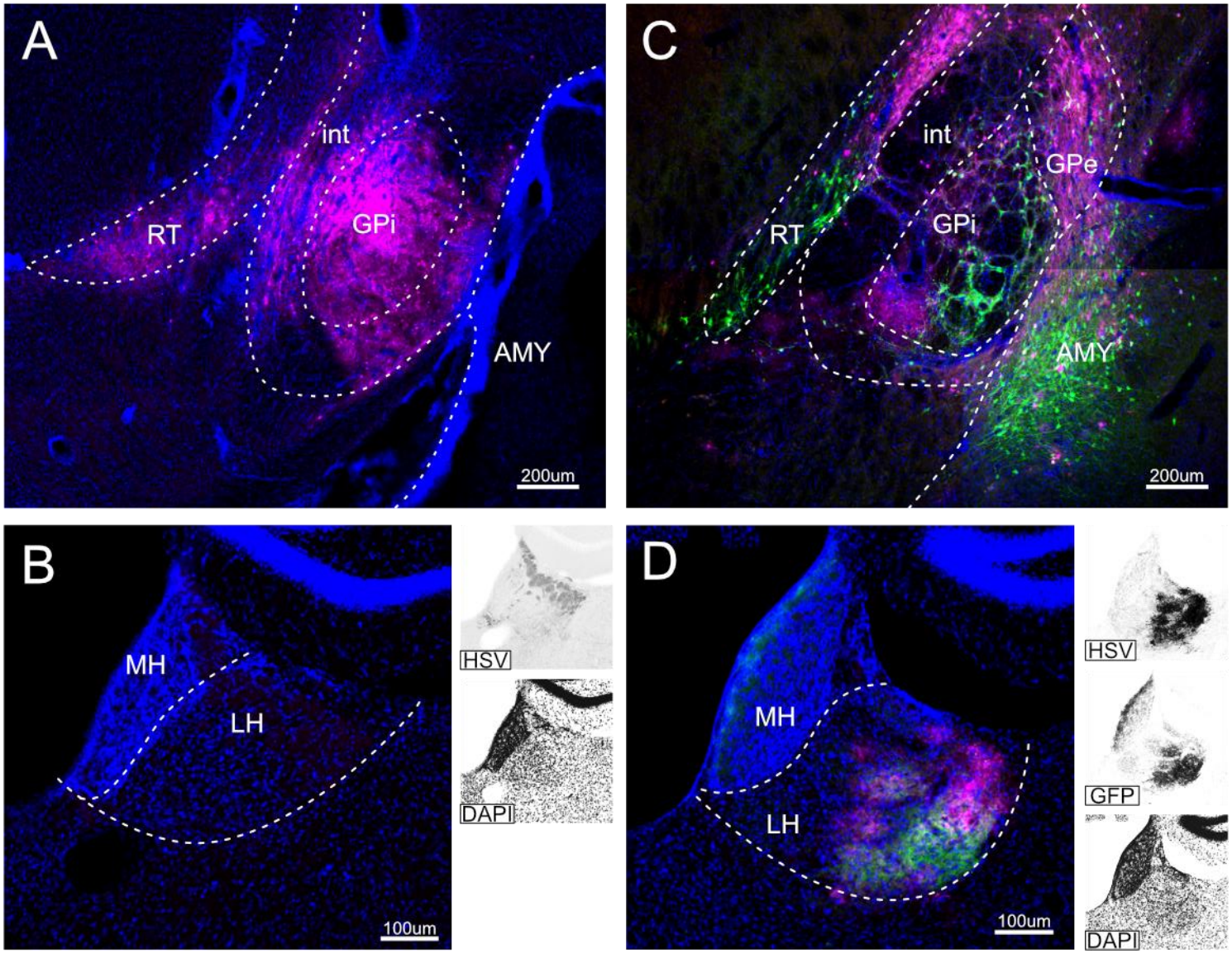
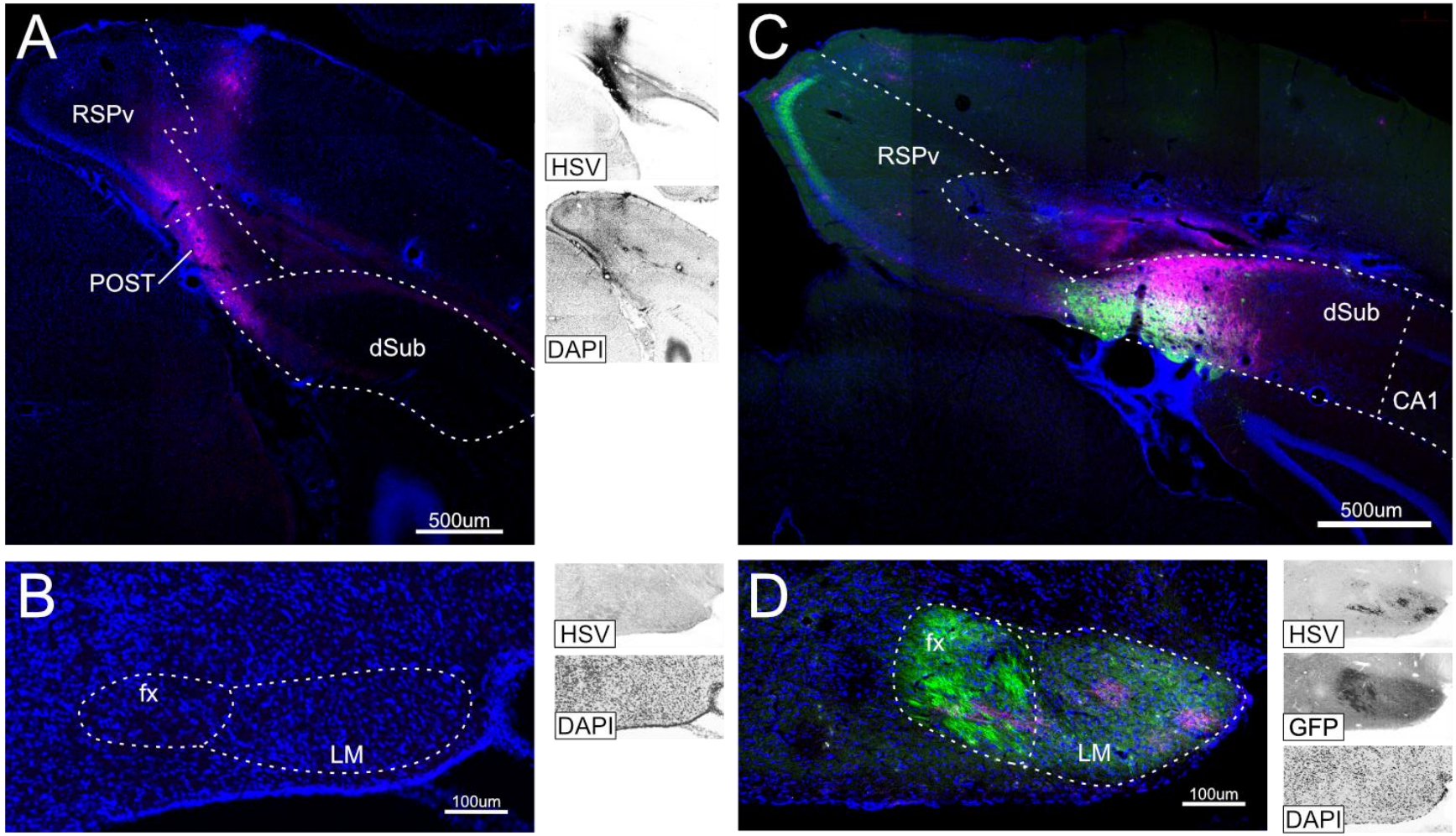

A
int
RT
GPi
AMY
200um
C
int
GPe
RT
GPi
AMY
200um
B
MH
LH
100um
HSV
DAPI
D
MH
LH
100um
HSV
GFP
DAPI

Figure 1.5: Transcomplementation of H129 LSL-TK-TT ΔUL6 with AAV in dorsal Subiculum vglut2 neurons results in HSV labeled neurons in Mammillary Body

(A) Injection site of an H129 LSL-TK-TT injection targeting dorsal subiculum in a Slc17a6-cre mouse. (B) Columns of the fornix and mammilary body from the same mouse. Split channels at right. (C) Injection site of AAV8-nef-2N-DIO-GFP-p2a-coUL6 and H129 LSL-TK-TT ΔUL6 injections targeting dorsal subiculum in a Slc17a6-cre mouse. (B) HSV positive neurons amongst GFP positive axons originating from dSub in the mammilary body. GFP in green, anti-HSV antibody in magenta, DAPI in blue. Right column shows split channels. RSPv, ventral retrosplenial cortex; POST, postsubiculum; dSub, dorsal subiculum; fx, columns of the fornix; LM, lateral mammillary body.


A
RSPv
POST
dSub
500um
HSV
DAPI
C
RSPv
dSub
CA1
500um
B
fx
LM
100um
HSV
DAPI
D
fx
LM
100um
HSV
GFP
DAPI

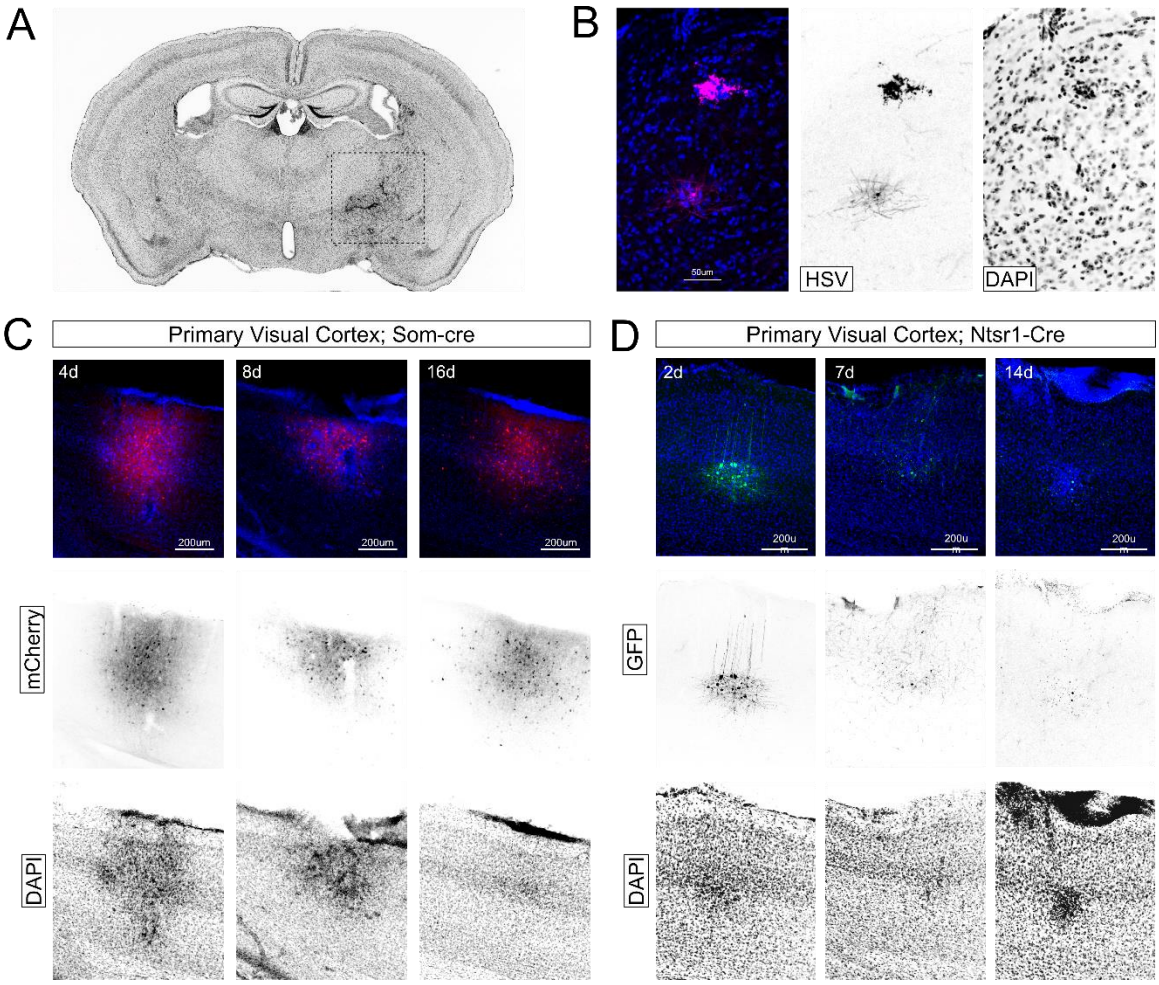


Figure 1.6: Cre-activation of H129 LSL-FP ΔUL6 leads to rapid transgene expression and cell death

(A) Whole brain slice showing condensed DAPI staining at the injection site (dotted box). 7 days post injection of H129 LSL-TK-TT ΔUL6, 21 days post injection of AAV8 nef-2N-DIO-GFP-2A-coUL6. Injection target location globus pallidus internal segment. (B) Example HSV labeled neurons present in LGN after transcomplementation in retina of Slc17a6-cre mouse. Anti-HSV antibody in magenta, DAPI in blue. Tissue collected 7 days after injection. Channels split at right. (C) Primary visual cortex tissue following injection of H129 LSL-TK-GFP ΔUL6-mcherry in Som-cre mice. Tissue collected at 4, 8, and 16 days post injection. GFP channel not shown. mCherry in red, DAPI in blue. Split channels below. (D) Primary visual cortex tissue following injection of H129 LSL-TK-GFP ΔUL6-mCherry in Ntsr1-cre mice. Tissue collected 2, 7, and 14 days post injection. mCherry channel not shown. DAPI in blue, GFP in green. Split channels below.

