Analysis of the Dopaminergic Systems Development and Function by Targeted Genome Editing

Authors: Kristine Østevold
Submitted: 20. May 2015
Published: 20. May 2015
Keywords: Dopamine, tyrosine hydroxylase, TALENs, CRISPRs, neurogenesis, prepulse inhibition

Master Thesis

Analysis of the Dopaminergic Systems Development and Function by Targeted Genome Editing

written by:

Kristine Østevold

Albert-Ludwigs-Universität Freiburg

Director:

Prof. Dr. Wolfgang Driever
Institut für Biologie
Albert-Ludwigs-Universität Freiburg

Co-Director:

Dr. Med. Alejandro de Nicola
IBYME CONICET
Universidad de Buenos Aires

December 2013 - October 2014

Prof. Dr. Wolfgang Driever
Director

Dr. Med. Alejandro de Nicola
Co-director

Kristine Østevold
Student

Acknowledgements

First of all I would like to thank Prof. Dr. Wolfgang Driever for giving me the opportunity to do my Master Thesis work in his lab. He has been an excellent director and motivator and he was always full of good advice.

Thank you to Dr. Med. Alejandro de Nicola for co-directing my thesis.

For my everyday survival in the lab, I want to thank Theresa Schredelseker. She has been a wonderful teacher and colleague and I would not have made it without her.

In addition, I would like to thank Meta Rath and Martha Manoli who were always willing to help when I was in need and I am happy to have been able to share happiness and sorrows with my good colleague Florian Veit. For excellent technical help and support Dr. Aristides Arrenberg and Dr. Jörn Schweitzer also deserve a thank you.

Last, I would like to thank Dr. Thomas Auer for time, good advice, materials and help with the establishment of the CRISPR/Cas9 system.

Abstract

Dopamine is a prominent modulatory neurotransmitter affecting a vast array of neural circuits, however, dopamine has also been shown to be an important regulator of neurogenesis both in the developing embryo and in the adult mammalian brain. A more thorough understanding of dopamine and its role in neurogenesis will be pivotal to exploit endogenously produced neural precursors for cell replacement in neurodegenerative disorders such as Parkinson's Disease.

Two genes important for dopaminergic function are *tyrosine hydroxylase (th)*, the rate limiting enzyme in the synthesis of catecholamines, and the re-uptake carrier *dopamine transporter (dat / slc6a3)*. Using TALENs technology, mutations were induced in the *th* and *dat* genes of zebrafish embryos. A knockout mutation in the *th* gene is expected to cause the complete lack of dopamine and noradrenalin production in all catecholaminergic cells, whereas a knockout of the *dat* gene is expected to cause the accumulation of dopamine in the synaptic cleft. TALENs targeting exon one of *th* induced mutant alleles with a 7 bp (th^{m1403}) deletion, an 11 bp (th^{m1404}) deletion and a 1 bp (th^{m1405}) insertion close to the translation start, causing truncated proteins devoid of their native functions thus resulting in the knockout of the gene. TALENs targeting the *dat* gene induced a mutant allele with a 10 bp deletion which, however, was not expected to cause the knockout of the gene since the native start ATG was deleted and a second in frame ATG is present 20 codons downstream of the start codon. The CRISPR/Cas9 system was used to induce a mutation at a different position of exon one of the *dat* gene and genotyping of the F_1 generation will determine if this has lead to the knockout of the gene.

Homozygous *th* mutants were analyzed for catecholaminergic neuron development in the brain and motor neuron development in the spine. No major differences were found in neither motor neuron cell numbers in the spine nor in catecholamine cell number in the brain when comparing $th^{-/-}$ and wildtype larvae at 3 days post fertilization. These results indicate that dopamine does not play a pivotal role in the generation of these neurons during embryogenesis, and question previous publications reporting changes in motor neuron number following pharmacological or antisense morpholino based manipulation of dopaminergic function. $th^{-/-}$ zebrafish larvae were additionally analyzed for prepulse inhibition responses, and preliminary data was consistent with previous publications suggesting that dopamine may modulate this behaviour.

Genome editing techniques were in addition employed to enable the targeted insertion of transgenes. The CRISPR/Cas9 system was used to establish the protocol and its functionality was demonstrated by the insertion of a *gal4* transgene at the eGFP locus of Tg(NeuroD:eGFP) x Tg(UAS:RFP, cry:eGFP) zebrafish embryos, driving red fluorescent expression in 1 of 6 of injected embryos. In the future this system can be used to establish reliable transgenic lines with a *gal4* at the *th* and *dat* loci driving expression of active proteins that may control axonogenesis of dopaminergic neurons, an interesting aspect useful for potential new regeneration therapies, or driving expression of neuroprotective proteins in order to attenuate neurodegenerative processes in zebrafish models of Parkinson's disease.

Abbreviations

aa	amino acid
AADC	Aromatic L-amino acid decarboxylase
ADHD	Attention deficit hyperactivity disorder
Amp	Ampicillin
bp	base pairs
CRISPR	Clustered regularly interspaced short palindromic repeats
DA	Dopamine
DBH	Dopamine beta hydroxylase
DNA	Deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
dpf	days post fertilization
dsDNA	double stranded DNA
eGFP	enhanced green fluorescent protein
gDNA	genomic DNA
GFP	Green fluorescent protein
GG1	Golden gate 1
GG2	Golden gate 2
hpf	hours post fertilization
gRNA	guide RNA
Inj	Injection
IHC	Immunohistochemistry
Kan	Kanamycin
kb	kilo base pairs
LLC	Long-latency C-start
mRNA	messenger RNA
NHEJ	Non homologous end joining
NLS	Nuclear localization signals
No	Number
ORF	Open reading frame
PCR	Polymerase chain reaction
PNK	Polynucleotide kinase
PPI	Prepulse inhibition
PTU	Phenylthiourea
RFP	Red fluorescent protein
RNA	Ribonucleic acid
RVD	Repeat-variable di-residue
Rxn	Reaction
SGZ	Subgranular zone
Shh	Sonic hedgehog
SLC	Short-latency C-start
Spec	Spectomycin
SVZ	Sub ventricular zone
TALEN	Transcription activator-like effector nuclease
Tm	Melting temperature
VTA	Ventral tegmental area
WISH	Whole mount in situ hybridization
WT	Wildtype

Contents

1 Introduction

1.1 The Brain and Dopaminergic Systems

The human brain is a marvellous organ. It controls all of the body's diverse and numerous functions, including our abilities to feel, think, reason, learn, plan ahead and act. The brain is also where our consciousness and our sense of "self" arise. What forms this intriguing organ are primarily two types of cells, neurons and glial. An adult human brain has around 100 billion neurons, which is just about as many stars as in the Milky Way Galaxy (Blinkov and Glezer, 1968). Neurons are signalling cells able to communicate with specific targets over short or long distances while glial cells, which are as many in number as neurons, have mostly supporting functions in the brain. Neurons can be classified according to which molecule they use to transfer their signals. These signal molecules are called neurotransmitters and can be amino acids, monoamines, peptides or other chemicals. One of the monoamine neurotransmitters is called dopamine (DA), and the neurons transmitting DA are called dopaminergic neurons.

Dopaminergic neurons project throughout the brain creating several different pathways which again form defined systems (*Fig. 1*). The main function of the dopaminergic systems is to modulate and change the behavior of other neurons and neural circuits. Neuromodulators do not directly excite or repress electrical activity in target neurons, however, they affect how the target neurons respond to direct neurotransmitters. In the nigrostriatal pathway, one of the best known dopaminergic system, dopaminergic neurons help regulate voluntary movement by modulating neurons in the basal ganglia motor loop. Further dopaminergic systems regulate other complex functions of the central nervous system such as cognition, motivation, reward, mood, pain, attention and learning. These functions are all essential for our everyday function and well being. A disruption of the different dopaminergic systems can therefore have severe consequences both for humans and animals.

1.2 Diseases of the Dopaminergic Systems

The most widely known illness of the dopaminergic systems is Parkinson's disease, which is the second most common neurodegenerative disorder after Alzheimer's (Shulman et al., 2011). Neurodegeneration is the progressive loss of structure or function of neurons, including their death. For Parkinson's disease this happens predominantly in the substantia nigra, which is the starting point of the nigrostriatal pathway. The neuronal loss in the substantia nigra affects motor control and can give rise to tremors and bradykinesia (slowness of movement), some of the most prominent symptoms of Parkinson's disease.

Dopaminergic dysfunction has also been related to several other diseases, including schizophrenia, attention deficit hyperactivity disorder (ADHD) (Volkow et al., 2011; Madras et al., 2005), depression and addiction. Symptoms of schizophrenia, especially psychoses, are believed to be caused by a disrupted and hyperactive dopaminergic signal transduction (Seeman, 2007). This disruption causes overactivation of DA receptors in the limbic system, which appears to be the primary system responsible for our emotional life. This "dopamine hypothesis" of schizophrenia is also supported by the fact that a large number of antipsychotic drugs have DA-receptor antagonistic effects.

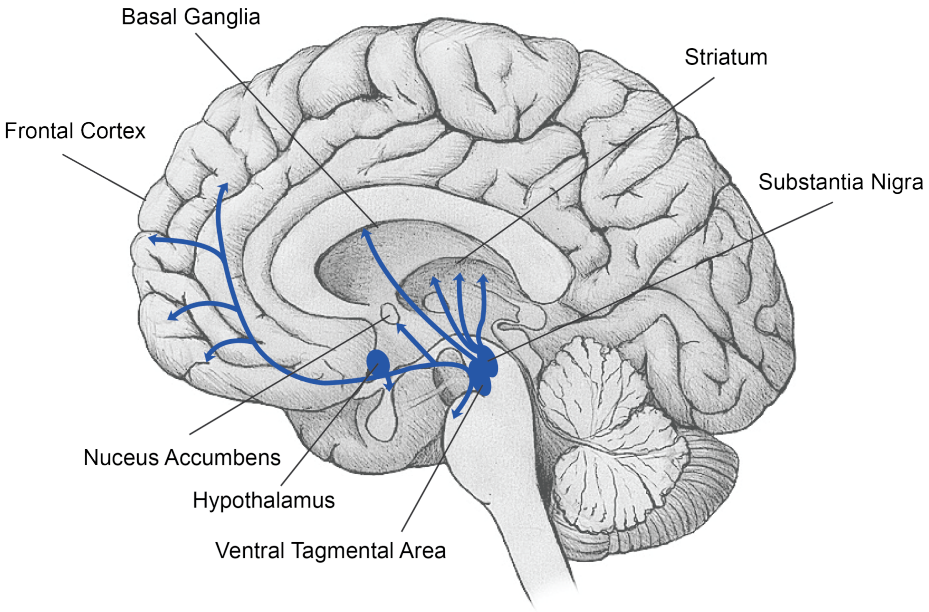

Figure 1: The major dopaminergic pathways of the brain are the mesolimbic pathway which transmits DA from the ventral tegmental area (VTA) to the limbic system, the mesocortical pathway which transmits DA from the VTA to the frontal cortex, the nigrostriatal pathway which transmits DA from the substantia nigra to the striatum and the tuberoinfundibular pathway which transmits DA from the hypothalamus to the pituitary gland influencing the secretion of certain hormones (Blinkov and Glezer, 1968).