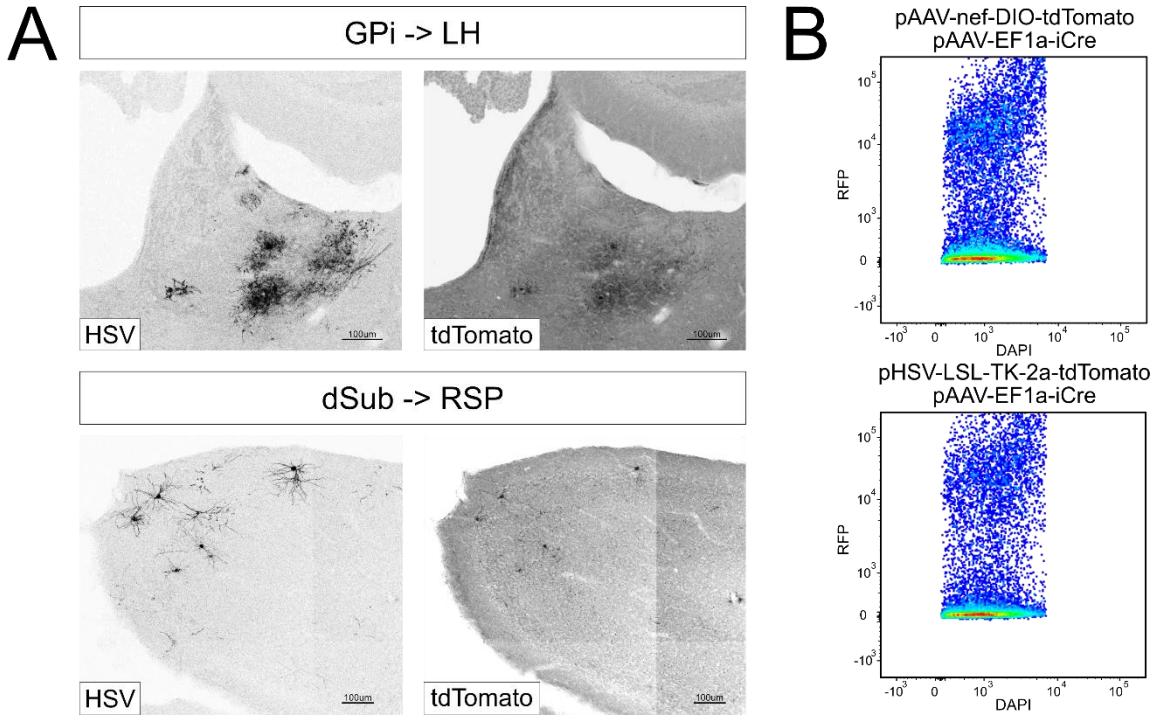


Figure 1.7: Sporadic and weak expression of tdTomato from the H129 LSL-TK-TT ΔUL6 virus

Examples of tdTomato expression from the cre-activated H129 LSL-TK-TT ΔUL6 virus show (A) weak expression in HSV antibody labeled neurons in lateral habenula (LH) from transcomplementation experiments targeting somatostatin neurons of the internal segment of glubus pallidus (GPi), and stronger expression of tdtomato (B) in HSV antibody labeled neurons of the retrosplenial cortex (RSP) from transcomplementation experiments targeting vglut2 neurons of the dorsal subiculum (dSub). All tissue collected 7 days after HSV injection. No antibody was used against tdTomato in these preparations. (C) Flow cytometry dot plots of HEK 293T cells transfected with the indicated plasmids.

Chapter 2: Spontaneous recombination across recombinase recognition sites contributes to off-target expression from common recombinase-dependent AAVs

Abstract

Dissection of neural circuits requires technologies allowing for precise and reliable transgene expression in specific neuronal cell types within specific brain regions. In combination with transgenic mouse lines expressing the recombinases cre or flp in molecularly defined cell types, recombinase-dependent adeno-associated viruses (AAVs) have become the tool of choice for localized cell-type-targeted expression of fluorescent proteins, calcium indicators, channelrhodopsins, rabies transsynaptic tracing components, and more. Unfortunately, application of this technique when expressing highly sensitive transgenes (ie, gene products that are directly or indirectly detectable at very low copy numbers) are impeded by apparently off-target, or "leak", expression from recombinase-dependent AAVs. Surprisingly, this leak persists when the transgene open reading frame (ORF) is disrupted such that it is only reconstituted following recombinase-mediated recombination. We investigated this phenomenon and find that while leak expression is indeed mediated by infrequent transcription from the inverted transgene in recombinant-dependent AAV designs, the persistent leak is due to recombination events across anti-parallel recombinase-specific recognition sites during AAV plasmid production in a process that has been previous described. Finally, we propose a solution to this issue by blocking the packaging of the AAV genomes originating from the offending recombinant plasmids by modification of the D-sequence of a single inverted terminal repeat (ITR). Beyond providing for more reliable and targeted transgene expression, and the expansion

of technologies requiring such precision, the underlying cause of leak described herein should be considered when designing AAVs and the experiments in which they are applied.

Introduction

Site-specific recombinases, of which cre and flp are members, resolve Holliday junctions formed between recombinase-specific 34 base pair sequences comprised of an eight base pair spacer sequence flanked by 13 base pair inverted repeat (Hoess et al., 1982; Lee and Saito, 1998; Voziyanov et al., 1999). Besides the canonical *loxp* and *frt* sites recognized by cre and flp, respectively, numerous orthogonal sites containing small mutations to the spacer region have also been described (Albert et al., 1995; Arakil et al., 1997; Schlake and Bode, 1994; Senecoff et al., 1988). Depending on whether pairs of these recognition sites are in a parallel or antiparallel orientation, the recombinase will cause excision or inversion, respectively, of the intervening DNA. The ability to manipulate DNA in such a manner has led to diverse applications, including the development of many transgenic mouse lines expressing recombinases in a restricted neuronal cell types (Taniguchi et al., 2011). While these lines can be crossed with recombinase reporter lines to drive transgene expression in a target population globally (Madisen et al., 2010), recombinase-dependent AAVs allow for targeting of specific regions and expression of different transgenes without the comparatively cumbersome process of transgenic mouse line production.

AAV is a small relatively non-toxic unenveloped virus of with a roughly 5kb single stranded DNA genome bookended by 145bp GC-rich inverted terminal repeats (ITRs). These single stranded ITRs form cruciform hairpins necessary for replication, packaging, episomal concatermization of viral genomes, and host genome integration (Chiorini et al., 1994; Ling et al., 2015; Yang et al., 1997, 1999). While recombinase-dependent AAV designs using the *lox*-STOP-*lox* and *frt*-STOP-*frt* system have been used (Kuhlman and

Huang, 2008), Double-Inverted Open reading frame (DIO) and Flip/Excision (FLEX) constructs (Cardin et al., 2009; Schnütgen et al., 2003), practically identical in their design, have gained the most widespread use for their limited size and purported less leaky nature when using strong promoters (Huang et al., 2014). Briefly, the DIO and FLEX designs use two pairs of orthogonal recognition sites in an overlapping antiparallel orientation around the desired transgene that is, in respect to the rest of the expression cassette, inverted and thus transcriptionally repressed. When exposed to the appropriate recombinase, the transgene ORF is reverted and locked in-sense with the promoter and 3' UTR, driving expression.

Our lab and others have noted evidence of transgene expression in off-target cell populations, termed "leak", from DIO and FLEX constructs expressing sensitive transgenes, such as TVA for targeting EnvA-pseudotyped rabies virus to specific neural cell types (Wall et al., 2010), of which small quantities are sufficient for EnvA binding and viral entry (Seidler et al., 2008). As this leak can confound or prohibit the use of these tools for certain applications, including rabies tracing of local circuits, efforts have been made to circumvent the issue, such as a TVA with decreased sensitivity (Miyamichi et al., 2013; Rong et al., 1998). Another, sometimes called "ATG-out" or "split-TVA" and working on the assumption that low-level expression occurs from the inverted ORF, disrupts the ORF in DIO or FLEX constructs such that the kozak sequence and initiating codon of the transgene are placed outside the first set of recombinase recognition sites, leaving the transgene ORF to be reconstituted following recombination (Kohara et al., 2014; Wall et al., 2010). While both designs decrease leak expression, our lab has found that neither abolish it entirely.

AAV production requires copious quantities of high concentration plasmid containing the recombinant genome. Due to the inverted homology of the ITRs, it is recommended that AAV plasmids are prepared in commercially available *E. coli* containing mutations to the *RecA* or *RecBJ* genes (Jungmann et al., 2017), part of the RecBCD pathway mediating bacterial homologous recombination (Dillingham and Kowalczykowski, 2008). However, previous work has shown that rates of recombination across short homologous regions less than 50 base pairs in length are unaffected by *RecA* mutation and likely occur due to strand breaks at the replication fork (Bi and Liu, 1996; Lovett et al., 2002). Further, recombination resulting from homology across short inverted repeats spanning an inverted tetracycline resistance gene resulted in plasmid doubling events that reconstituted transcription and imparted resistance (Bi and Liu, 1996).

Given the short size of recombinase-specific sites, the fact that ORF disruption fails to abrogate leak expression, and the previous literature finding that recombination across inverted repeats can reconstitute an inverted transgene, we hypothesized that spontaneous recombinase-independent inversion occurs during the AAV production process. Our investigation found recombinants in all AAV preparations and AAV plasmids we tested and confirmed that intra- or interplasmid recombination indeed occurs across recombinase sites in *RecA*-mutant competent *E coli*. In addition, we show that while this recombination occurs at a low rate in most AAV plasmids, its nature is such that some AAV preparations could be largely or entirely comprised of recombinants. By independently disrupting spontaneous inversion and the transgene ORF, we show that both must be disrupted to fully abrogate leak. Further, while leak expression from an intact ORF is only detectable in

highly sensitive systems, spontaneous inversions can drive low but detectable levels of expression of fluorescent proteins.

Finally, we implemented a small change in AAV plasmid to block packaging of the offending recombinant plasmid in an attempt to produce “leakless” recombinase-sensitive AAV preparations. Recent work has shown that a 20nt modifications to the D-sequence of one of either AAV ITRs results in selective packaging of specific senses of the AAV ssDNA genome without disrupting replication (Ling et al., 2015; Wang et al., 1996, 1997). Given the mirrored nature of the plasmids resulting from the offending recombination events, we implement a single-ended D-sequence deletion design that impinges minor costs on packaging of unaffected genomes while blocking packaging of genome containing reverted transgenes. While this strategy worked to block the offending genome orientation, the resulting AAV preparations had no effect on leak expression, suggesting that a further mechanism of leaky transcription that is specific to AAV.