To become better equipped to tackle diseases of the dopaminergic systems, more knowledge need to be gained on the dopaminergic systems development and functions. So far, no medical treatment is available neither to cure nor to stop the progression of Parkinson's disease, and for treating the symptoms of schizophrenia, effective drugs without hindering side effects are scarce. Conventional Parkinson's therapies aim to replace the progressive loss of DA by the administration of DA agonists or precursors, however, over time the ongoing destruction of dopaminergic neurons overwhelms and alleviation by drugs is limited. For this reason, current research aspire to develop cell replacement therapies, replacing the dopaminergic cells which were lost with new healthy cells taking over the vital production of DA. In this aspect, it is especially crucial to understand how dopaminergic neurons and systems first develop to eventually be able to recapitulate DA cell development in vitro.

1.3 Dopamine and Development

The understanding of how organisms develop has become essential for the understanding of organ function and many areas of biomedical research. When studying genetic control of cell growth, differentiation and morphogenesis, a powerful insight into the processes that give rise to tissues, organs and anatomy is obtained. This in turn can provide important knowledge on disease pathogenesis and competence to find new strategies for medical treatments.

One of the primary events in early embryonic development is gastrulation, where the until then single layered embryo (blastula) is reorganized into a three layered structure

(gastrula). Following gastrulation organogenesis begins and from a specialized region of the outer layer, the ectoderm, the central nervous system develops. What makes a cell decide to become a nervous tissue cell and not for example a skin cell depends on external and internal cues which orchestrate neurogenesis in both a temporal and spatial manner. Using modern molecular biology it has been made possible to identify a number of complex molecular networks and signalling cascades regulating the induction and proliferation of neural progenitor cells and their subsequent differentiation into mature neurons.

For dopaminergic neurons in the midbrain, Sonic hedgehog (Shh), fibroblast growth factor 8 (Fgf8) and int/Wingless family member 1 (Wnt1) create gradients along both the rostral-caudal and the dorso-ventral axes which act as differentiation signals (Allodi and Hedlund, 2014; Wang et al., 1995). Later in development, Foxa2 (a forkhead transcription factor) which is induced by Shh, is essential for the further generation of dopaminergic neuron progenitors. Many other transcription factors have also been identified both in DA progenitors and mature DA neurons and they have been shown to be activated at different stages in development (*Fig. 2*). Although a wealth of information has been obtained on DA neuron development, and neurons expressing dopaminergic markers have been successfully cultured in a Petri dish, these are not stable when transplanted in vivo, showing that the picture of dopaminergic development is not complete (Swistowski et al., 2010; Bjorklund et al., 2002).

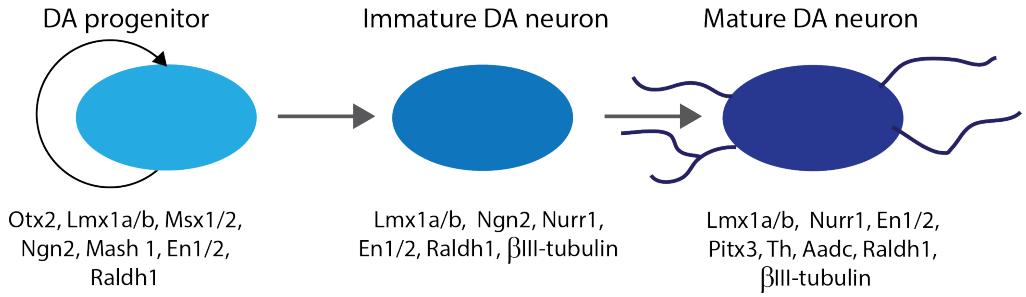

Figure 2: The sequential timing of transcription factor activation in midbrain DA progenitors (Simon et al., 2001; Puelles et al., 2004; Andersson et al., 2006; Kele et al., 2006). (Picture modified from Ang (2006)).

In recent years, it has become clear that neurogenesis not only takes place in the developing embryo, but also in the adult vertebrate brain (Eriksson et al., 1998). Research show that mature cells in all neural lineages are generated and integrated into neural circuits throughout adulthood. This adult neurogenesis takes place in at least two distinct regions of the forebrain, the sub-ventricular zone (SVZ) and the subgranular zone (SGZ) of the hippocampus (Alvarez-Buylla and Garcia-Verdugo, 2002; Kaplan and Hinds, 1977). It is believed that the new neurons generated here are important for olfaction, mood regulation and memory processes in mammals (Nilsson et al., 1999). Additionally, a recent publication showed that new neurons are also added to the striatum of the adult human brain (Ernst et al., 2014). The striatum, as mentioned earlier, is the target of the dopaminergic nigrostriatal pathway which is heavily affected in Parkinson's disease. The discovery of adult neurogenesis here might have large implications for future treatment strategies as it opens the possibility to force neuronal precursors in the striatum into a dopaminergic fate. Normally, no dopaminergic neurons are present in the striatum,

however the striatum receives dopaminergic input from axons projected by dopaminergic cells localized in the substantia nigra. Therapies aiming to replace the cells in the substantia nigra would face the challenge of correct axon guidance into the striatum, a problem avoided by placing the dopaminergic cells there directly.

Adult neurogenesis, like embryonic neurogenesis, is highly regulated. Interestingly, the neurotransmitter DA has itself been indicated as an important player in this regulation. This means that DA has a dual role in our body, both as a neurotransmitter and as a regulator controlling neural stem cell niche activity (Takamura et al., 2014; Hoglinger et al., 2014). DA receptors are found on both multipotent and restricted neuronal progenitor cells in the SVZ and SGZ, and DA seems to stimulate adult nerurogenesis here. DA receptors are also greatly abundant in highly proliferative zones of the brain during development, suggesting a regulatory role for DA in embryonic neurogenesis as well (Diaz et al., 1997).

Supporting the theory that DA plays a role in neurogenesis, a paper was published last year strongly indicating DA involvement in spinal motor neuron generation and regeneration in zebrafish (Reimer et al., 2013), both in the developing embryo and in the adult lesioned spine. The motor neuron generation was shown to be promoted at the expense of V2 interneurons, both of which progenitors express the transcription factor *lim3*. The homeobox gene *Hb9 ; mnx1* is required for the down-regulation of *lim3*, something which enhances motor neuron generation over V2 interneurons (Arber et al., 1999). *hb9* is also claimed to be needed for the continued expression of *isl1* in postmitotic motor neurons (*Fig. 3*). *isl1* is a LIM homeodomain transcription factor which is required for the survival and specification of spinal motor neurons. Both *hb9* and *isl1* are used as specific markers for motor neurons in the developing embryo. Reimer et al. (2013) analyzed *hb9* and *isl1* positive cells under different forms of DA depletion, and they suggest that the role of DA in the promotion of motor neurons over V2 interneurons is elicited through the D4a receptor which activates the major signalling pathway in developmental generation, the hedgehog pathway (*Fig. 4*). Moreover, the authors claim that the same mechanisms for the promotion of spinal motor neuron generation during development seems to be reactivated in the generation of motor neurons in the lesioned spinal cord of adult zebrafish.

To further study the role of DA in development and neurogenesis, two genes in particular are of great interest, namely *tyrosine hydroxylase (th)* and *dopamine transporter (dat ; slc6a3)*. The protein encoded by *th* is an enzyme that catalyses the first and rate-determining step in the formation of catecholamines (*Fig. 5*). Catecholamines are monoamines, and the most abundant ones in the human body are epinephrine/adrenaline, norepinephrine/noradrenaline and DA. *th* is expressed in all catecholaminergic neurons, including the dopaminergic neurons. Dat is a membrane transport protein which transports DA from the synaptic cleft and back into the presynaptic boutons and thus terminating DA action (*Fig. 6*). *dat* is expressed in all dopaminergic neurons.

1.4 Targeted Knockout Mutations

A basic approach to find out more about the function of a certain gene and its potential role in development is to find out what happens when the gene is no longer active. One way of doing this is to create a knockout organism, where the gene in question is modified

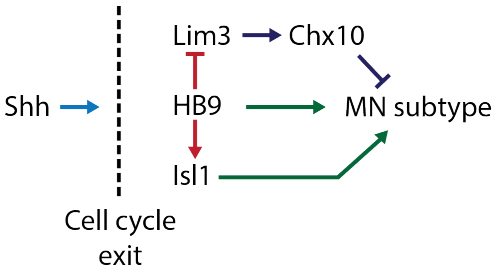

Figure 3: *hb9* down-regulates *lim3*, which is needed for expression of *chx10*, a marker for V2 interneurons. At the same time *hb9* upregulates *isl1*, a transcription factor which enhances motor neuron (MN) generation. (Picture modified from Arber et al. (1999).)

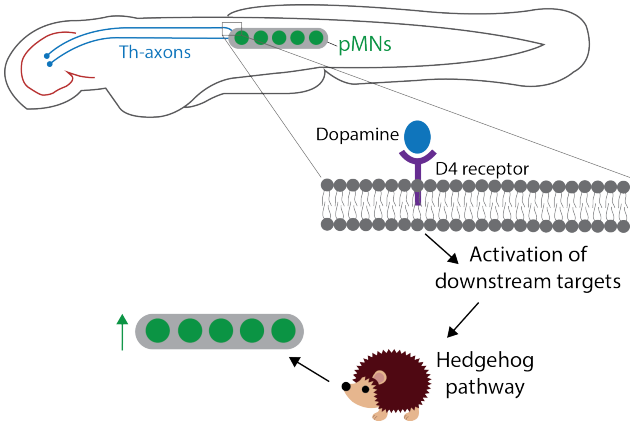

Figure 4: Dopamine from the brain promotes motor neuron generation through the D4a receptor which elicits effects on the hedgehog pathway.

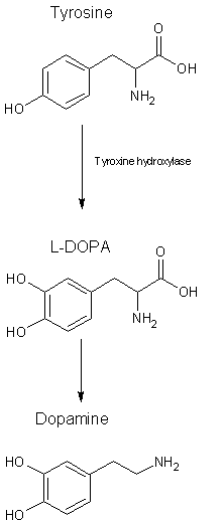

Figure 5: Tyrosine hydroxylase catalyses the rate determining step of dopamine and catecholamine formation, transforming tyrosine to L-DOPA.

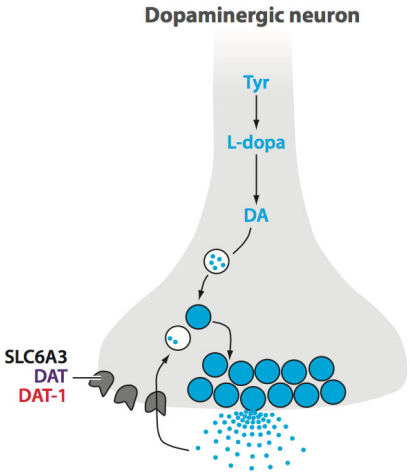

Figure 6: The dopamine transporter transports dopamine from the synaptic cleft back into the neuron, thus terminating dopamine action (picture modified from (Flames and Hobert, 2011).

to encode a non-functional protein. The first *Th* and *Dat* mouse knockout models were developed around 20 years ago.

1.4.1 *Th*$^{-/-}$ Mice Models

The targeted disruption of the *Th* locus in mice resulted in severe catecholamine depletion and perinatal lethality (Kobayashi et al., 1995) (Zhou and Palmiter, 1995) (Zhou et al., 1994). The histopathological changes found in the homozygous mutants included some blood congestion, bradycardia and unusual histology of the cardiomyocytes, suggesting that catecholamines are necessary for proper cardiovascular function. The defects of the cardiovascular system were indicated as the primary cause of death during prenatal and

neonatal stages. When pregnant females were treated with L-DOPA, the product of Th enzymatic activity, all *Th*$^{-/-}$ mutants were rescued to birth, further sustaining the statement that catecholamines are essential for fetal survival.

To be able to better study the consequences of Th deficiency on the development of the catecholaminergic systems, another mouse model was developed where the expression of the *Th* gene was restored in noradrenergic cells. This was done by targeting the *Th* coding sequence to the noradrenergic-specific *Dopamine-β hydroxylase (Dbh)* promoter, giving mice that were only DA deficient, but had the ability to produce adrenaline and noradrenaline (Zhou and Palmiter, 1995). The *Dbh-Th* transgene prevented perinatal lethality, indicating that DA function is not required before birth. *Dbh-Th, Th*$^{-/-}$ mice appeared normal until postnatal day 10-15 where they began to show hypoactivity and became significantly smaller than their littermates. However, they showed normal reflexes, including the startle reflex. By four weeks of age, all homozygous mutants died, indicating that DA function is important for post-natal survival. Conversely, the same paper showed that DA did not seem to have an effect on DA containing pathways or normal neurogenesis, seeing as the brains of the DA deficient mice apparently had developed normally. AADC (the enzyme that converts L-DOPA to DA) positive cells could be found were DA neurons normally exist and no other anatomical changes in the brains were found.

In a different research group, cathecholamine levels in *Th*$^{-/-}$ mice were measured, both at stage E12.5 and at birth, determining the effect of loss of *Th* on catecholamine metabolism (Kobayashi et al., 1995). No detectable or very low levels were found of adrenaline and noradrenaline, however interestingly, DA was discovered both in the body and in the head of the *Th*$^{-/-}$ mice with a level of 37.5% (head) of normal values at E12.5 and 41.7% (body) of normal values at birth.

1.4.2 *Dat*$^{-/-}$ Mice Models

The creation of *Dat* knockout mice showed that the loss of the DA transporter makes DA linger in the synaptic cleft for around 300 times longer than normal, and its diffusion distance becomes ten times higher than in wildtype mice (Benoit-Marand et al., 2000). This elongates and enforce DA action, and *DAT*$^{-/-}$ mouse models show hyperlocomotion as one of several phenotypes (Giros et al., 1996). So far, the *Dat*$^{-/-}$ mice have not been used in greater extent to study the effect of DA on the development of the central nervous system, however, model organisms lacking the DA transporter have been popular tools to study ADHD, a disorder which is characterized by inattention, impulsiveness and hyperactivity (Leo and Gainetdinov, 2013).

1.5 Prepulse Inhibition

As earlier described, DA and the dopaminergic systems play important roles in many physiological functions, one of which is called prepulse inhibition (PPI) of the startle response. In PPI, a weaker startle stimulus (prepulse) has the ability to reduce the startle response to a stronger startle stimulus (*Fig. 7*)(Hoffman and Ison, 1980). PPI tests can be used as a quantitative and clinically relevant measure of the brains ability to preform sensorimotor gating (Burgess and Granato, 2007). Sensorimotor gating is a measure to control behaviour in a natural environment where sensory stimuli are plentiful, and it

is partly accomplished through selective transmission, or gating, of signals going to the motor systems.

Many persons with the psychiatric disorder schizophrenia show a marked impairment in their ability to perform sensorimotor gating and animal model studies demonstrate that an increased amount of DA in the nucleus accumbens can cause sensory gating failure similar to that seen in schizophrenic patients (Braff and Geyer, 1990). Pharmacological treatment of mice with DA agonists such as amphetamine also gives a decrease in PPI (Zhang et al., 2000), while DA antagonists have been shown to enhance PPI, supporting the claim that DA neurotransmission is involved in the modulation of PPI. *th* and *dat* knockout animal models could be useful to further study the role of DA in PPI on a genetic level, something which could help unravel some of the fundamental questions regarding schizophrenia (Braff et al., 1992).

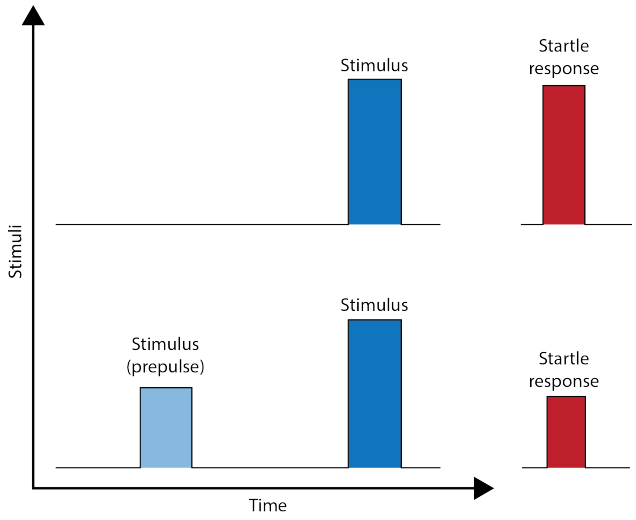

Figure 7: In prepulse inhibition, a weaker startle stimulus (prepulse) has the ability to reduce the startle response to a stronger startle stimulus (Hoffman and Ison, 1980).

1.6 Knock-in Lines

The knock-in of an exogenous DNA sequence, an ORF, or a whole gene at a particular locus in an organisms genome can be of great value for research on gene expression and function. In theory, any sequence or gene can be knocked in at any locus of the genome. For research, transgenes that aim to visualize expression or to provide different functions in specific cell types are particularly useful. Fluorescent proteins can be knocked in at a gene locus of interest, specifically labelling it and making expression patterns, morphology and cell migration easily observable using epifluorescence microscopy (Feng et al., 2000). By knocking in transcriptional activators or enhancers it is possible to control gene expression, something which is also very practical in the study of genes. A very potent system for this is called the GAL4:UAS system (*Fig. 8*) (Brand and Perrimon, 1993). In this system, lines are created by the knock-in of the transcriptional activator *gal4* (yeast transcription activator protein) which is expressed in a tissue specific way or in just a subset of a tissue. GAL4 lines are then crossed to reporter lines which have an enhancer, called GAL4-UAS (Upstream Activation Sequence), engineered to drive ex-

pression of a functional protein. This will result in the activation of UAS:target in the cells that produce Gal4. With a *gal4* at the *th* or *dat* locus, the expression of genes in dopaminergic cells can be controlled using different UAS reporter lines. In the future, this system can potentially be used to control axonogenesis of these cells, an interesting aspect which can be useful for possible new regeneration therapies. It can also be achievable to drive expression of neuroprotective proteins and then evaluate their effect under conditions of dopaminergic neurodegeneration.

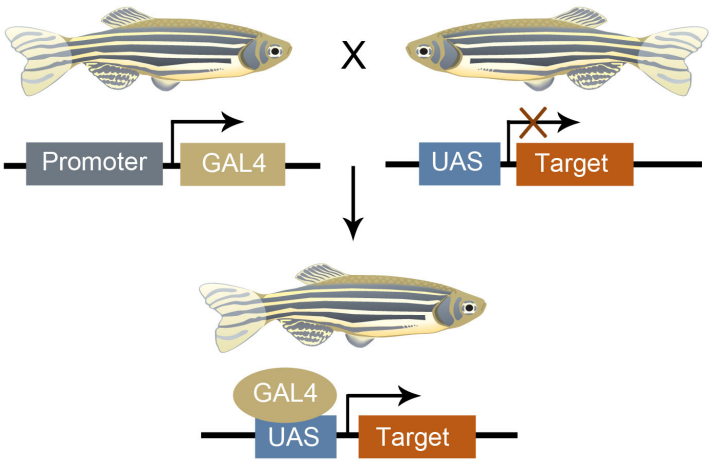

Figure 8: In the GAL4:UAS system, the transcriptional activator *gal4* is expressed in a specific way and has the potential to drive the expression of UAS regulated targets by binding to the UAS enhancer.

As technologies develop, the desire for new and diverse knock-in lines grows. In the field of optogenetics for example, where light is used to control neuron activity, the knock-in of different light sensitive proteins is crucial, and to measure dopaminergic neuron activity based on Ca^{2+} release patterns, GCaMPs (a genetically encoded calcium indicator) must be knocked in. As the list of wanted genes for knock-in at the same locus gets longer, special knock-in technologies may come in handy such as the PhiC31 integrase vector system. PhiC31 is a site-specific recombinase originally present in the PhiC31 phage. The integrase recognizes binding sites termed attB (attachment site Bacterium) and attP (attachment site Phage), and the completed recombination results in sites (attL and attR) incompatible with the PhiC31 integrase, making the integration reaction irreversible (*Fig. 9*) (Groth et al., 2000). Once an attP landing site has been knocked into the genome, the PhiC31 integrase can be used to easily insert any wanted DNA fragment flanked by attB sites, making it a perfect method for the knock-in of different genes at the same locus.

1.7 Targeted Genome Editing

When creating specific knock-in or knockout organisms, the ability for targeted genome editing is essential. Lately, two new promising systems for this have been developed named TALENs and CRISPRs.

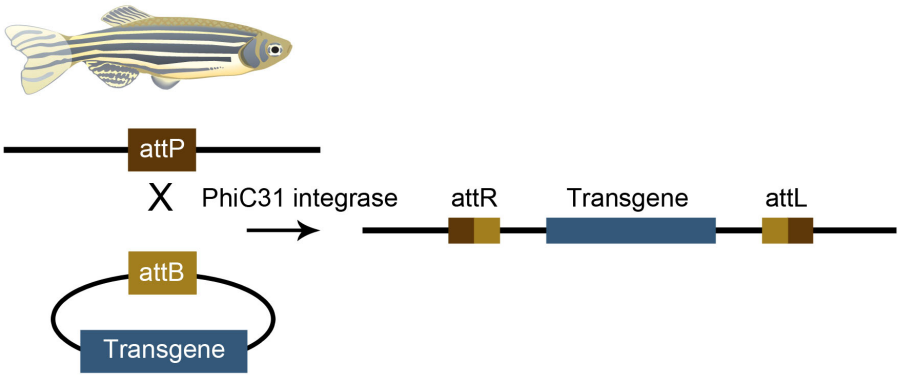

Figure 9: The PhiC31 integrase can be used to irreversibly knock in genes into genomes with an attP landing site present.

1.7.1 TALENs

Transcription activator-like effector nucleases (TALENs), are DNA binding proteins engineered to create targeted double strand breaks (DSB). One of the two major building blocks of a TALEN is the TAL (transcription activator-like) effectors which are derived from the plant pathogenic bacteria *Xanthomonas spp* (Joung and Sander, 2013). TAL effectors, originally used by the bacteria to aid in the infection of plants, contain a central domain of tandem repeats which recognize and bind to DNA flanked by nuclear localization signals (NLS) and an acidic transcriptional activation domain (AD) at the C-terminal end (*Fig. 10*). The different motifs enable the TAL effectors to bind to promoter sequences and modulate host gene expression through the activation of transcription (Cermak et al., 2011).