Materials and Methods

Plasmid Production

NEB Stable cells (NEB C3040) were used for DNA preparations and all blue/white assays. Chemical transformations were followed as described by the manufacturer. All minipreps were grown in a shaker at 220rpm in 2mls of LB containing 0.1mg/ml Ampicillin at 30C for 20hr, unless otherwise noted.

Blue/White Screening

To produce minipreps of DNA, 2 mls of LB containing Ampicillin were inoculated with a single colony and grown at 30C for 12hrs. The bacterial solution was the diluted to 10 cfu/ul from an estimated starting concentration of 10^6 cfu/ul. 50ul of this dilution was plated on 100ug/ml carbenicillin plates made with Chromomax IPTG/X-Gal solution (Fisher BP4200). Three plates were grown for each miniprep. Plates were grown at 37C overnight and moved to 4C before being imaged with a gel imager. For colony counting, a matlab script was used to identify blue and white colonies in each image before manual correction and addition of missed colonies.

Plasmids and Cloning

The blue/white screening plasmid pUC19-X-inv_LacZ was made by restricting pUC19 (NEB N3041) with BsmBI, EcoRI, and HindIII and collected both the backbone and LacZ fragments. Recombinase specific and shuffled 34bp LacZ flanking sites were designed as oligos to be annealed such that the final product contained overhangs matching

those left by the restriction sites. Sequences of these 34bp sites can be found in Figure 4B. To anneal, oligos were resuspended in molecular grade water to 100uM and 2ul of each oligo solution was added to the reaction mix with 1ul of T4 Polynucleotide Kinase (NEB M0201), 2ul of 10x T4 PNK Buffer (NEB BS0201) and 13ul of molecular water. The PNK reaction was conducted as follows: 37C for 10 minutes, 65C for 10 minutes, 95C for 30 seconds, 60C for 10 minutes, and cooling at 0.1C/second to 20C. The reaction solution was diluted 20-fold and 1ul of this annealed oligo solution was added to each T4 ligation along with the pUC19 backbone and LacZ fragment. Shuffled *loxp* sequences used in both model pUC19 and pAAV plasmids were generated by randomly shuffling the *loxp* sequence, choosing sequences without more than 3 consecutive nucleotide repeats, and tested for homology using ThermoFisher's Multiple Primer Analyzer. All plasmids were sequenced by Eton across the cloned regions prior to use.

For disruption of the D-sequence, an extant nef-ATG-out-DIO expression cassette was cloned into the AAV backbone of the pLC1 plasmid (Ling et al., 2015), an AAV plasmid with a substituted D-sequence in the 5' ITR, which was generously provided by Dr. Arun Srivastava.

AAV PCR

AAV preparations were treated with DNAse1 (epicenter D9905K) at a final concentration of 0.1U/ul for 10 minutes at 37C. The solution was then immediately heated to 98C for 10 minutes to denature the DNAseI and capsid proteins. 1ul of this solution was used as a PCR template in a standard 50ul PCR reaction.

qPCR

For pUC19 experiments, the forward and reverse primers for the lacz amplicon used to measure total copy number were GCCTCTTCGCTATTACGCCA and TCACTGGCCGTCGTTTTACA, respectively. Forward and reverse primers used to assay for recombinant plasmids were GTGCTGCAAGGCGATTAAGT and TGTGGAATTGTGAGCGGATA respectively. Primer3 was used to design all qPCR primers. All qPCR reactions were carried out with SYBR Green (ThermoScientific K0251) on a Lightcycler 480II for 45 cycles with an annealing temperature of 60C.

qPCR standards were 500ng 500bp gBlocks made by IDTDNA and based upon the expected plasmid sequence of following recombination. Separate gBlocks were used for *loxp* and *frt* based sequences. These sequences contained templates for both the lacZ and recombinant amplicons.

loxp_template:

GTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTAC
TGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAA
AATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGG
GCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGT
GCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT
GTAAAACGACGGCCAGTGAATTCGATAACTTCGTATAATGTATGCTATACGA
AGTTATCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGT
TATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAG
CCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTG
CCCGCTTTCCAGTCGGGAAACCTGTCGTGCC

frt_template:

GTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTAC
TGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAA
AATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGG
GCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGT
GCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT
GTAAAACGACGGCCAGTGAATTCGAAGTTCCTATACTTTCTAGAGAATAGGA
ACTTCGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGT
TATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAG

CCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTG
CCCGCTTTCCAGTCGGGAAACCTGTCGTGCC

Mammalian Cells and Transfections

All cells were grown in 10% Fetal Bovine Serum (FBS; HyClone SH 30070.03) DMEM media (Gibco 11995-040) at 37C and 5% CO_2. Cre-reporter HEK cells were made by lentiviral transfection of a CAG-loxp-TdTomato-loxp-GFP expression cassette and selection of the top 1% of RFP+ cells. However, potentially due to multiple integrated copies per genome of the selected cells, these cells rarely switch from green to red as the expression cassette suggests. For each ug of DNA to be transfected, a 3x volume of 1mg/ml PEI was added to a 6x volume of Opti-MEM (Gibco 51985-043) and the appropriate volume of DNA was added to the same 6x volume of Opti-MEM. These two solutions were mixed and left at room temperature for 15 minutes, after which they were applied dropwise to the media of each target cell plate or well. After 4 to 6 hours, the media was replaced.

Flow Cytometry

Forty-eight hours following transfection, cells were trypsinated, collected in 2ml Eppendorf tubes, and spun at 500 rcf for 3 minutes. The media was aspirated and the cellular pellet resuspended in 1ml of cold 1% FBS in dPBS (Gibco 14190-144). The cells were spun down again, the media aspirated, and cells resuspended in 1ml of cold 10um DAPI in 1% FBS in dPBS. Cells were spun down, the media aspirated, and cells resuspended in 500ul of cold 1% FBS in dPBS. This solution was moved to filter-topped tubes (Falcon 352235) on ice and run on a LSRII in the Salk Flo Cytometry Core Facility.

Methods Section Analysis

Text of methods sections from Neuron articles were obtained by a Python script utilizing the Elsevier API (https://dev.elsevier.com/) and elsapy python module (https://github.com/ElsevierDev/elsapy). Nature Neuroscience methods sections were obtained using Python and beautifulsoup. For both journals, only peer reviewed articles were queried, specifically Articles and Reports for Neuron and Articles and Brief Communications for Nature Neuroscience. From these texts, searches were made for the terms "AAV" and "DIO" or "flex". All data was reviewed to omit false positives (eg, amino acid sequences containing "AAV" or use of the lipophilic dye "DiO").

Viruses

AAV plasmids were produced in NEB Stable cells grown for 20hrs in 250ml of LB containing the appropriate antibiotic at 30C. These plasmids were provided to the Salk Vector Core for AAV production. All plasmids were restricted with XmaI to test for gross recombination prior to submission. Titer for AAV8 nef-DIO-FlpO was 2.32 x 10^{12} GC/ml.

AAV crude preps were also used. To produce these crude preps, 10ug of the AAV plasmid DNA and 15ug of the helper plasmid pDP8 (Grimm et al., 2003) were transfected by PEI into a 70-80% confluent 15cm plate of HEK 293T cells. Three days later, the media was aspirated, the cells washed with 10mls of cold dPBS, and scraped off and collected in a 1.5ml Eppendorf tube. Cells were then spun at 510rcf for 3 minutes, the supernatant

aspirated, and the cells resuspended in 1ml of cold gradient buffer (10mM Tris pH 7.6, 150mM NaCl, 10 mM $MgCl_2$, 100mL, 0.22um filtered). Three rounds of freezing in liquid nitrogen, thawing in a 37C water bath, and vortexing were performed to lyse the cells. Finally, the lysed cells were spun at 14K rpm for 10 minutes at 4C and the supernatant collected and pushed through a 0.22um filter (Millipore SLGV033RS) into a sterile 2ml freezer tube. The crude prep was then aliquoted and stored at -80C. These crude preps, typically for small batch use, were not tittered.

Mice

RCF-tdT mice, a gift from the lab of Dr. Sam Pfaff, are flp-reporter mice containing the Rosa-CAG-Frt-Stop-Frt-tdTomato allele and were made from crossing the flp and cre reporter mouse line Ai65 (RCFL-tdT; Jackson Laboratory 021875) with a CMV-cre mouse line. Ai14 (Jackson Laboratory 007914) expresses tdTomato in a cre-dependent manner from a Rosa-CAG-Lox-Stop-Lox-tdTomato allele. The Som-ires-flp (Jackson Laboratory 028579) and PV-ires-cre (Jackson Laboratory 008069) were crossed with the aforementioned strains to specifically target the somatostatin and parvalbumin interneuron populations, respectively.

Both male and female mice were used. Injected were conducted at ages between 80 and 150 days. Stereotactic coordinates used for all injections targeting primary visual cortex were, from bregma: AP -3.28, ML 2.5, DV 0.5.

Tissue Processing and IHC

Mice were anesthetized with euthasol and intracardially perfused using 50mls of PBS and 100mls of 4% paraformaldahyde (PFA). Following perfusion, brains were dissected and placed in 5mls of a 50/50 mixture of 4% PFA and 30% sucrose in PBS at 4C. Having sunk in this solution, brains were transferred to 5mls of 30% sucrose in PBS at 4C. Following sinking, brains were cut on a microtome at 50um and slices were stored in solution of 1% NaN_3 in PBS.

Tissue staining started with a 5 minutes wash of 0.1% Triton X-100 in PBS followed by a 2 to 3 hour block comprised of 5% normal donkey serum in 0.5% Triton X-100 in PBS at room temperature. Primary staining occurred overnight at 4C. Primary antibodies including rabbit anti-somatostatin (Peninsula Labs T-4103) and rabbit anti-parvalbumin (Swant PV27) were applied at 1:1000 in block. Following primary staining, tissue was washed for 10 minutes in PBS three times. The secondary antibody Alexaflour 647 donkey anti-rabbit (Invitrogen A31573) was diluted at 1:500 in PBS and applied to the tissue for 2 to 3 hours at room temperature. Following secondary staining, tissue was washed thrice for 10 minutes in PBS. 10um DAPI was applied for 10 minutes following by three five-minute PBS washes. Tissue was stored in 1% NaN_3 in PBS at 4C prior to mounting.

Results

Commonly used recombinase-dependent AAV constructs suffer from off-target expression when expressing sensitive transgenes

To quantify the degree to which recombinase-dependent AAVs are being used in neuroscience, the text of methods sections from all peer-reviewed articles from the journals of Neuron and Nature Neuroscience were collected for the years between 2005 and 2017. These texts were searched for the terms “AAV” and “DIO” or “flex” and the results reviewed to ensure returns did not refer to acronyms other than adeno-associated virus or recombinase-dependent AAVs, respectively. The results show a widespread and rapid adoption of the technique over the last 10 years; in 2017, 39 of the 129 peer-reviewed articles queried from Nature Neuroscience (30.23%) and 67 of the 289 peer-reviewed articles queried from Neuron (23.18%) used recombinase-dependent AAVs, a 49.95% and 166.18% increase, respectively, from the prior year.

To illustrate leak expression, recombinase-dependent AAVs expressing an orthogonal recombinase were used, due the widespread availability of reporter lines, the increased application of intersectional approaches requiring such viral designs (Capelli et al., 2017; Li et al., 2018; Penzo et al., 2015), and the highly sensitive nature of the recombinase system, as only four peptide copies are required to mediate recombination (Chen et al., 2000; Guo et al., 2002). These AAVs were the injected in primary visual cortex of transgenic mouse lines crossed such that transgene expression should be restricted to common interneuron subclasses. AAV8 nef-fDIO-iCre injection into an Ai14 x Som-

flp mouse resulted in sparse tdTomato labeling around the injection site (Fig 1B). Staining against somatostatin, however, reveals many tdTomato positive cells unlabeled by the somatostatin antibody (white arrows). Similarly, leak is evident from AAV8 nef-DIO-FlpO injection of RCF-tdT x PV-Cre mouse results in many tdTomato positive neurons unlabeled by the parvalbumin antibody (Fig 1C). These results provide prime examples of typical leak expression from recombinase-dependent AAVs.

Recombinase-dependent AAV viral and plasmid preparations contain inverted transgenes

PCR primers were designed to selectively amplify template DNA containing transgenes that had reverted to be in-sense with the promoter and downstream transcription termination motifs (Fig 2A, top). To ensure any inversions were not specific to virus produced by the Salk vector core, we tested DNA extracted from AAV preparations from the Salk as well as from the UNC and UPenn vector cores. In addition, both DIO and FLEX constructs were tested to assess if any inversions were specific to either design. All three AAVs produced a doublet of bands (Fig 2A) which, when sequenced, invariably revealed reversion of the transgene to be in-sense with the wpre element. The sequence difference differentiating these two bands was the presence of either a single or triplicate of recombinase sites in a pattern indicative of inversion across one set of orthogonal recognition sites (Fig 1B).

To see whether this same pattern of reversion was inherited from the plasmids used to produce AAV preparations, we applied a similar strategy of parallel and antiparallel

PCR primers (Fig 3A) to a preparation of pAAV-nef-DIO-FlpO-wpre-pA plasmid. Test conditions spanning either end of the transgene (Fig 3A, lanes 3 & 6) produced doublet and triplicate bands similar to those seen in PCR of AAV genomes (Fig 1A) and of the predicted size for reversions across orthogonal sites. Sequencing of the largest and smallest of these bands produced sequences identical to those found from AAV genomes. Based on these sequences, the lox-sites across which recombination is expected to have occurred to have produced these bands is indicated (Fig 3A, bottom-right). Interestingly, a third band containing two lox sited between the reverted transgene and wpre element was also found, the sequence of which suggested recombination across orthogonal sites *loxp* and *lox2722* (Fig 3A, bottom-right, center).

A larger band, or bands, was typically, but not always, found in these and similar PCR experiments (for example, Fig 3A, dotted box). Sequencing attempts of this fragment using the wpre primer failed to return results. Sequencing using primers against the ends of the FlpO coding region revealed that this band contained the FlpO sequence flanked by inverted wpre sequences (Fig. 3B & C). Given these sequence results, the single band, and the difficulty of sequencing with the wpre plasmid, we reasoned that this band resulted from single-primer amplification across a region containing flanking inverted sequences, which suggested a duplicative nature of the recombinant plasmid.

Plasmid doubling due to a recombination across homologous recombinase sites leads to transgene reversion

Due to the presence of two nested sets of recombinase recognition sited in DIO and FLEX constructs, we developed a simplified model plasmid system to ascertain the nature of spontaneous transgene inversions. In these plasmids, the LacZ element of the common cloning plasmid pUC19 was inverted and flanked by antiparallel 34 base pair sequences (Fig 4A). To assess whether the mechanism of inversion was specific to the structure of site-specific recombinase recognition site or due to gross homology we designed these 34bp flanking regions to be either the canonical common recombinase recognition sequences (*loxp* for cre and *frt* for flp) or shuffled *loxp* sequences lacking any of the sequence structure of typical recombinase binding motifs in either a homologous or non-homologous setup (Fig 4B).

To assess the rate of recombination in each condition, NEB Stable *E. coli* were transformed with these plasmids and plated on XGal/IPTG plates, from which a single white colony was grown in 2ml LB media at 30C. This preparation was then diluted and plated on XGal/IPTG plates, revealing low numbers of spontaneous blue colonies (Fig 4C). Counting of both white and blue colonies revealed a recombination rate slightly greater than 1/1000 colonies across the three conditions containing homologous regions, regardless of sequence. No blue colonies were found in the non-homologous condition (Fig 4D).

To more sensitively detect reversion events, we conducted qPCR in which the targeted amplicon was specific to plasmids in which the LacZ was in-sense with the upstream operon. Absolute quantification against a synthesized dsDNA standard

suggested recombinant frequencies between 1/10000 and 1/1000 when compared to LacZ copy number, and detected recombinants in the non-homologous condition at a level considered equivalent to noise (Fig 4E).

It was our hypothesis that transgene reversion events, at the molecular level, were similar in nature to those of typical recombinase reactions, in which the nucleotides within a set of antiparallel *lox* or *frt* sites are flipped while the surrounding genetic material remains unchanged. Based on this hypothesis, we surmised that the rate of recombination from white to blue colonies should be the same as those from blue to white, and conducted the same set of experiments on blue colonies selected from those conditions in which they were found. Surprisingly, no white colonies were found on any X-Gal/IPTG plates from bacterial preparations started from a blue colony (Fig 4F & G). Further, qPCR of these plasmid preparations revealed a recombinant:LacZ ratio of roughly 1:2 (Fig 4H).

Given these results, the pattern of sequences found in AAV genome and plasmid PCR experiments, and the previously published model of recombination between inverted repeats (see Fig 5, Bi and Liu, 1996) we predicted that recombination across homologous sites at different positions within or between plasmids resulted in a plasmid doubling event leaving the inverted LacZ coding sequence in-sense with the operon (Fig 4I). Given the nature of this doubling, we further predicted that DNA from blue colonies would produce unique bands when cut with a restriction enzyme expected to have a single recognition site on the original plasmid. SacI restriction of DNA from preparations grown from white and blue colonies revealed bands equivalent to those predicted by such a doubling (Fig 4J), confirming this hypothesis.

Both intact ORF and spontaneous reversions contribute to leak from recombinase-dependent AAV plasmids

Having deduced the nature of these recombination events and knowing that disruption of the inverted transgene ORF leads to decreased leak, we hypothesized that infrequent transcription from the inverted ORF was responsible for leak expression but that previous strategies seeking to block leak by disrupting the ORF by moving the transgene start codon outside of the recombinase recognition sites found persistent leak due to reconstitution of this ORF by recombination. To test this hypothesis we employed a model AAV plasmid design with a single pair of antisense 34bp *loxp* shuffled sequences bookending the inverted transgene (Fig 5A). To probe the role of spontaneous reversions in leak expression of cre from these constructs we used either homologous (HS) or nonhomologous (NHS) shuffled *loxp* sequences, which left the plasmids sensitive or insensitive to spontaneous reversion, respectively. To ascertain the role of leak expression from the inverted ORF the constructs were further modified by removal of the start codon (ATG-Out; AO) from the transgene and placing it outside the shuffle *loxp* sites and downstream from the promoter. Finally, a HEK-LSL-GFP cell line expressing GFP under cre control and mCherry constitutively (Fig 5B) was used as a reporter of leaky cre expression. Transfection of this cell line with pCAG-fDIO-Cre in the absence of flp results in a small number of GFP+ cells, indicative of leak (Fig. 5B column 2).

Flow cytometry of HEK-LSL-GFP cells transfected with an empty pCAG vector and iCre controls (Fig. 5C, column 1) resulted in obvious and strong GFP+ expression in the positive control. To measure the degree of leak expression in subsequent transfection of test constructs, a rectangular gate was constructed such that all GFP+ cells in the positive

control were included, while all cells from the empty vector condition were avoided. Transfection of cre leak test conditions (Fig. 5C, columns 2 & 3, top) revealed the highest degree of leak in conditions in which the ORF was maintained. When the ORF was disrupted but recombination was not (Fig 5C, column 2, bottom), as is the case in previous attempts to mitigate leak from AAVs, leak expression was diminished but not extinguished. Finally, very few GFP+ cells were found in the condition in which both the ORF and spontaneous reversion were disrupted (Fig 5C, bottom right panel), confirming our hypothesis that abolishment of leak requires disruption of both recombination and the transgene ORF. Summary data from six transfections of test conditions and the empty vector (Fig. 5D) shows consistency in this pattern across transfections as measured by percent of cells in the gated region.

To describe the role of leak expression in less-sensitive, or accumulative, transgenes, we performed the same experiments by transfecting the same plasmids with EYFP instead of iCre as the transgene in HEK 293T cells (Fig. 5E). In these experiments, a gate was similarly draw encompassing the EYFP+ cells in the positive control, but the bottom bounding line of the rectangular gate was drawn just above the bulk of cells in the empty vector condition (Fig. 5E, left column). While no leak conditions showed expression levels as high as those seen in this positive control, as a percentage of parent population the HS condition in which both ORF and spontaneous reversion are left unhampered (Fig. 5E, middle column, top panel) showed a significant increase in the number of low-expressing GFP+ cells (Fig. 5F), indicating that low, but present, levels of leak expression still occur from less sensitive transgenes.

Disruption of the 5' ITR D-sequence blocks packaging of AAV genomes containing reverted transgenes but not leak expression

To make a leakless recombinase-dependent AAV we targeted the packaging of genomes containing the reverted transgene into AAV particles by taking advantage of the mirrored nature of the recombination event and the requirement of specific ITR motifs for genome packaging. During production, both positive and negative ssDNA AAV genomes are produced from the plasmid (Fig 6A). Assuming recombination only occurs across homologous orthogonal recombinase recognition sites (which we have shown is not necessarily the case (Fig. 3A) but is omitted here for brevity), there are two first order recombinant plasmids, each of which contains two ssDNA AAV genomes, creating 4 different potential ssDNA AAV genomes (Fig. 6B). Due to the mirrored nature of the template AAV genome sequences present in the plasmids resulting from recombination, both the 5' and 3' ITRs bounding the AAV genomes in these recombinant plasmids are inherited either from the 5' or 3' end of the original plasmid, as shown in Fig. 6B in which the green ITRs are labeled 5' or 3' in reference to their original orientation.

Recent work has shown that disruption of the 20nt D-sequence located at the base of either AAV ITR blocks the packaging of ssDNA genomes containing that substitution at their 5' end (Fig 6C). Based our previous results implicating the transgene ORF in leaky expression, we designed a recombinase-dependent plasmid in which the kozak sequence and start codon are outside the recombinase-specific sites and a p2a element is placed upstream in sense with the transgene coding sequence, such that upon recombinase-dependent inversion the ORF will be reconstituted, transcription will commence, and the peptide resulting from readthrough of the *lox* or *frt* sites between the start codon and

transgene coding sequence will be cleaved off (Fig 6C). Further, the D-sequence of the 5' ITR of this plasmid is modified, such that upon recombination (Fig 6E) the resulting AAV genome containing a reconstituted ORF will be bookended by two D-sequence mutated ITRs. As such, we hypothesized that AAVs produced from this plasmid design would not package any genomes disposed to leak expression, at the cost of packaging only one sense, or half, of the intended genome (Fig 6F). We designated this design D-sequence Substituted (5DS), and PCR performed against AAV8 nef-fDIO-iCre and AAV8 NDS-nef-ATG-fDIO-p2a-iCre with primers designed to detect all possible transgene orientations confirmed that the NDS design contained no detectable genomes in which the promoter was in-sense with the transgene (Fig 6G). Put another way, no genomes with reconstituted ORFs were packaged. Surprisingly, injection of this AAV8 5DS- nef-ATG-fDIO-p2a-iCre virus in a Som-flp x Ai14 mouse had little discernable effect on leak expression, as numerous tdTomato positive cells unlabeled somatostatin positive neurons were found (Fig 6G). This finding suggests that some AAV-specific mode of transcription of the inverted transgene, even with the start codon removed, leads to leak expression.

Discussion

Recombinase-dependent AAVs have rightly become a popular tool in systems neuroscience. Coupled with transgenic mice, they allow for the targeted expression of transgenes to specific neuron populations. However, we've shown that current DIO and FLEX designs are not immune from off-target expression. To measure the extent of leak expression, we chose to use recombinases as reporters of leak expression (Fig 5C). These experiments showed clear leak expression in those plasmids in which the ORF was intact or recombination could occur. In contrast, the same experiments conducted using EYFP revealed levels of leak expression which, though detectable by flow cytometry (Fig 5D), are likely of negligible concern in typical fluorescent protein applications. This expression was only detected in conditions in which both recombination and the transgene ORF were undisrupted (Fig 5D, HS condition, middle column, top). It remains unclear whether this increased expression is due to reversion of the transgene or from transcription of the inverted transgene, as we did not detect elevated levels of EYFP expression in either the AO HS or NHS conditions. One hypothesis is that the extra amino acids on the N-terminal of EYFP following readthrough of the shuffled *loxp* sequence in the AO HS condition leads to decreased expression of florescence. Regardless, these experiments confirm that DIO and FLEX designs leak when containing enzymatically or accumulatively detected transgenes. Whether an investigator should be concerned about AAV leak from these constructs depends on the intended application and the nature of the transgene being expressed.