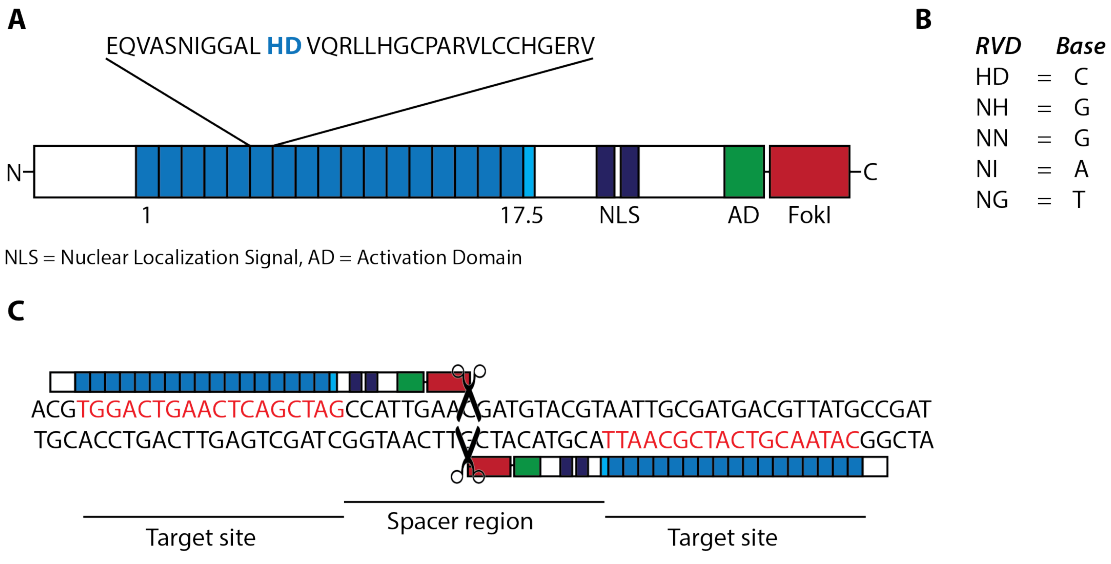

Figure 10: TALEN structure and function. **A:** TAL effector domain with a FokI nuclease domain. **B:** RVD specificity. **C:** TALENs work in pairs and the two FokI nucleases come together in the spacer region creating a double strand break.

Target specificity of TAL effectors originates from the central domain which consists of

12 to 27 highly conserved repeated domains of 33 to 35 amino acids each, with the exception of the last repeat which is truncated and only consist of the first 20 amino acids (Weber et al., 2011). Within each repeated domain, amino acid 12 and 13 are highly variable and together they are called a repeat-variable di-residue (RVD). Computational and functional analyses have revealed that the RVDs hold the key to where on the DNA a specific TAL effector will bind, as each RVD recognizes and binds to a specific nucleotide (Miller et al., 2010). Discovering the TAL effector code has made it possible to design and engineer TAL effectors that bind to practically any place on the DNA by choosing and combining repeats with the appropriate RVDs. However, for the TAL effector proteins to serve as a tool in genetic engineering, they also need a nuclease domain, which is the second major building block of a TALEN.

The TALEN nuclease domain consists of the cleavage domain of a FokI restriction endonuclease. FokI is special in the way that it cleaves double stranded DNA non-specifically. Thus, after binding, it can cleave any DNA sequence 9/13 nucleotides downstream of the binding site (Wah et al., 1998). TALENs work in pairs, each binding to its own DNA strand at each side of the wanted cutting site, called the spacer region. Here, the two FokI nucleases come together to form a dimer and create a DSB. All in all this means that a TALEN can be designed to target any sequence of the DNA and create a DSB at any desired site.

Golden Gate TALEN assembly Golden Gate cloning is a method for the assembly of many DNA fragments, from a number of different parental plasmids, in one single step (Engler et al., 2009). The method is based on type IIs restriction enzymes and restriction ligation. Type IIs restriction enzymes have the ability to cut outside of their recognition site. Thus, two DNA ends flanked by said restriction site can be designed in such a way that after digestion, the restriction site is removed and complementary 4nt overhangs are created. These ends can then be ligated together flawlessly without the presence of the original restriction site. To achieve a certain order of the DNA fragments in the final construct, one can simply give the DNA ends different 4nt overhangs, each complementary to the following fragment. It has been shown that the Golden Gate cloning can efficiently assemble up to ten fragments in an orderly fashion into a target vector, making it an excellent method for building TALENs.

To create a TALEN it is essential to combine a specific set of RVDs in an explicit order. This can be done using the Golden Gate cloning method, assembling plasmids each containing one RVD for one specific position in a TALEN (*Fig. 11*) (Cermak et al., 2011). These plasmids are flanked by a type IIs restriction enzyme site, and when cut the overhangs created decide the order in which they will be ligated together in the target vector. Since TALENs are usually longer than 10 RVDs, and thus needs the assembly of more than 10 plasmids, they are made in two separate cloning steps. In the first step, TALEN RVDs are assembled into two different intermediate vector backbones, pFUS_A and pFUS_B respectively. These vectors are flanked by the same restriction site as the RVD plasmids, with overhangs complementary to the first and the last of the RVDs. In the last cloning step, the two intermediate vectors are combined together with the last RVD into the final target vector backbone. This restriction ligation is done by a second type IIs restriction enzyme with recognition sites present in the intermediate vectors and the final backbone.

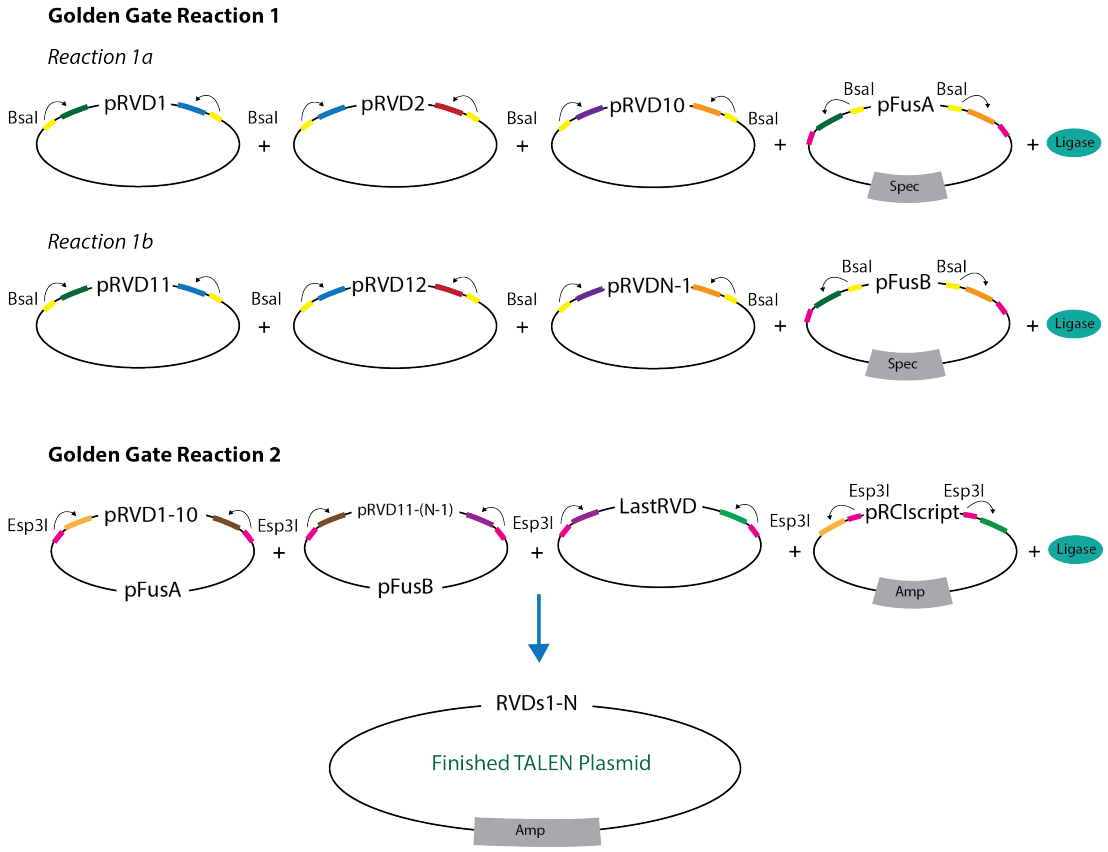

Figure 11: The Golden Gate TALEN assembly can be used to combine specific RVDs in an explicit order for the creation of any TALEN using type II restriction enzymes and ligation.

1.7.2 CRISPR/Cas Systems

The newly developed clustered regularly interspaced short palindromic repeat (CRISPR) / CRISPR-associated (Cas) system is also used to create DSB of the DNA, and thus edit the genomes of diverse organisms (Mali et al., 2013). The system is based on small RNAs and a Cas9 nuclease which detect and silence foreign nucleic acids sequence-specifically. It is originally derived from bacteria and archea which use the system as part of their adaptive immunity.

Bacteria and archea incorporate short invading phage or plasmid DNA fragments (protospacers) into CRISPR loci, which consist of short palindromic repeats and short interspaced regions (*Fig. 12*). The foreign DNA is integrated into the spacer regions and when the whole CRISPR locus is transcribed, the CRISPR RNAs (crRNA) serve as guides to identify foreign DNA and target it for nuclease-mediated cleavage by the Cas9 nuclease. This silences the foreign DNA, consequently providing adaptive immunity (Hruscha et al., 2013). For the Cas9 to cleave dsDNA site-specifically it requires a complimentary base-pared structure formed between the targeting crRNA and the target protospacer DNA, as well as a short protospacer adjacent motif (PAM) located right after the complementary region in the target DNA. So far, the CRISPR/Cas9 system is the only genome engineering tool described that is based on base pairing, and not the possibly less specific protein-DNA interactions, making the system very attractive.

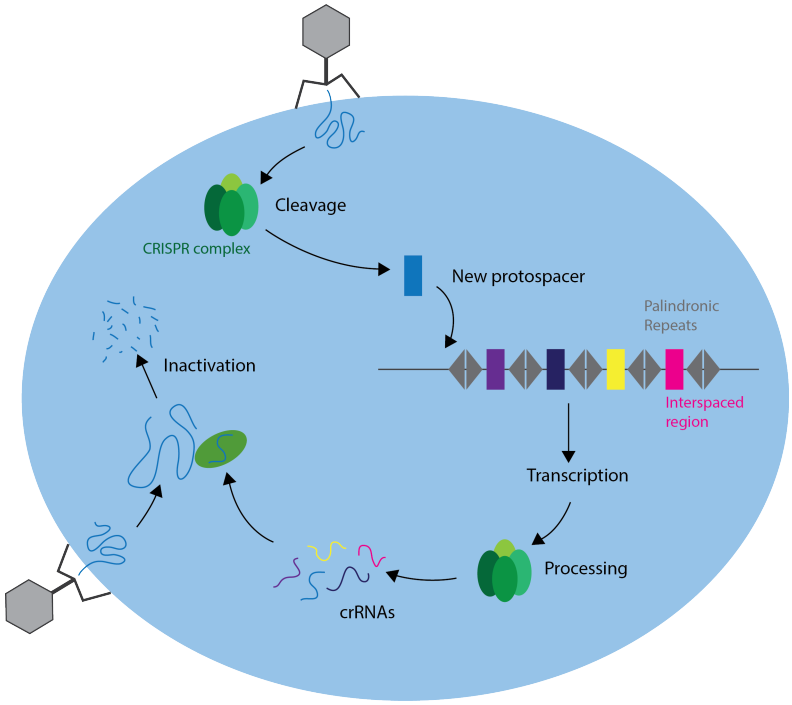

Figure 12: An overview of the immunization and immunity process of the CRISPR/Cas system (picture inspired by Horvath and Barrangou (2010))

In the lab, the CRISPR/Cas9 type II system of *Streptococcus pyogenes* has been modified and optimized for use in vertebrate cells to edit genomes (Mali et al., 2013)(Horvath and Barrangou, 2010). Using the same Cas9 and providing it with different crRNAs, in the lab called guide RNAs (gRNA), DSBs can be induced at basically any location in the genome (*Fig. 13*). It has also been shown that the system can be used for targeted knock-in of transgenes (Auer et al., 2013).