We show here that the presence of recombinant plasmids in a model system, as measured by spontaneous blue colony frequency and absolute quantification by qPCR, to be between 1/1000 and 1/10000 recombinants per non-recombinant (Fig 4D &E). Due to the overlapping nature of the two sets of recombinase recognition sites in DIO and FLEX designs, the duplicative nature of the recombinant plasmid, and the low GC content of the interstitial sequences flanking the recombinase recognition sites, we were unable to design a reliable method for quantifying the number of reverted genomes in AAV preparations or the plasmids used to make them. Given the size restrictions on genome packaging and potential differences in replication kinetics when multiple genomes of varying sizes are present in the same host cell, it is hard to estimate if the frequency of recombinants amongst AAV genomes is similar to that seen in the plasmid populations used to make them. Previous work has shown that recombination between inverted repeats decreases as intervening sequence size increases (Bi and Liu, 1996), suggesting variance in recombination frequency in AAV designs harboring different sized transgenes. Of potentially greater concern are the rare cases in which entire AAV preparations contain genomes from recombinant plasmids, which would occur if the wrong bacterial colony was unwittingly selected for plasmid production and if quality control checks, typically based on Xma1 and SmaI restriction at the ITRs, were not stringent enough to detect the difference. Indeed, in our blue/white screens we found colonies that were only partially blue, suggesting recombination events after plating, which in an AAV preparation would lead a large percentage of the AAVs to contain genomes with reverted transgenes.

We considered several approaches for abolishing leak expression. The simplest, given the duplicative nature of the recombination event, is size exclusion of the offending

recombinant by gel electrophoresis. While we may be unaware of a technique by which to perform this with high yield, we avoided this technique as it would prove cumbersome due to the high quantity of plasmid DNA needed for AAV production.

It is possible that solutions remain to be investigated at the level of plasmid production. It is our understanding that the *RecA*-independent recombination event is not dependent on a single gene product or pathway, but is rather a very specific recombination event that is the unwanted consequence of single strand breaks occurring during plasmid replication. As such, this process may be difficult to target by mutation of a single gene or pathway. However, there may yet be a bacterial mechanism upon which manipulation would produce a competent strain with decreased *RecA*-independent recombination which would allow for leakless production of recombinase-dependent AAVs without plasmid modification.

Unaware of such a mechanism, we targeted the AAV packaging mechanism for selective removal of the recombinant plasmid population from the AAV genome. Since AAV genomes with modified D-sequences in both ITRs go unpackaged (Wang et al., 1996, 1997) and the duplicative nature of the recombination leads a single ITR from the original plasmid to be at either end of the offending AAV genome containing the reverted transgene (Fig 6E & F), the D-sequence substitution should provides a simple means by which to exclude recombinants from packaging. Beside the removal of the start codon for disruption of the ORF, the D-sequence modification requires only a 20nt change of the standard AAV plasmid, making retrofitting and implementation relatively simple for labs wishing to adapt the system to their own constructs. The system would come at the cost of titer, however, as only a single strand of the ssDNA genome will be packaged, the titer will be relatively

halved in these AAVs. But this loss may be amply counteracted, as previous studies have shown that AAVs containing singly substituted D-sequences have higher transgene expression levels, potentially due to the removal of the transcriptionally inhibitory D-sequence (Ling et al., 2015).

This technique was successful in blocking the packaging of genomes in which recombination had resulted in ORF reconstitution (Fig 6G), but, surprisingly, this method failed to prevent leak expression of cre in a Som-flp x Ai14 mouse, as evidenced by tdTomato positive cells unlabeled by the anti-somatostatin antibody, many with obvious non-somatostatin neuron (eg, pyramidal) morphologies (Fig 6H). There are two potential hypotheses explaining this expression. The first is that recombinase expression from the mouse genome is itself leaky, resulting in off target activation. While this is certainly possible, we have not noticed the degree of off-target expression seen here in crosses with reporter mice in which any off-target expression would be obvious at a global scale, making this explanation unlikely. The second is that a property inherent of the AAV genome structure contributes to leak expression. Indeed, previous work has found that the ITR alone can act as a promoter to drive expression from AAV constructs (Flotte et al., 1993), and experiments in our lab have shown that removal of the start codon from cre recombinase in mammalian expression plasmids still results in leak (data not shown). As such, it could be that recombinase expression from the AAV8 NDS-nef-ATG-fDIO-iCre construct occurred from transcription of the inverted transgene driven by the ITR, in which case a minor modification of the ITR or incorporation of a polyA stop may be enough to halt transcription. Future experiments will investigate this phenomenon in the hope of limiting its effect.

The expanding application of recombinase-dependent AAVs in neuroscience demands increasingly precise and replicable systems for neural circuit dissection. The issue of off-target transgene expression investigated here provides details regarding a problem that has long plagued an established technique. It is our hope that in continuing to investigate the issue we will come upon a solution that could greatly expand the viral toolkit available to the modern neuroscientist. Even without a solution, the work presented here provides important caveats to be considered by those designing, producing, and using recombinase-dependent AAVs.

Acknowledgements

Chapter 2, in full, is in preparation for submission for publication. The dissertation author will be the first author of this work. Second author will be Hannah Collins, and Dr. Edward Callaway will be the third and final author as principle investigator.

References

Albert, H., Dale, E.C., and Lee, Elsa and Ow, David W. (1995). Site-specific integration of DNA~into wild-type and mutant lox sites placed in the plant genome. Plant J. *7*, 649–659.

Arakil, K., Araki, M., and Yamamura, K. (1997). Targeted integration of DNA using mutant fox sites in embryonic stem cells. *25*, 868–872.

Bi, X., and Liu, L.F. (1996). DNA rearrangement mediated by inverted repeats. Proc. Natl. Acad. Sci. U. S. A. *93*, 819–823.

Capelli, P., Pivetta, C., Esposito, M.S., and Arber, S. (2017). Locomotor speed control circuits in the caudal brainstem. Nature *551*, 373–377.

Cardin, J.A., Carlén, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L.H., and Moore, C.I. (2009). Driving fast-spiking cells induces gamma rhythm and controls sensory responses. Nature *459*, 663–667.

Chen, Y., Narendra, U., Iype, L.E., Cox, M.M., and Rice, P. a (2000). Crystal structure of a Flp recombinase-Holliday junction complex: assembly of an active oligomer by helix swapping. Mol. Cell *6*, 885–897.

Chiorini, J. a, Wiener, S.M., Owens, R. a, Kyöstió, S.R., Kotin, R.M., and Safer, B. (1994). Sequence requirements for stable binding and function of Rep68 on the adeno-associated virus type 2 inverted terminal repeats. J. Virol. *68*, 7448–7457.

Dillingham, M.S., and Kowalczykowski, S.C. (2008). RecBCD Enzyme and the Repair of Double-Stranded DNA Breaks. Microbiol. Mol. Biol. Rev. *72*, 642–671.

Flotte, T.R., Afione, S. a, Solow, R., Drumm, M.L., Markakis, D., Guggino, W.B., Zeitlin, P.L., Carter, B.J., Flottessli, T.R., and Afiones, S.A. (1993). Expression of the cystic fibrosis transmembrane conductance regulator from a novel adeno-associated virus promoter. J. Biol. Chem. *268*, 3781–3790.

Grimm, D., Kay, M.A., and Kleinschmidt, J.A. (2003). Helper virus-free, optically controllable, and two-plasmid-based production of adeno-associated virus vectors of serotypes 1 to 6. Mol. Ther. *7*, 839–850.

Guo, C., Yang, W., and Lobe, C.G. (2002). A Cre Recombinase Transgene with Mosaic , Widespread Tamoxifen-Inducible Action. *18*, 8–18.

Hoess, R.H., Ziese, M., and Sternberg, N.A.T. (1982). P1 site-specific recombination: Nucleotide sequence of the. *79*, 3398–3402.

Huang, Z.J., Taniguchi, H., He, M., and Kuhlman, S. (2014). Cre-dependent adeno-associated virus preparation and delivery for labeling neurons in the mouse brain. Cold Spring Harb. Protoc. *2014*, 190–194.

Jungmann, A., Leuchs, B., Rommelaere, J., Katus, H.A., and Müller, O.J. (2017). Protocol for Efficient Generation and Characterization of Adeno-Associated Viral Vectors. Hum.

Gene Ther. Methods *28*, 235–246.

Kohara, K., Pignatelli, M., Rivest, A.J., Jung, H.Y., Kitamura, T., Suh, J., Frank, D., Kajikawa, K., Mise, N., Obata, Y., et al. (2014). Cell type-specific genetic and optogenetic tools reveal hippocampal CA2 circuits. Nat. Neurosci. *17*, 269–279.

Kuhlman, S.J., and Huang, Z.J. (2008). High-resolution labeling and functional manipulation of specific neuron types in mouse brain by Cre-activated viral gene expression. PLoS One *3*.

Lee, G., and Saito, I. (1998). Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination. Gene *216*, 55–65.

Li, Y., Zeng, J., Zhang, J., Yue, C., Zhong, W., Liu, Z., Feng, Q., and Luo, M. (2018). Hypothalamic Circuits for Predation and Evasion. Neuron *97*, 911–924.e5.

Ling, C., Wang, Y., Lu, Y., Wang, L., Jayandharan, G.R., Aslanidi, G. V., Li, B., Cheng, B., Ma, W., Lentz, T., et al. (2015). Enhanced Transgene Expression from Recombinant Single-Stranded D-Sequence-Substituted Adeno-Associated Virus Vectors in Human Cell Lines *In Vitro* and in Murine Hepatocytes *In Vivo*. J. Virol. *89*, 952–961.

Lovett, S.T., Hurley, R.L., Sutera, V.A., Aubuchon, R.H., and Lebedeva, M.A. (2002). Crossing over between regions of limited homology in Escherichia coli: RecA-dependent and RecA-independent pathways. Genetics *160*, 851–859.

Madisen, L., Zwingman, T. a, Sunkin, S.M., Oh, S.W., Zariwala, H. a, Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., et al. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. *13*, 133–140.

Miyamichi, K., Shlomai-Fuchs, Y., Shu, M., Weissbourd, B.C., Luo, L., and Mizrahi, A. (2013). Dissecting local circuits: Parvalbumin interneurons underlie broad feedback control of olfactory bulb output. Neuron *80*, 1232–1245.

Penzo, M.A., Robert, V., Tucciarone, J., De Bundel, D., Wang, M., Van Aelst, L., Darvas, M., Parada, L.F., Palmiter, R.D., He, M., et al. (2015). The paraventricular thalamus controls a central amygdala fear circuit. Nature *519*, 455–459.

Rong, L., Gendron, K., Strohl, B., Shenoy, R., Wool-Lewis, R.J., and Bates, P. (1998). Characterization of determinants for envelope binding and infection in tva, the subgroup A avian sarcoma and leukosis virus receptor. J. Virol. *72*, 4552–4559.

Schlake, T., and Bode, J. (1994). Use of Mutated FLP Recognition Target (FRT) Sites for the Exchange of Expression Cassettes at Defined Chromosomal Loci. Biochemistry *33*, 12746–12751.

Schnütgen, F., Doerflinger, N., Calléja, C., Wendling, O., Chambon, P., and Ghyselinck, N.B. (2003). A directional strategy for monitoring Cre-mediated recombination at the cellular level in the mouse. Nat. Biotechnol. *21*, 562–565.

Seidler, B., Schmidt, A., Mayr, U., Nakhai, H., Schmid, R.M., Schneider, G., and Saur, D. (2008). A Cre-loxP-based mouse model for conditional somatic gene expression and

knockdown in vivo by using avian retroviral vectors. Proc. Natl. Acad. Sci. *105*, 10137–10142.

Senecoff, J.F., Rossmeissl, P.J., and Cox, M.M. (1988). DNA recognition by the FLP recombinase of the yeast 2 μ plasmid. A mutational analysis of the FLP binding site. J. Mol. Biol. *201*, 405–421.

Taniguchi, H., He, M., Wu, P., Kim, S., Paik, R., Sugino, K., Kvitsani, D., Fu, Y., Lu, J., Lin, Y., et al. (2011). A Resource of Cre Driver Lines for Genetic Targeting of GABAergic Neurons in Cerebral Cortex. Neuron *71*, 995–1013.

Voziyanov, Y., Pathania, S., and Jayaram, M. (1999). A general model for site-specific recombination by the integrase family recombinases. Nucleic Acids Res. *27*, 930–941.

Wall, N.R., Wickersham, I.R., Cetin, A., De La Parra, M., and Callaway, E.M. (2010). Monosynaptic circuit tracing in vivo through Cre-dependent targeting and complementation of modified rabies virus. Proc. Natl. Acad. Sci. U. S. A. *107*, 21848–21853.

Wang, X.S., Ponnazhagan, S., and Srivastava, A. (1996). Rescue and replication of adeno-associated virus type 2 as well as vector DNA sequences from recombinant plasmids containing deletions in the viral inverted terminal repeats: selective encapsidation of viral genomes in progeny virions. J. Virol. *70*, 1668–1677.

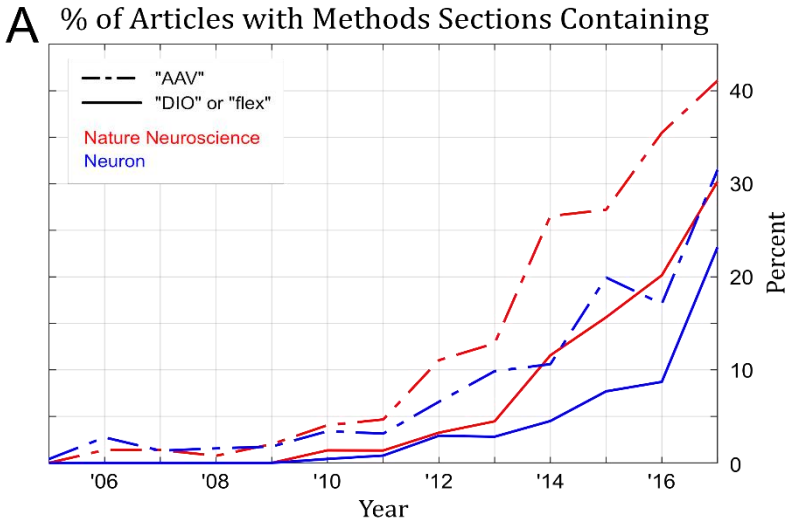
Wang, X.S., Qing, K., Ponnazhagan, S., Srivastava, A., Wang, X., Qing, K., and Ponnazhagan, S. (1997). Adeno-associated virus type 2 DNA replication in vivo : mutation analyses of the D sequence in viral inverted terminal repeats . Adeno-Associated Virus Type 2 DNA Replication In Vivo : Mutation Analyses of the D Sequence in Viral Inverted Terminal Repeats. *71*, 3077–3082.

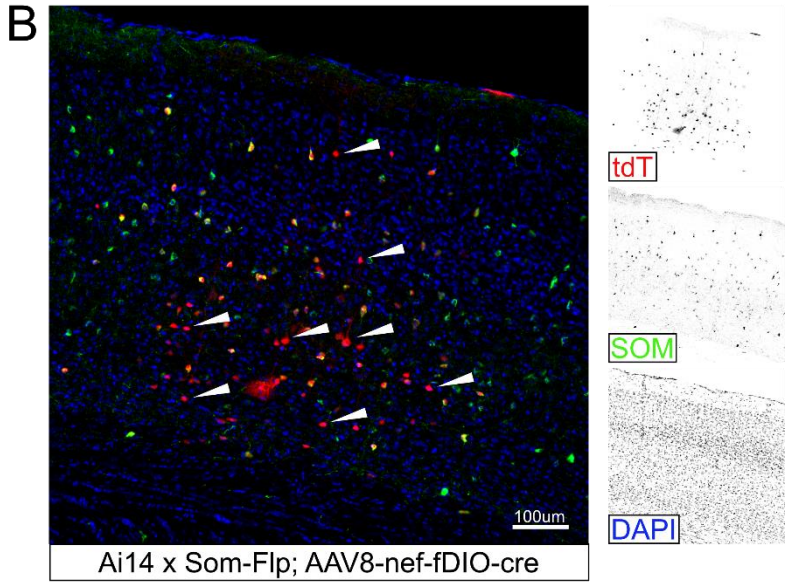
Yang, C.C., Xiao, X., Zhu, X., Ansardi, D.C., Epstein, N.D., Frey, M.R., Matera, A.G., and Samulski, R.J. (1997). Cellular recombination pathways and viral terminal repeat hairpin structures are sufficient for adeno-associated virus integration in vivo and in vitro. J. Virol. *71*, 9231–9247.

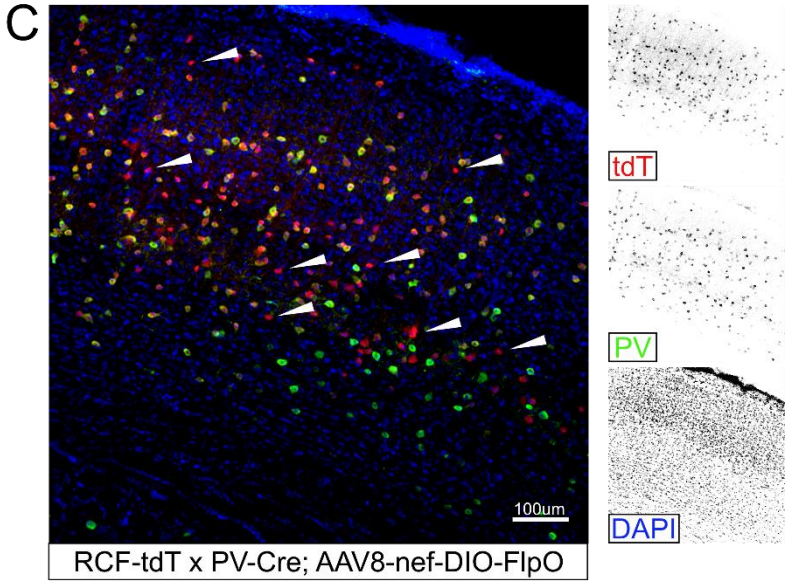
Yang, J., Zhou, W., Zhang, Y., Zidon, T., Ritchie, T., and Engelhardt, J.F. (1999). Concatamerization of adeno-associated virus circular genomes occurs through intermolecular recombination. J. Virol. *73*, 9468–9477.

Figure 2.1: Recombinase-dependent AAVs, an increasingly common tool in neuroscience, suffer from off-target "leak" expression

(A) Frequency of use of the term "AAV" (dashed line) or "DIO" or "flex" in reference to AAV (solid line) in the methods sections of peer-reviewed articles in Nature Neuroscience (red) and Neuron (blue). By year, 2005 to 2017. (B) Injection of AAV8-nef-DIO-FlpO into primary visual cortex of parvalbumin-cre x RCF-tdT mouse line. Anti-parvalbumin antibody staining in green, tdTomato in red, DAPI in blue. (C) Injection of AAV8-nef-fDIO-iCre into primary visual cortex of somatostatin-cre x Ai14 mouse line. Anti-somatostatin staining in green, tdTomato in red, DAPI in blue. White arrows indicate example neurons displaying recombinase-dependent expression unlabeled by the target cell-type antibody. Individual channels at right.


A
% of Articles with Methods Sections Containing
"AAV"
"DIO" or "flex"
Nature Neuroscience
Neuron
40
30
20
10
0
Percent
'06
'08
'10
'12
'14
'16
Year


B
100um
tdT
SOM
DAPI
Ai14 x Som-Flp; AAV8-nef-fDIO-cre


C
100um
tdT
PV
DAPI
RCF-tdT x PV-Cre; AAV8-nef-DIO-FlpO

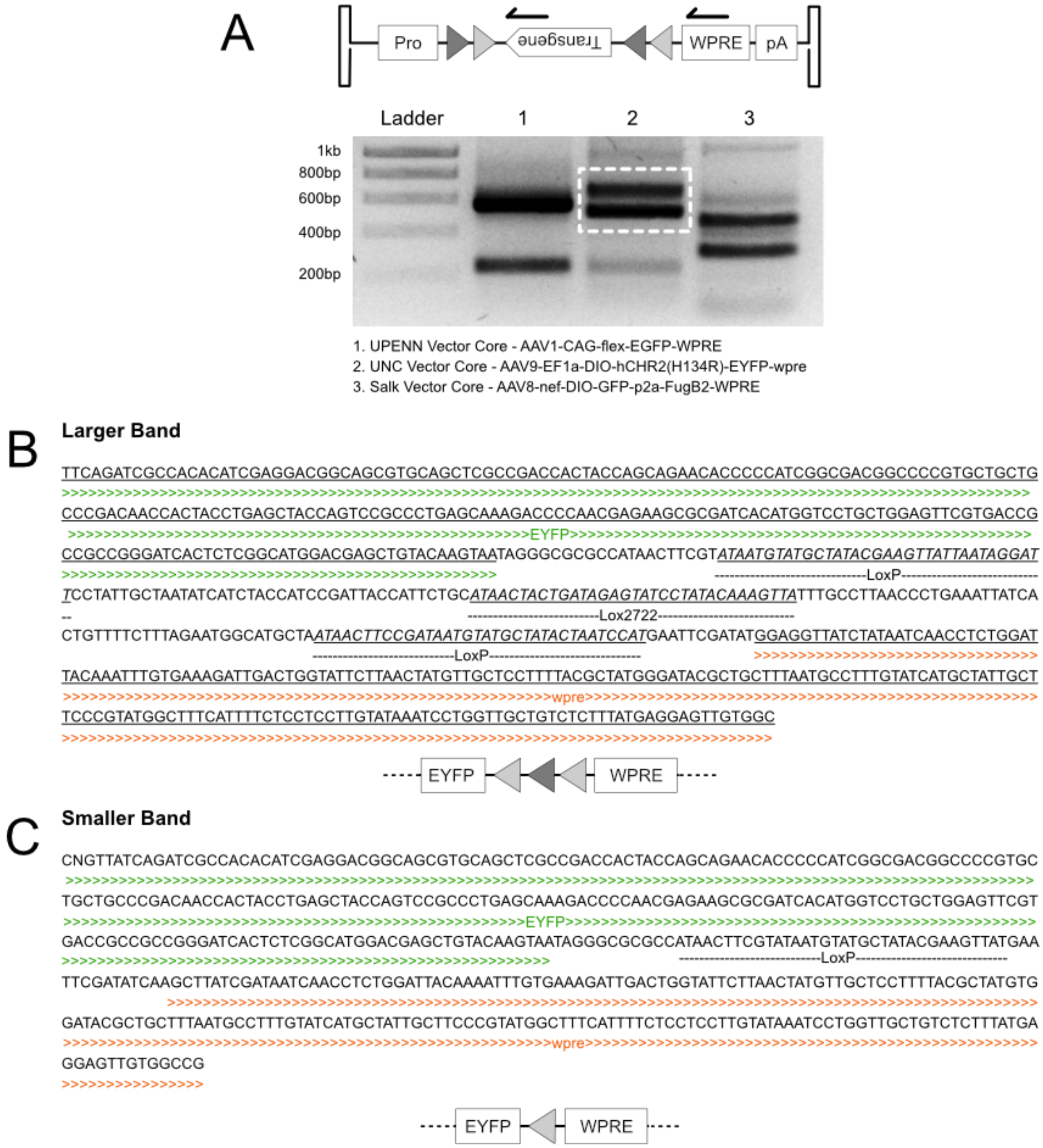


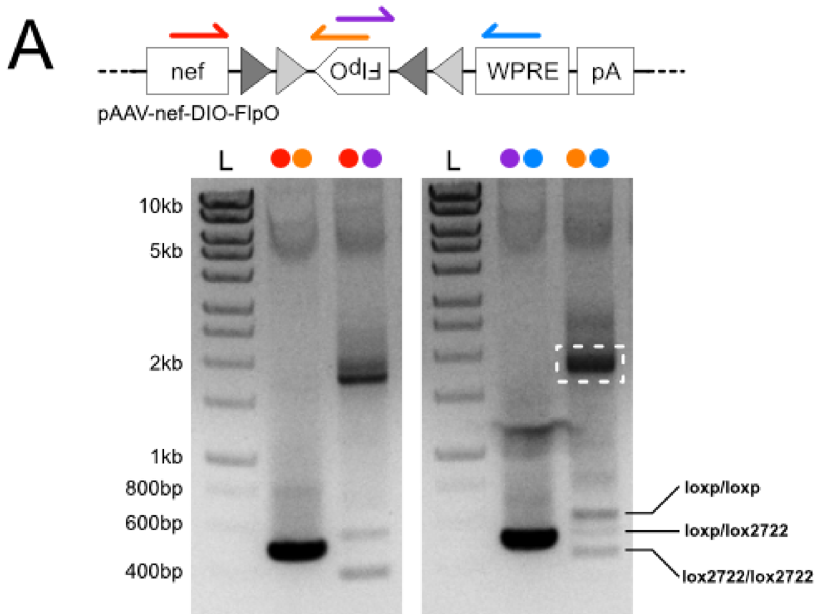
Figure 2.2: PCR products from in-sense primers across recombinase sites reveal reverted transgenes in AAV genomes

(A) PCR of FLEX and DIO AAV genomes using in-sense primers across recombinant sites creates short band doublets. General recombinase-dependent DIO or FLEX AAV design (top) with in-sense PCR primers (black). Gel of PCR products using the AAVs listed with sources (below). Dotted box in lane 2 denotes products used for sequencing. Sequencing of (B) larger and (C) smaller of the two bands show the transgene in-sense with the WPRE element and separated by three or one recombinase sites (triangles).