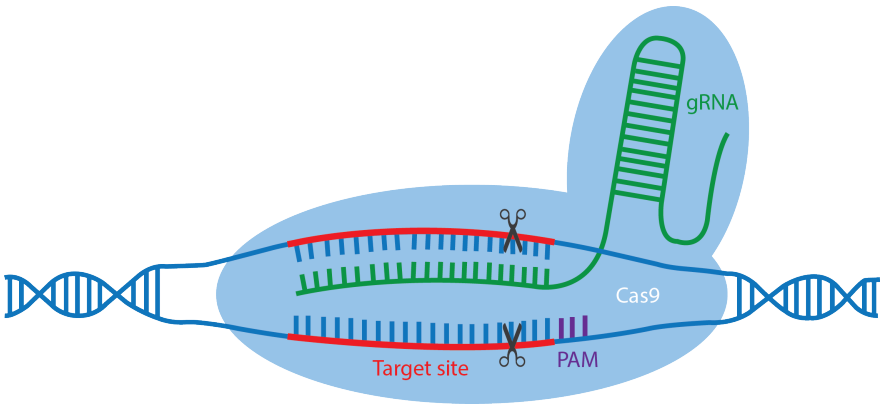

Figure 13: The CRISPR/CAS type II system of *Streptococcus pyogenes* has been modified and optimized for use in vertebrate cells to edit genomes.

gRNAs consist of two parts, the first of which is 20 nucleotides long and is used to target a specific sequence of DNA. The most restrictive requirement when deciding on a DNA target site is that it must be followed by a PAM sequence. Recent studies addi-

tionally indicate that the rest of the target sequence can have a significant effect on the DSB inducing efficiencies of the gRNA and that it is beneficial with a high GC content and having G as the last base of the gRNA to maximize efficiencies (Gagnon et al., 2014).

The second part of the gRNA is a constant and is equal for all gRNAs independent of their targets. gRNAs can be easily made by cloning the 20 nucleotide target specific DNA sequence into a gRNA expression vector backbone containing the constant part of the gRNA and a promoter for transcription (Hwang et al., 2013). The gRNA can then be transcribed in vitro using an RNA polymerase. As a non cloning alternative, the gRNAs can also be made annealing two oligonucleotides together, one containing a promoter, the target specific region and an overlap region and the other containing the overlap region and the constant region. After filling in the strands with a polymerase, in vitro transcription can be performed (Gagnon et al., 2014).

Regarding optimal injection amounts of Cas9 mRNA and gRNA, the literature have split opinions. Functional CRISPRs have been reported after injection into zebrafish embryos of as little as 0.15ng Cas9 mRNA and 0.0125ng gRNA and as much as 0.6ng Cas9 mRNA and 0.4ng gRNA. This breach in suggested amounts for injection is not that surprising considering the novelty of the CRISPR/Cas9 technique. Time and thorough experience is needed to find the optimal conditions, not only for injection amounts, but for all parts of the system.

1.7.3 Mutagenesis Mediated by CRISPRs and TALENs

Following a double strand break created either by TALENs or CRISPRs, a cell will automatically initiate its DNA repair machinery, ligating the two pieces of DNA back together. This can happen either by homologous recombination or, more frequently, by non homologous end joining (NHEJ) (Mao et al., 2008). NHEJ is an error prone process which often causes misrepair of the genome and gives rise to in/del mutations where either a small sequence of the DNA is lost or gained (Heidenreich et al., 2003). In two of three cases, this will lead to a frameshift mutation and thus to the disruption of gene function and a knockout organism. Moreover, the DNA-repair machinery can also be taken advantage of in the creation of knock-in mutations. If another piece of DNA is presented to the cell, either in form of a donor plasmid, a PCR product or a short single stranded oligonucleotide, the cell can be "tricked" to insert this sequence at the site of the double strand break during the repair. This process can as well be mediated both by homologous recombination, if the donor DNA have homology arms with the DNA around the cut site, or by NHEJ.

CRISPRs and TALENs have become extremely popular over the last few years due to their potency to easily edit the genome of a diverse range of organisms, including the zebrafish.

1.8 Zebrafish as a Model Organism

The zebrafish (Danio rerio) has in later years become a powerful animal model to study vertebrate development and disease related processes (Mosimann et al., 2013). As an aquatic non-mammal the zebrafish differs a lot from humans, however, their brains also

show many similar arrangements to that of mammals. Over the years, many essential mechanisms of the central nervous system have been enlightened using the zebrafish as a model organism, and many of these principles can be transferred to more complex organisms including humans. Furthermore, there has been developed several approaches for modifications of the zebrafish genome giving rise to a range of different zebrafish lines.

The zebrafish growing popularity is due to several practical qualities, including its small size, a short generation time and the complete extrauterine development of the embryos. Furthermore, the developing embryos are transparent, making them ideal for in vivo imaging of early developmental events and only two days after fertilization, the precursors for most major organs are already present (Kimmel et al., 1995) (*Fig. 14*). Within 4 or 5 days the fish present a fully functional nervous system as well as complex behaviours. Moreover, since the zebrafish brain also undergoes continuous neurogenesis during adulthood, it is a good system to study adult neurogenesis and regeneration (Kizil et al., 2012).

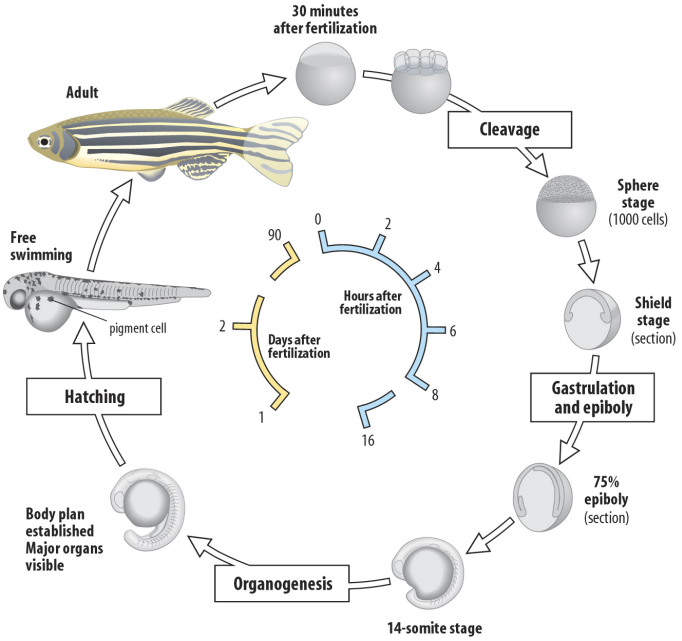

Figure 14: An overview of the different stages of zebrafish development (Wolpert and Tickle, 2010).

2 Hypothesis & Aims

I hypothesized that dopamine (DA) is a signal that modulate neural stem cell and precursor cell proliferation and differentiation both in the developing brain and in the adult. Additionally, a pivotal role of DA in prepulse inhibition was assumed. Both the loss of *th* and the loss of *dat* should therefore affect prepulse inhibition responses. Furthermore, I presumed that the knock-in of transgenes at the *th* or *dat* loci should give expression of the transgene in all cells that normally express *th* or *dat*.

2.1 Dopamine and Neurogenesis

The role of DA in neurogenesis has until now not been fully elucidated. *Th* knockout mice models gave indications that the disruption of the *Th* gene does not have any severe effect neither on the development of catecholaminergic neurons nor on dopaminergic pathways (Kobayashi et al., 1995; Zhou and Palmiter, 1995; Zhou et al., 1994). On the other hand, the finding of relatively high DA levels in $Th^{-/-}$ mice suggests that the dopaminergic neurons in effect did not develop in complete absence of DA. Due to the high DA levels found in $Th^{-/-}$ mice a compensatory mechanism for the production of catecholamines, and especially DA, independent of *Th* it was postulated (Kobayashi et al., 1995). Furthermore, it has also been shown that cells producing melanin do this by converting tyrosine to L-DOPA using the enzyme tyrosinase (Kumar et al., 2011). This means that L-DOPA is available for theoretical further conversion to DA by AADC, which is present in *Th* knockout mice (Kobayashi et al., 1995). Altogether it is clear that further investigation should be done to clarify the actual role of DA in the development of catecholaminergic neurons and DA containing pathways. It is additionally of interest to use *th* knockouts to investigate the role of DA in the development of other neuronal subtypes, and based on the recent publication by Reimer et al. (2013), especially spinal motor neurons. Reimer et al. (2013) claim that DA from the brain promotes spinal motor neuron generation and regeneration in zebrafish , strongly reinforcing the hypothesis that DA effects neuronal development. In the light of this, it is of even grater interest to re-investigate the role of DA in catecholaminergic neuron development as well.

Further supporting the DA and neurogenesis hypothesis, Reimer et al. (2013) demonstrated that the same mechanisms for the promotion of spinal motor neuron generation where activation of the D4a receptor elicits downstream targets interacting with the hedgehog pathway, seem to be reactivated in the generation of motor neurons in the lesioned spinal cord of adult zebrafish. This concurs with other reports that DA plays an important role in adult neurogenesis. Already 10 years ago, Hoglinger et al. (2014) provided clear evidence that proliferative precursors in the adult mammalian subependymal zone express DA receptors and receive dopaminergic signals. Moreover, they found that experimental depletion of DA decreases precursor cell proliferation in the subependymal zone and also in the SGZ (Hoglinger et al., 2004). The same year, it was shown that the destruction of dopaminergic neurons in the substantia nigra and VTA significantly reduce the amount of proliferating neural precursors in the SVZ (Baker et al., 2004). Thus, DA has been identified as an endogenous regulator of adult neurogenesis. A more thorough understanding of this regulation is pivotal towards endogenously produced neuroprecursors being used for cell replacement strategies in neurodegenerative disorders.

In zebrafish, striatal dopaminergic neurons are continuously added from stem cells, and thus these neurons are a good candidate system to determine whether loss of dopaminergic input may affect dopaminergic neurogenesis in the brain (Mahler et al., 2010). In this context, it would be of interest to disclose if DA could be involved in projections and synapse formation of dopaminergic neurons from dopaminergic groups 2, 4 and 6 to the subpallium, considering that a number of experiments suggest DA signalling playing a role in the establishment of synaptic connections (Todd, 1992). These projections in zebrafish are homologue to the projections from the A11 group to the striatum in mammals, a pathway of particular significance due to its involvement in Parkinson's disease.

2.1.1 Targeted Mutations

As a first step towards confirming my hypothesis stating that DA is a signal modulating neural stem cell and precursor cell proliferation and differentiation, I aim to genetically confirm the results from the publication by Reimer et al. (2013) using homozygous *th* knockout zebrafish. Additionally, the knockout will be used to re-evaluate the effect of DA on neurogenesis of catecholaminergic neurons in the brain. Knocking down genes with morpholinos or pharmacological substances as done by Reimer et al. (2013) rarely result in a complete gene knockout, meaning that some DA might still be produced by Th. Pharmacological and morpholino interference with DA signalling may also have elicited off-target effects in zebrafish, as drugs have not yet been thoroughly characterized in fish and morpholinos can act outside of their targets (Bedell et al., 2011). Moreover, the *orthopedia a (otpa)* mutant used by Reimer et al. (2013) does not affect all DA neurons but instead affects neurons in the hindbrain and in the spine. Thus, these models are far from optimal for the study of DA's effect on neurogenesis.

Furthermore, research shows that the DA knockout mouse model *Dbh:Th, $Th^{-/-}$*, where the ability to produce noradrenaline is restored in noradrenergic cells, also was far from ideal seeing as evidence was provided showing that DA is co-released together with noradrenaline from noradrenergic neurons in the cerebral cortex (Devoto et al., 2001; Devoto and Flore, 2006). If so, DA function in the *Dbh:Th, $Th^{-/-}$* strain may be partially rescued in dopaminergic function because the noradrenergic neurons will be able to release DA. All in all, this implicates that to attain a good DA knockout model, *th* should be expressed neither in dopaminergic neurons nor in noradrenergic neurons which was the case for the $Th^{-/-}$ mouse model. However, as mentioned earlier, this $Th^{-/-}$ phenotype led to perinatal lethality due to cardiovascular malfunction, restricting the research on these mice. Zebrafish on the other hand, should be a better model for a *th* knockout as it has been shown that zebrafish larvae can develop largely normal without cardiovascular function until at least 5 or 6 days post fertilization. As zebrafish embryos develop rapidly, this should suffice in time to thoroughly study the embryonic development of dopaminergic neurons in the brain and motor neurons in the spine as well as behavioural patterns. Additionally, seeking to eliminate possible presence of DA as observed in the $Th^{-/-}$ mice, other zebrafish mutant lines, for instance melanin deficient lines, are available for crossing making it possible to create *th* mutants where no tyrosine is converted to L-DOPA neither in *th* cells nor in cells normally producing melanin.

Nonetheless, when making a Th deficient zebrafish, this implies that the animals will not only be deficient of DA, but of all catecholamines and one must keep in mind that

any potential phenotype of these mutants is not necessarily only due to the lack of DA alone. Hence, to corroborate and support the statement that DA affects neurogenesis, the disclosure of the effect of an increased level of DA is essential, something which can be made possible with the creation of a *dat* knockout zebrafish line.

Consequently, *th* and *dat* knockout zebrafish lines will be created and used to investigate motor neuron development in the spine where differences in cell number will be compared between mutants and wildtypes. The same mutants will furthermore be used to study the development of catecholaminergic neurons in the brain. For the $th^{-/-}$ zebrafish, a significant decrease in neuron number in the brain and ventral spinal cord as compared to wildtype is expected given the loss of DA signalling which was suggested to be necessary for proper motor neuron development by Reimer et al. (2013). Furthermore, if DA signalling does in fact affect neuron development, it can be hypothesized that an augmentation of DA should inversely affect neurogenesis when compared to loss of DA, and could in fact increase the number of catecholaminergic neurons in the brain and motor neurons in the spinal cord.

2.2 Dopamine and Prepulse Inhibition

Prepulse inhibition (PPI) of the startle response has been successfully applied in zebrafish as a measure of sensorimotor gating (Burgess and Granato, 2007). Burgess and Granato show that the acoustic startle response in zebrafish larvae is modulated by weak prepulses in a way that can be compared to PPI in mammals. Zebrafish show a startle response by bending their body in a C-shape, followed by a smaller counter bend and swimming. The startle responses occur in two waves termed short-latency C-start (SLC) responses and long-latency C-start (LLC) responses, however only SLC is modulated by prepulses.

Burgess and Granato claim, in agreement with several other studies, that DA is important for the modulation of PPI. An increased amount of DA is found in the nucelus accumbens of animals with sensory gating failure, and it has further been shown that drugs such as the DA agonist amphetamine can significantly decrease PPI. To further study the role of DA in the modulation of PPI, I aim to use *th* and *dat* knockout zebrafish in PPI experiments. Since *dat* deficient animals will have an accumulation of DA in the synaptic cleft, it is reasonable to believe that a similar decrease in PPI will be observed in these fish as observed in animals with high DA levels in the nucleus accumbens. A complete absence of DA, such as in the *th* knockout, should also cause disturbances in the startle responses in the PPI test if DA in fact plays a central modulating role here.

2.3 The Creation of Knock-in Lines

The establishment of a targeted knock-in system and the creation of knock-in lines will be of great use for further studies of DA and the development of the dopaminergic systems. Applying different transgenic technologies, a large number of zebrafish lines expressing certain transgenes under specific promoters have already been generated. These transgenic technologies however, are often not specifically targeted, meaning that the expression of the transgene is not driven by the endogenous promoter but rather cloned sequences of what is believed to be the promoter sequence placed randomly into the genome. Using regions upstream of the *th* gene of as much as 154 kbp in addition to

large parts of the *th* zebrafish gene itself, a stable expression of eGFP was achieved in zebrafish catecholaminergic neurons (Tay et al., 2011). Unfortunately, stable transgenes showed insertion site position effects eliminating expression in some of the dopaminergic groups (Fernandes et al., 2012). A targeted knock-in, using TALENs or CRISPRs, will give the possibility to drive the expression of transgenes endogenously, using the full enhancer regions of the gene in question and doing this at the *th* or *dat* loci is expected to give rise to transgenic lines expressing the desired gene in all catecholaminergic neurons.

TALENs have already been reported to successfully mediate the knock-in of transgenes into the zebrafish genome (Zu et al., 2013). The authors' knock-in strategy was based on homologous recombination, using a transgene flanked by large homology arms. With this approach an eGFP DNA sequence appeared to be knocked into the *th* locus as revealed by PCR. However, no stable eGFP expression was achieved. Recent research shows that also the CRISPR/Cas9 system can be efficiently used for targeted knock-in of genes (Auer et al., 2013). (Auer et al., 2013). based their knock-in strategy on non homologous end joining (NHEJ), an event which in the cell is both more frequent and more efficient than homologous recombination, and achieved stable expression of the transgene. Moreover, it was shown that knock-in efficiencies were highest when the transgene was presented to the cell as part of a circular donor plasmid containing the CRISPR target sequence which then was linearized by the CRISPR system inside the cell. It has also been reported that CRISPRs can be used to mediate the knock-in of short single stranded oligonucleotides quite efficiently (Gagnon et al., 2014).

In theory, it should not make any difference if the creation of the double strand break (DSB) in the genome and the linearization of the donor plasmid is mediated by CRISPRs or TALENs. I aim to take advantage of either the TALEN or CRISPR/Cas9 system to establish a method for the targeted knock-in of transgenes, aspiring to knock in the ORFs of eGFP, Gal4 or an attP landing site for the use with the PhiC31 integrase at the *dat* and *th* loci in zebrafish. Upon successful integration, six new zebrafish lines will be created where each of the transgenes are expressed in either the *th* or *dat* expression pattern.

3 Materials & Methods

3.1 Materials

3.1.1 Chemicals, Reagents and Enzymes

Name	Formula	Provider	Number
2x YT broth		MP Biomedicals	3012-031
5x Transcription buffer		Fermentas	00031807
100 bp, 1 kb DNA ladder		Fermentas	00129569 (1 kb) 00137239 (100 bp)
Agar Bacteriology grade BC		AppliChem	A0949
Agarose	[C12H18O9]n	Bioron	604005
Ampicillin	$C_{16}H_{18}N_3NaO_4S$	AppliChem	A0839
Anti-digoxigenin-alkaline phosphatase conjugated goat anitbody		Roche Applied Science	11093274910
Anti-rabbit Alexa 488 goat		Invitrogen	A-11034
ATP	$C_{10}H_{14}N_5Na_2O_{13}P_3$	AppliChem	A1348
BCIP (5-Bromo-4-chloro-3-indolyl phosphate)	$C_8H_8BrClNO_4$	AppliChem	A1117
BSA (Bovine Serum Albumin), proteinase-free		Sigma-Aldrich	A3059
Blocking reagent		Roche Applied Science	11096176001
DIG RNA labeling mix		Roche Applied Science	11277073910
DMSO (dimethylsulfoxide)	C_2H_6SO	Carl Roth GmbH	A994.1
DNAseI, RNAse free		Roche Applied Science	04716728001
dNTPs		Fermentas	dGTP:R0161 dTTP:R0171 dATP:R0141 dCTP:R0151
EDTA (Ethylendiaminetetraacetic acid)	$C_{10}H_{16}N_2O_8$	Sigma-Aldrich	60-00-4
Esp3I		Fermentas	ER0452
Ethanol	CH_3CH_2OH	AppliChem	A1612
Ethidiumbromide (1%)	$C_{21}H_{20}BrN_3$	AppliChem	A1152
Formamide	CH_3NO	AppliChem	A2156
Glycerol	$C_3H_8O_3$	AppliChem	A1123
Goat serum		Sigma-Aldrich	G9023
Heparin		Sigma-Aldrich	H9399
Hydrogen chloride	HCl	Carl Roth GmbH	7647-01-0
IPTG	$C_9H_{18}O_5S$	AppliChem	A1008
Kanamycin	$C_{18}H_{36}N_4O_{11}SO_4$	Carl Roth GmbH	T832.1
Lithium chloride	LiCl	Carl Roth GmbH	3739.1
Loading dye(6x)		Fermentas	R0611
Magnesium chloride	$MgCl_2$	AppliChem	A1036
Maleic acid	$C_4H_4O_4$	Sigma-Aldrich	M-0375
Methy Cellulose		Sigma-Aldrich	M-0387