Figure 2.3: PCR products from in-sense primers across recombinase sites using AAV plasmids as templates replicate findings from AAV genomes and provide evidence for a plasmid duplication event

(A) PCR experiments using four sets of primers with either sense or antisense orientation in reference to the expected design (top) of the template pAAV-nef-DIO-FlpO plasmid. Antisense primers (lanes 2 and 5) produce a band at the expected size. In-sense primer pairs (lanes 3 and 6) produce doublet or triplet bands less 1kb. Cre-recognition site pairs labelling each band (bottom right) indicate the lox sites across which recombination occurred to produce the sequence found for each of these bands. The sequences in (B) and (C) are from the band found in the dotted box in (A) using the flp-specific primers shown.

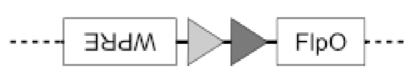
A



B

Sequencing Primer ●

TGAAAGCCATACGGGAAGCAATAGCATGATACAAAGGCATTAAAGCAGCGTAT
CCACATAGCGTAAAAGGAGCAACATAGTTAAGAATACCAGTCAATCTTTCACA
WPRE
AATTTTGTAATCCAGAGGTTGATTATCGATAAGCTTGATATCGAATTC*ATAACTT*
*CGTATAGCATACATTATACGAAGTTAT*TCTTTGCACCATTCTAAAGAATAACAGT
loxp
GATAATTTCTGGGTTAAGGCAA*ATAACTTCGTATAGGATACTTTATACGAAGTTA*
lox2722
*T*GCTAGCGCCGCCACCATGGCTCCTAAGAAGAAGAGGAAGGTGATGAGCCA
GTTCGACATCCTGTGCAAGACCCCCCCCAAGGTGCTGGTGCGGCAGTTCGT
FlpO
GGAGAGATTCGAGAGGCCCAGCGGCGAGAAGATCGCCAGCTGTGCCGCCG
AGCTGACCTACCTGTGCTGGATGATCACCCACAACGGCACCGCCATCAAGA
GGGCCACCT



C

Sequencing Primer ●

CACTACTTCGCCCTGGTGTCCAGGTACTACGCCTACGACCCCATCAGCAAGG
AGATGATCGCCCTGAAGGACGAGACCAACCCCATCGAGGAGTGGCAGCACA
FlpO
TCGAGCAGCTGAAGGGCAGCGCCGAGGGCAGCATCAGATACCCCGCCTGG
AACGGCATCATCAGCCAGGAGGTGCTGGACTACCTGAGCAGCTACATCAACA
GGCGGATCTGAGGCGCGCC*ATAACTTCGTATAATGTATGCTATACGAAGTTAT*
loxp
GAATTCGATATCAAGCTTATCGATAATCAACCTCTGGATTACAAAATTTGTGAAA
GATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTG
WPRE
CTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCAT

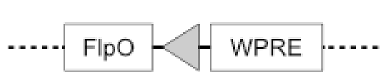


Figure 2.4: Plasmid duplication event resulting from homologous recombination across recombinase-specific sites results in transgene reversion

(A) Design of inverted LacZ expression plasmid based on pUC19 expression vector. The LacZ coding sequence is inverted antisense to the Lac operon and start codon and bookended by two 34bp antisense recognition sites. (B) Sequence of four pairs of tested recombinase recognition sites and shuffled *loxp* sequences in positions 1 and 2 as denoted in (A). (C) XGal/IPTG plate of diluted bacterial miniprep grown from a white colony harboring the pUC19-loxp-inv_LacZ plasmid. Inset shows blue colonies among white colonies. (D) Frequency of blue colonies on XGal/IPTG plates grown from white colonies. From 3 minipreps grown across 3 plates. Horizontal lines signify mean value across minipreps. Average colony number per plate and standard deviation shown below. (E) Relative amounts of reverted LacZ coding sequences to total LacZ copies in DNA minipreps grown from white colonies by qPCR. Horizontal lines signify mean value across minipreps (circles), vertical lines are standard deviation, and boxes 95% confidence intervals. (F) XGal/IPTG plate of diluted bacterial miniprep grown from a blue colony found in plates shown in (C) from a transformation of the pUC19-loxp-inv_LacZ plasmid. Inset shows blue colonies. (G) Frequency of white colonies on XGal/IPTG plates grown from blue colonies. From 3 minipreps grown across 3 plates. Horizontal lines signify mean value across minipreps. Average colony number per plate and standard deviation shown below. (H) Relative amounts of reverted LacZ coding sequences to total LacZ copies in DNA preparations grown from blue colonies by qPCR. Horizontal lines signify mean value across minipreps (circles), vertical lines are standard deviation, boxes 95% confidence intervals. (I) Schematic of proposed inter- or intraplasmid recombination events across antisense homologous recombinase sites at different relative positions resulting in transgene reversion and plasmid doubling. P and P' denote mirrored plasmids. Approximate SacI site denoted by blue dashed circle. Blue arrow denotes transcription from ORF following recombination. (J) SacI restriction profile of minipreps grown from white and blue colonies from conditions in which blue colonies were present. First and last lanes are a 1kb ladder.

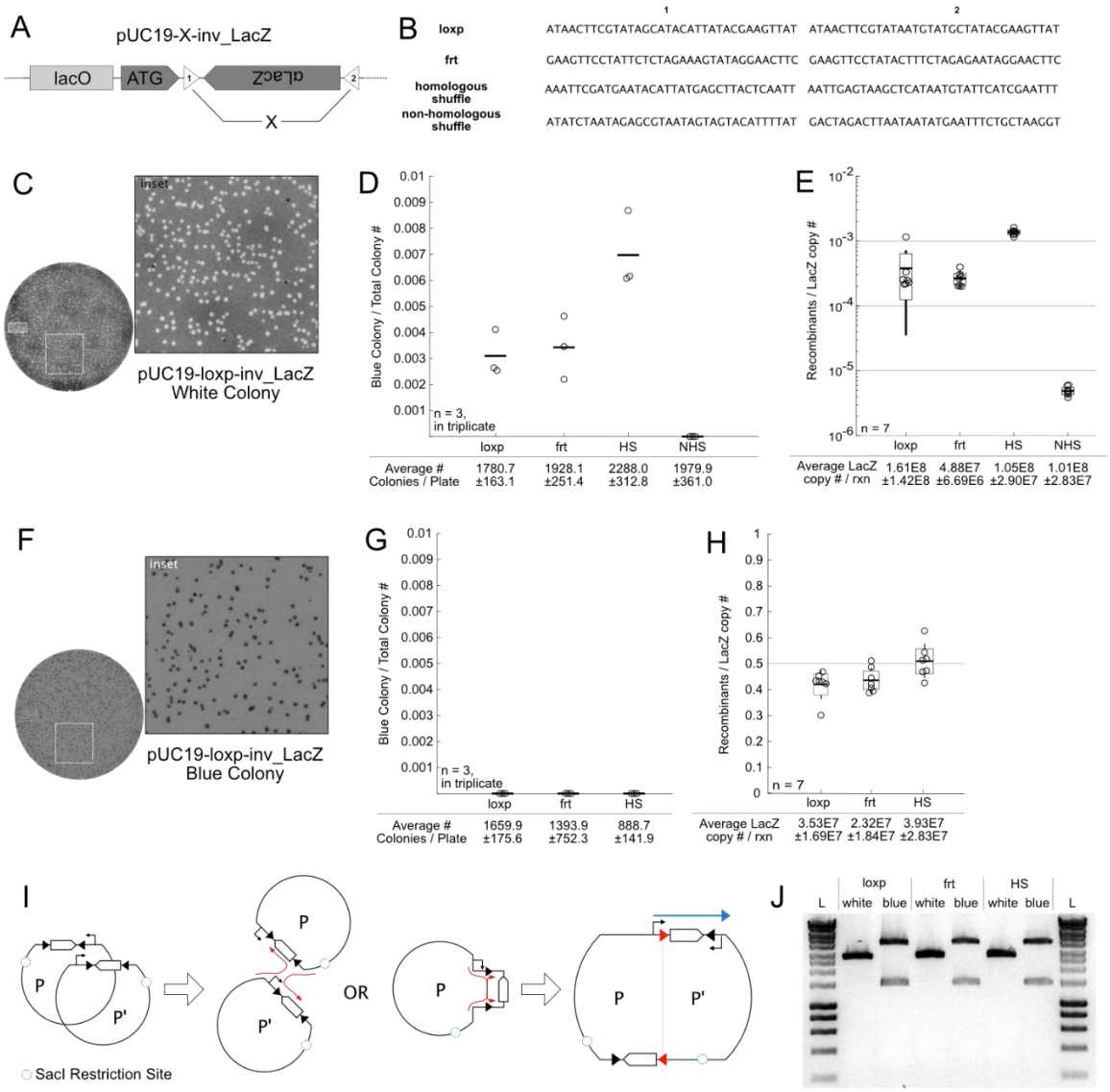

A
pUC19-X-inv_LacZ
lacO
ATG
1
lacZ
2
X
B
1
2
loxp
ATAACTTCGTATAGCATACATTATACGAAGTTAT
ATAACTTCGTATAATGTATGCTATACGAAGTTAT
frt
GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC
GAAGTTCCTATACTTTCTAGAGAATAGGAACTTC
homologous shuffle
AAATTCGATGAATACATTATGAGCTTACTCAATT
AATTGAGTAAGCTCATAATGTATTCATCGAATTT
non-homologous shuffle
ATATCTAATAGAGCGTAATAGTAGTACATTTTAT
GACTAGACTTAATAATATGAATTTCTGCTAAGGT
C
inset
pUC19-loxp-inv_LacZ
White Colony
D
Blue Colony / Total Colony #
n = 3,
in triplicate
loxp
frt
HS
NHS
Average # Colonies / Plate
1780.7 ±163.1
1928.1 ±251.4
2288.0 ±312.8
1979.9 ±361.0
E
Recombinants / LacZ copy #
n = 7
loxp
frt
HS
NHS
Average LacZ copy # / rxn
1.61E8 ±1.42E8
4.88E7 ±6.69E6
1.05E8 ±2.90E7
1.01E8 ±2.83E7
F
inset
pUC19-loxp-inv_LacZ
Blue Colony
G
Blue Colony / Total Colony #
n = 3,
in triplicate
loxp
frt
HS
Average # Colonies / Plate
1659.9 ±175.6
1393.9 ±752.3
888.7 ±141.9
H
Recombinants / LacZ copy #
n = 7
loxp
frt
HS
Average LacZ copy # / rxn
3.53E7 ±1.69E7
2.32E7 ±1.84E7
3.93E7 ±2.83E7
I
P
P'
OR
SacI Restriction Site
J
loxp
frt
HS
L
white
blue
white
blue
white
blue
L

Figure 2.5: Disruption of both transgene ORF and spontaneous reversion are required to abolish leak expression

(A) Schematic of model AAV plasmid constructs designed to disrupt transgene ORF (AO plasmids) or reversion by recombination (NHS plasmids). Triangles denote 34bp shuffled *loxp* sites; triangles of the same shade are homologous. (B) HEK 293T line constitutively expressing mCherry and GFP under cre control reveal leak expression from plasmids expressing cre in a flp-dependent manner. Images show GFP (top row) and RFP (bottom row) channels of images of cell monolayers transfected 48hrs earlier with the plasmids listed above. Arrows in column two denote leak expression from the flp-dependent cre expression plasmid in the absence of flp. (C) Example flow cytometry results showing degree of cre-leak in HEK-mcherry-LSL-GFP cells across cre-expressing plasmids 48hrs after transfection. Controls (left column) show example GFP signal in empty vector and iCre transfection conditions. Dotted line denotes gate bounding GFP+ cells without any cells falling within it in empty vector condition. Percentage to the upper right of the doted gate signify percent of total shown cell population contained within the gate. Columns two and three show test conditions as outlined in (B) where the transgene is iCre. (D) Summary of leak across conditions as measured by percent of cells falling within the dotted gate. N = 6 transfections. iCre positive control omitted to maintain scale. Horizontal lines signify mean value across transfections (circles), vertical lines are standard deviation, and boxes 95% confidence intervals. (E) Example flow cytometry results showing degree of EYFP expression in HEK cells across EYFP-expressing plasmids 48hrs after transfection. Controls (left column) show example EYFP signal in empty vector and EYFP transfection conditions. Dotted line denotes gate bounding GFP+ cells drawn just above the bulk of cells in the empty vector condition. Percentage to the upper right of the dotted gate signify percent of total shown cell population contained within the gate. Columns two and three show test conditions as outlined in (B) where the transgene is EYFP. (F) Summary of leak across conditions as measured by percent of cells falling within the dotted gate. N = 6 transfections, except for NHS in which N = 5. EYFP positive control omitted to maintain scale. Horizontal lines signify mean value across transfections (circles), vertical lines are standard deviation, and boxes 95% confidence intervals.

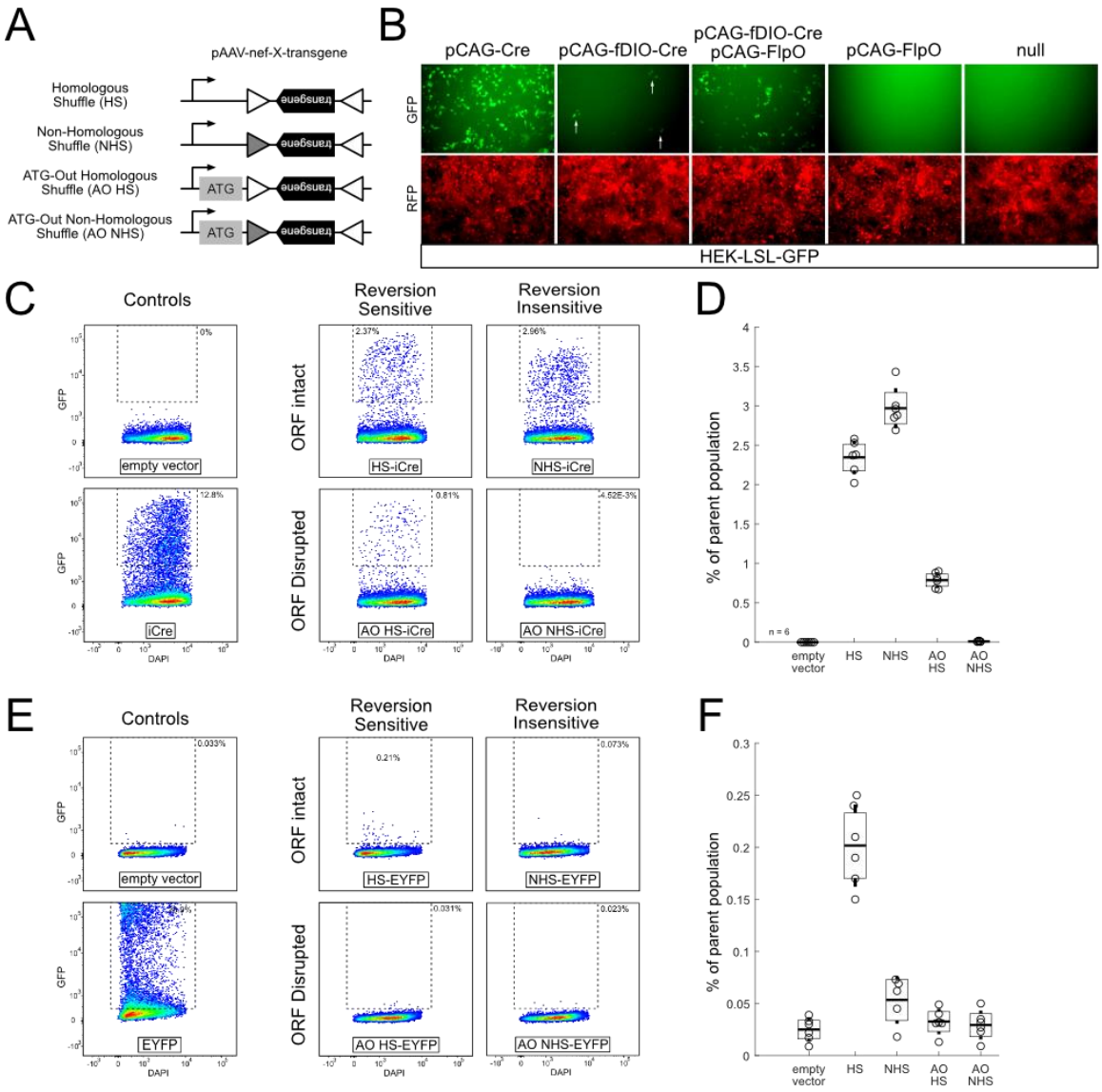
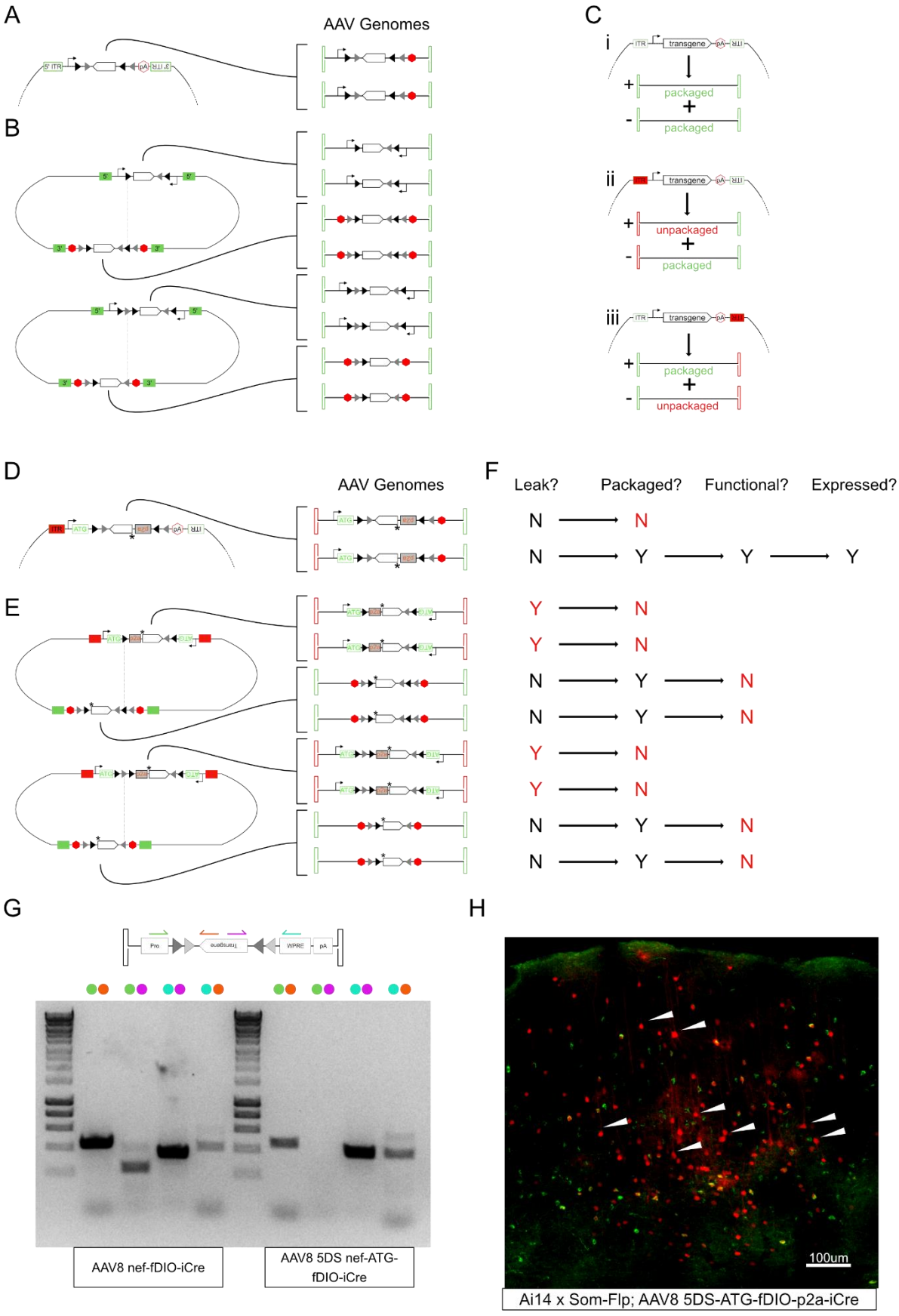

A
pAAV-nef-X-transgene
Homologous Shuffle (HS)
Non-Homologous Shuffle (NHS)
ATG-Out Homologous Shuffle (AO HS)
ATG-Out Non-Homologous Shuffle (AO NHS)
ATG
transgene
B
pCAG-Cre
pCAG-fDIO-Cre
pCAG-fDIO-Cre pCAG-FlpO
pCAG-FlpO
null
GFP
RFP
HEK-LSL-GFP
C
Controls
Reversion Sensitive
Reversion Insensitive
ORF intact
ORF Disrupted
empty vector
iCre
HS-iCre
NHS-iCre
AO HS-iCre
AO NHS-iCre
GFP
DAPI
0%
12.8%
2.37%
2.66%
0.81%
4.62E-3%
D
% of parent population
n = 6
empty vector
HS
NHS
AO HS
AO NHS
E
Controls
Reversion Sensitive
Reversion Insensitive
ORF intact
ORF Disrupted
empty vector
EYFP
HS-EYFP
NHS-EYFP
AO HS-EYFP
AO NHS-EYFP
0.033%
0.21%
0.073%
0.031%
0.023%
F
% of parent population

Figure 6: Disruption of 5' ITR D-sequence packaging signal blocks AAV packaging of genomes containing reverted transgenes but not leak expression

(A) Typical recombinase dependent AAV plasmid design and AAV genome. (B) Expected first-order recombinant plasmids and resultant AAV genomes resulting from recombination events across like recombinase-recognition sites. (C) Packaging profile of ssDNA AAV genomes from plasmids containing unmodified ITR D-sequences (i), 5' modified ITR (ii, in red), and 3' modified ITR (iii, in red). (D) Plasmid design and resulting AAV genomes from a recombinase-dependent AAV design with both a disrupted ORF and 5' ITR modified D-sequence, in red. (E) Expected first-order recombinant plasmids and resultant AAV genomes from the plasmid in (D). (F) Flowchart of each AAV genomes' propensity to leak, be packaged, and functionally express the transgene in the presence of recombinase. (G) PCR of AAV genomes using primers targeting all possible transgene orientations following recombination. Templates, listed at bottom, are AAV8 nef-fDIO-iCre and the ITR modified AAV8 5DS-nef-ATG-fDIO-iCre. (H) Example image of an AAV8-5DS-ATG-fDIO-p2a-iCRE injection in primary visual cortex of an Ai14 x SomFlp mouse. White arrows indicate off target non-somatostatin cre expression. Anti-somatostatin staining in green, tdTomato in red.


A
AAV Genomes
B
C
i
transgene
packaged
packaged
ii
unpackaged
packaged
iii
packaged
unpackaged
D
AAV Genomes
E
F
Leak?
Packaged?
Functional?
Expressed?
N N
N Y Y Y
Y N
Y N
N Y N
N Y N
Y N
Y N
N Y N
N Y N
G
AAV8 nef-fDIO-iCre
AAV8 5DS nef-ATG-fDIO-iCre
H
100um
Ai14 x Som-Flp; AAV8 5DS-ATG-fDIO-p2a-iCre