Methanol	CH_3OH	AppliChem	A3493
Methylene Blue	$C_{16}H_{18}N_3SC$	Sigma-Aldrich	MB1
NBT (Nitro tetrazolium blue chloride)	$C_{40}H_{30}Cl_2N_{10}O_6$	AppliChem	A1243
Paraformaldehyde	$[CH_2O]_8 - 1_{00}$	Sigma-Aldrich	76240
PfuUltra II Fusion HS DNA Ploymerase		Agilent Technologies	600670
Phenol red	$C_{19}H_{14}O_5S$	Sigma-Aldrich	143-74-8
Phenylthiourea	$C_6H_5NHCSNH_2$	Sigma-Aldrich	P-7629
Plasmid Safe DNase		epicentre	E3101K
Potassium Chloride	KCl	Carl Roth GmbH	6781
Proteinase K		AppliChem	A3830
Quick Ligase		New England Biolabs GmbH	M2200L
Rabbit anti TH (zebrafish)		Ryu et al. (2007)	
Red sea salt		Red Sea Deutschland	05021447
Restriction enzymes and buffers		New England Biolabs GmbH	ApaI:R0114S BamHI:R0136 BamHI-HF:R3136S BbsI:R0539S BsaI:R05355 BssHII:R0199S ClaI:R0197S EcoRI:R0101S HindIII:R0104S KpnI:R0142S NotI:R0189S PmeI:R0560S PvuI:R0150S SacI:R0156S SpeI:R0133S SphI:R0182S XbaI:R0145S XhoI:R0146S NEBuffer1.1:B7201S NEBuffer2.1:B7202S NEBuffer3.1:B7203 CutSmart:B7204
RiboLock RNAse Inhibitor		Fermentas	00115303
Sheep serum		Sigma-Aldrich	52263
S.O.C. medium		Invitrogen	15544-034
Sodium chloride	NaCl	AppliChem	A3597
Spectinomycin dihydrochloride pentahydrate	$C_{14}H_{24}N_2O_7$x 2 HCl x5 H_2O	Sigma-Aldrich	S4014-5G
T3, T7 Polymerases		Fermentas	00084729(T3) 00086741(T7)
T4 DNA ligase		Thermo Scientific	EL0014
T4 DNA ligase buffer		Thermo Scientific	B69
T4 DNA polymerase		Thermo Scientific	EP0061
T4 Polynucleotide Kinase		New England Biolabs GmbH	M0201S
T4 Polynucleotide Kinase buffer		New England Biolabs GmbH	B0201S
T7 endonuclease I		New England Biolabs GmbH	M0302S
Taq polymerase		Bioline GmbH	BIO-21107

Tango buffer		Fermentas	BY5
Torula RNA		Sigma-Aldrich	R-6625
Tricain (Ethyl3-aminobenzoate methansulfonate)	$C_{10}H_{15}NO_5S$	Sigma-Aldrich	A5040
Tris	$C_4H_{11}NO_3$	AppliChem	A2264
Tween-20 (Polyoxyethylen-20-sorbitan-monolaurate)		AppliChem	A1389
X-gal	$C_{14}H_{15}BrClNO_6$	AppliChem	A1007

Table 1: List of chemicals and enzymes

3.1.2 Solutions

Solution	Content
Ampicillin working solution	100mg/mL
Annealing buffer	10mM Tris, 1mM EDTA, 50 mM NaCl
BCIP	50mg/mL BCIP in 100% DMF
Blocking solution for WISH	2% BSA, 5% sheep serum in PBST
Blocking reagent for IHC	10% blocking reagent, 100mM Maleic acid, 150mM NaCl
Blocking solution for IHC	5% goat serum, 1% blocking reagent, 1% BSA in PBTD
BSA 10x solution	0.1g/mL in mpH_2O
E3 Embryo medium	5mM NaCl, 0.17mM KCl, 0.33mM $CaCl_2$, 0.33mM $MgSO_4$, 0.01% methylene blue in mpH_2O
Egg water	0.3g/L see salt in mpH_2O
50% Formamide	50% Formamide, 2x SSC, 0.1% Tween20 in mpH_2O
25% Formamide	25 Formamide, 2x SSC, 0.1% Tween20 in mpH_2O
80 % Glycerol	80 % Glycerol in PBST, 1mM EDTA
Hybridisation mix (Hyb)	50% Formamid, 5x SSC, 5 mg/mL Torula-RNA, 50 µg/mL Heparin, 0.1% Tween20, in mpH_2O
Kanamycin working solution	50mg/mL
Lysis buffer	10mM Tris pH8, 1mM EDTA pH8, 30mM KCl, 30mM NaCl, 30mM $MgCl_2$, 0.5 Tween20, 0.5% NP40
Lysis buffer 2	1mM TrisHCl pH 8.0, 50mM KCl, 0.3% Tween20, 0.3% NP40, 1mM EDTA in mpH_2O
Mounting solution	7.5g methyl cellulose, 12mL tricain, 288mL egg water
Mounting medium for IHC	80% glycerol, 1% agarose in PBST
NBT	100mg/mL NBT in 70% DMF in mpH20
NTMT buffer	0.10M TrisHCl (pH 9.5), 0.10M NaCl, 0.05M $MgCl_2$, 0.1% Tween-20, 4.5mg NBT, 1.75mg BCIP in 25mL mpH_2O
4% PFA	20g PFA, 50 mL 1x PBS
20x PBS stock solution	35mM KH_2PO_4, 208mM NaH_2PO_4, 54mM KCl, 2.74M NaCl diluted in 1L mpH_2O, pH 7.4
1x PBST	50mL 20x PBS stock solution, 0.1% Tween-20 up to 1L with mpH_2O, pH adjusted with NaOH 7.5
PBTD	1% DSMO in PBST
Phenol red working solution	0.5% Phenol red in mpH_2O

Phosphate buffer	800mL 0.5M Na_2HPO_4, 200mL 0.5M NaH_2PO_4
10x PTU stock solution	2mM PTU in mpH_2O
1x PTU working solution	100 mL 10x PTU, 900 ml methylene blue egg water
Proteinase K stock solution	20mg/mL in mpH_2O
Proteinase K working solution (WISH)	10 µg/mL Proteinase K in PBST
Spectinomycin stock solution	50mg/mL
20x SSC stock solution	3M NaCl, 300mM NatriumCitrat, up to 1L with mpH_2O, pH adjusted with HCl to 7
2x SSCT working solution	100mL 20x SSC stock solution, 0.1% Tween-20 in 1L mpH_2O
0.2x SSCT	10mL 20x SSC stock solution, 0.1% Tween20 in 1L mpH_2O
Stop Solution WISH	400mL PO_4-buffer (0.5M, pH 5.5), 80g NaCl, 2g KCl, 20 mL EDTA (0.05M), 10 mL Tween-20, up to 1L mpH_2O
1x TAE (Tris-acetate-EDTA)	40mM Tris, 1mM EDTA, 20mM acetic acid in mpH_2O
25x Tricane stock solution	4g/L Tricane powder in mpH_2O, 1M TrisHCl (pH9)
TE buffer	10mM Tris, 1mM EDTA in mpH_2O, pH 8.0
1x Tricaine working solution	4.2mL Tricaine stock solution in 100mL egg water
YT broth Amp/Kan/Spec	16 capsules 2x YT broth, 500 µL Amp/Kan/Spec stock solution, up to 500mL mpH_2O, autoclaved before addition of antibiotics
YT Agar broth Amp/Kan/Spec	16 capsules 2x YT broth, 6.5g Agar, 500 µL Amp/Kan/Spec stock solution, up to 500mL mpH_2O, autoclaved before addition of antibiotics.

Table 2: List of solutions

3.1.3 Equipment

Product	**Manufacturer**
Capillaries, TW100F-4	World Precision Instruments
Centrifuge 5417 R	Eppendorf
Centrifuge (large), Heraeus Megafuge 40R	Thermo Scientific
Cover glasses 18x18mm	Menzel Gläser
DAQ Timer, PCI-6221	Natioanl Instruments
Electrophoresis Power Supply (Power Pac 300)	BioRad Laboratories
Electrophoresis Chambers, Owl Easycast	Thermo Scientific
E.Z.N.A Cycle Pure Kit, D6492-00	OMEGA Bio-Tek
E.Z.N.A DNA Plasmid Mini Kit I, D6942-02	OMEGA Bio-Tek
E.Z.N.A DNA Plasmid Midi Kit, D6904-03	OMEGA Bio-Tek
Freezer -20°C, Premium	Libherr
Freezer -80°C, MDF-U6086S	Sanyo
Gel documentation system (Vilber), In_nity-3000	Vilber Lourmant
High-Speed Camera, FASTCAN-1024PCI	Photron Europe LTD
Incubator fish, Heraeus	Kendro Laboratory Products
Incubator 37°C, 55°C	Memmert GmbH
Leica MZ12 Stereomicroscope	Leica Mikrosysteme Vertrieb GmbH

Leica MZ16F fluorescent microscope	Leica Mikrosysteme Vertrieb GmbH
Leica MZ16FA fluorescent microscope	Leica Mikrosysteme Vertrieb GmbH
Leica DFC350Fx Camera	Leica Mikrosysteme Vertrieb GmbH
mMESSAGE mMACHINE T3 Transcription Kit, AM1348	Ambion
MEGAscript T7 transcription Kit, AM1333	Ambion
MEGAscript SP6 Transcription Kit, AM1330	Ambion
mMessage mMachine T7 ULTRA Transcription Kit, AM1345	Ambion
Microneedle puller, P-30	Sutter Instruments
Mini Shaker, 4810	Brüel and Kjær
NanoDrop Spectrophotometer, ND-1000	Preqlab Biotechnologie GmbH
NucleoSpin Gel and PCR Clean-up, 740609	Macherey-Nagel
PCR Machine, PTC-100 Programmable Thermal Controller	MJ Research Inc.
Pico Pneumatic Pump, PV280	World Precision Instruments
Plasmid Midi Kit	Quiagen
Refrigerator 4°C	Libherr
RNeasy Mini Kit	Quiagen
Test-tube-rotator, 34528	Snijders
Thermomixer, compact	eppendorf
Vortex, Vortex Genie 2	Scientific industries inc.
Water Bath	Memmert GmbH
Zeiss Axioskop 2	Carl Zeiss MicroImaging GmbH
Zeiss LSM510	Carl Zeiss MicroImaging GmbH

Table 3: List of equipment

3.1.4 Plasmids

Name	**Obtained From**	**Published**
TALEN plasmids	Addgene (Kit No. 1000000024)	(Cermak et al., 2011)
pME-MCS	Chien Lab	(Kwan et al., 2007)
pME-EGFP-CAAX	Chien Lab	(Kwan et al., 2007)
pME-Gal4VP16	Chien Lab	(Kwan et al., 2007)
pKhsp82MOS (attP landing site)	Konrad Baslers Lab, Zurich	(Bischof et al., 2007)
pBs zTH (Th probe)	Jochen Holzschuh	(Holzschuh et al., 2001)
pBs-KS+isl-1 zf (Isl1 probe)	Okamoto Lab	(Inoue et al., 1994)
pBs-DAT (DAT probe)	Jochen Holzschuh	(Holzschuh et al., 2001)
HB9 probe	Dirk Meyer Lab, Innsbruck	
Zebrafish-gRNA-0007 (th)	Addgene (42247)	(Hwang et al., 2013)
eGFP gRNA	Filippo del Bene Lab, Paris	(Auer et al., 2013)
pT7EGFPgRNA	Addgene (46760)	(Jao et al., 2013)
pCS2-nCas9n	Addgene (47929)	(Jao et al., 2013)
Cas9 Paris	Filippo del Bene Lab, Paris	(Auer et al., 2013)
MLM3613	Addgene (42251)	(Hwang et al., 2013)

JDS246	Addgene (43861)	Keith Joung
eGFPbaitE2A-KalTA4 (eGFP-Gal4 donor plasmid)	Filippo del Bene Lab, Paris	(Auer et al., 2013)

Table 4: List of plasmids

3.2 Fish Care

All fish care and experiments were conducted according to German animal protection laws. Up until 5 days of age, embryos were incubated at 28.5 °C in egg water mixed with 1/3 methylene blue. Embryos utilized for behavioural analysis were raised in E3 embryo medium in an incubator with a normal night day light cycle. To suppress pigmentation of embryos that were to be used for fluorescent screening and in situ hybridizations, PTU (Phenylthiourea) was added to the water. Embryos to be raised to adulthood were transferred to the main fish facility at day 5 after fertilization.

To genotype single fish, fin-clipping was performed on juvenile or adult fish. Fish were then first anesthetized by being put into water containing tricain. Once the fish no longer moved they were put on a piece of parafilm paper and a small corner of the back fin was cut off with a scalpel. The fin-clipped fish were kept in separate tanks containing fish water and methylene blue until the genotype was clear, and they were either euthanizated according to law or put in new fish tanks in the fish facility.

An overview over fish stocks used in this work is listed in table 5.

Genotype	**Stock Number**
Wildtype AB/TL	14354, 14790, 14658, 14439, 15606, 15150, 15040, 14440, 15775, 14042
Tg(otpb:gfp)	19760
th F0 TAL inj	15941
dat F0 TAL inj	15972, 15962, 15972
th F1 outcross	16179, 16180, 16181, 16182
dat F1 outcross	16220, 16221, 16283, 16387, 16388, 16389
th m1403 F1 het	16181
th m1404 F1 het	16180, 16181
th m1405 F1 het	16181
dat F1 het	16221
Tg(UAS:RFP, cry:eGFP)	Received from Thomas Auer
Tg(neuroD:eGFP)	Received from Thomas Auer

Table 5: Fish stocks

3.3 Common Laboratory Procedures

3.3.1 Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) is a method used to amplify specific sequences of DNA (Bartlett and Stirling, 2003). Double stranded DNA is first denatured at temperatures around 93-98 °C, giving access to the base sequence otherwise hidden within the double helix. Short oligonucleotides (15-25 nucleotides) called primers are designed to bind to each end of the targeted DNA sequence (the amplicon). The temperature at which the primers bind to the template DNA depends on the sequence and the melting temperature (Tm) of the primers in question, but it usually ranges between 50-65°C. Starting at the 3' end of the bound primers, a polymerase copies the amplicon using the denatured DNA

as a template. The polymerases used for PCR normally work at 72°C. By repeating the whole process 20-35 times, thousands to millions of copies of the particular DNA sequence can be obtained.

A PCR reaction must contain DNA template, forward and reverse primers, reaction buffer, nucleotides (often added to buffer) and a polymerase. In this work *PfuUltra* II Fusion HS DNA Ploymerase (Stratagene) was used for the production of sequence accurate PCR products, while the MyTaq^{TM} DNA polymerase (Bioline) was used for all other purposes. The PCR reactions were made following the protocol of the polymerase manufacturer and the quick changes of temperatures needed for a PCR reaction was obtained with a thermocycler. An example of a PCR program used in this work is shown below.

1.	98°C	3 min
2.	98°C	20 sec
3.	2-5°C below the primer Tm	20 sec
4.	72°C	20 sec
5.	Repeat steps 2-4 35 times	
6.	72°C	3 min

Primers for PCRs were designed using either the CLC Sequence Viewer software (CLC bio) or manually using the ApE plasmid editor (by M. Wayne Davis). Primers were analyzed using the online OligoAnalyzer 3.1 from Integrated DNA Technologies (`https://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/`) before they were ordered desalted from Sigma-Aldrich. Primers were diluted in mpH_2O to a concentration of 100 µM. Primer mixes were made mixing forward and reverse primer with a final concentration of 20 µM each.

3.3.2 Agarose Gel Electrophoresis

Agarose gel electrophoresis is a method used to separate DNA or RNA molecules based on their size and electrical charge (Sambrook and Russel, 1968). The negatively charged nucleic acid molecules move at different speeds through the gel when an electric field is applied. Gel electrophoresis can be used to separate DNA or RNA fragments ranging from 50 base pairs to several kilobases. For this work, agarose gels were made mixing agarose with TAE buffer, the amounts of which depend on the percentage and the size of the gel to be made. 0.7% agarose gels are good for separating DNA of 800-12000 bp, while 2% gels separate smaller DNA pieces (50-2000) better. The DNA or RNA bands were made visible under UV-light by the addition of ethidium bromide (0.02% final concentration) to the agarose gel.

3.3.3 Transformation

The protocol was adjusted from Addgene protocols, Bacterial Transformation (`http://www.addgene.org/plasmid-protocols/bacterial-transformation/`). Chemically potent bacterial cells (Top10 or DH5α) to be transformed were removed from the -80°C freezer and let to thaw on ice. 1-5 µL of the plasmid to be transformed was then added to the bacteria. After resting on ice for 20 minutes, the bacteria were put in a water bath of 42°C for 40 seconds. The bacteria were then put back on ice, and 100 µL S.O.C medium

was added before incubation on a shaking heating block (37°C, 300 rpm). After 45-60 minutes, the bacteria were plated on agar plates and left to incubate at 37°C over night.

3.3.4 T7 Endonuclease Assay

The T7 endonuclease assay can be used to detect mutations induced either by TALENs or CRISPRs and it is based on the fact that the T7 endonuclease recognizes and cuts non-perfectly matched DNA (`https://www.neb.com/products/M0302-T7-Endonuclease-I`). genomic DNA (gDNA) from CRISPR or TALEN injections is denatured and re-annealed with the aim of combining mutated and non-mutated strands of DNA to create mismatched sequences which would be recognized by the endonuclease. This means that only in samples containing DNA with a mutation, the nuclease will create a cut, and two bands will be visible on an agarose gel in contrast to the one band which would be expected if no mutation was present and all DNA is perfectly matched. The protocol used in this work was based on the protocol published by Keith Joung (Reyon et al., 2012).

3.3.5 Sequencing

All sequencing was performed by GATC Biotech. Samples for sequencing were cleaned with the E.Z.N.A Cycle Pure Kit (OMEGA) before delivery.

3.4 TALENs

3.4.1 TALEN Design

The TALENs for both the *th* and *dat* loci were designed considering they were going to be used to create both knockout and knock-in mutants. For the knockout mutants, it was essential that the in/del mutation resulting from the TALEN cut would lie at the beginning of the open reading frame of the gene in question. This way, one did not run the risk of having a functional protein, just missing the last part of its amino acid sequence. For the knock-ins, it was found preferable to use the start codon of the inserted transgene, and try to eliminate the ATG of the original gene. Because of this, the TALENs were designed to cut close to the ATG of each loci, hoping that it would be deleted after the cut. Additionally, for screening of the TALEN cutting efficiency, it was found convenient to design the TALENs to cut at a restriction enzyme recognition site, making the screen possible by a restriction digest.

th and *dat* sequence information was extracted from Ensembl Genome Browser (`http://www.ensembl.org/index.html`) (*th* and *dat*) and NCBI (`http://www.ncbi.nlm.nih.gov/pubmed`) (*dat*), and TALEN target sites were found using online programs. Effector Nucleotide Targeter 2.0 (`https://tale-nt.cac.cornell.edu/`) (Doyle et al., 2012) was used for pairs 1 and 2 and Mojo Hand (`http://www.talendesign.org/`) (Bedell et al., 2012) was used for pairs number 3 (*Table 6*). The targeted restriction enzyme recognition sites were EcoRI and SphI for the Th and DAT TALENs respectively.

Before starting the assembly of the TALENs, wildtype ABTL fish from the labs fish facility were sequenced, making sure that no single nucleotide polymorphisms were present which could interfere with the binding of the TALENs. Wildtype ABTL embryos were

Th & DAT TALEN Target Sites

Pair	TALEN Target Site		
	TAL 1	Spacer	TAL 2 / *Reverse Compliment*
Th1	CCATCTGTTCCGCGCGCGCG	ccatctgaac*atg*cc*gaattc*a	AGCAGCTCCACATCTTCCAC *GTGGAAGATGTGGAGCTGCT*
Th2	CCGCGCGCGCGCCATCTGAA	c*atg*cc*gaattc*aagca	GCTCCACATCTTCCACAAAG *CTTTGTGGAAGATGTGGAGC*
Th3	TCCGCGCGCGCGCCATCTGA	ac*atg*cc*gaattc*a	GCAGCTCCACATCTTCCACA *TGTGGAAGATGTGGAGCTGC*
DAT1	GTCCATCAGAGCATCTCCTC	ctcagtctcctgt*gcatgc*cc*atg*c	TGAGAGGCAGACCGGCGGTC *GACCGCCGGTCTGCCTCTCA*
DAT2	CAGAGCATCTCCTCCTCAGT	ctcctgt*gcatgc*cc*atg*ctg	AGAGGCAGACCGGCGGTCAC *GTGACCGCCGGTCTGCCTCT*
DAT3	CCTCCTCAGTCTCCTGT	*gcatgc*cc*atg*ctga	GAGGCAGACCGGCGGTC *GACCGCCGGTCTGCCTC*

1. In the spacer, the native atg is marked in red, while the restriction site EcoRI is marked in blue and SphI in green.

Table 6: Th & DAT TALEN Target Sites.

obtained by crossing of adult fish, and gDNA was then generated from a pool of 10-15 embryos. The embryos were lysed by adding 90 µL lysis buffer and heating at 95°C for 15 minutes before adding 10 µL proteinase K (10mg/mL). After incubation at 55°C for at least 2 hours (could go over night), the proteinase was heat inactivated (15 minutes at 75°C). Finished gDNA was stored at -20°C. A PCR on the gDNA was performed on the wildtype gDNA using the primer mixes Th_TAL and DAT_TAL with annealing temperatures of 50 and 56°C respectively. The cleaned PCR products were then sent for sequencing, using both forward and reverse primer of the sets Th_TAL and DAT_TAL.

3.4.2 Golden Gate TALEN assembly

The Golden Gate TALEN assembly was done following Dan Voytas protocol (Cermak et al., 2011).

Golden Gate 1 (GG1) Reaction Each TALEN of a pair was assembled in two cloning steps. In the GG1 reaction, all RVDs of a TALEN were assembled in the right order into two separate intermediate vector backbones, pFUS_A and pFUS_B. 24 GG1 reactions were mixed as listed below, with RVDs 1-10 together with pFUS_A and RVDs 11-(N-1) together with pFUS_B. A table of the 24 reactions and the respective RVD plasmids can be seen in table 7.

150ng	of each RVD vector
150ng	of appropriate pFUS vector
1 µL	BsaI restriction enzyme
1 µL	Quick Ligase (NEB)
2 µL	10X T4 DNA ligase buffer
up to 20 µL	H_2O

Pair	TAL	Rxn	RVD Plasmids	Last RVD	Target Vector
Th1	1	1a	pHD1 pHD2 pNI3 pNG4 pHD5 pNG6 pNH7 pNG8 pNG9 pHD10	NH	pFUS_A
		1b	pHD1 pNH2 pHD3 pNH4 pHD5 pNH6 pHD7 pNH8 pHD9		pFUS_B
	2	2a	pNH1 pNG2 pNH3 pNH4 pNI5 pNI6 pNH7 pNI8 pNG9 pNH10	NG	pFUS_A
		2b	pNG1 pNH2 pNH3 pNI4 pNH5 pHD6 pNG7 pNH8 pHD9		pFUS_B
Th2	1	3a	pHD1 pHD2 pNH3 pHD4 pNH5 pHD6 pNH7 pHD8 pNH9 pHD10	NI	pFUS_A
		3b	pNH1 pHD2 pHD3 pNI4 pNG5 pHD6 pNG7 pNH8 pNI9		pFUS_B
	2	4a	pHD1 pNG2 pNG3 pNG4 pNH5 pNG6 pNH7 pNH8 pNI9 pNI10	HD	pFUS_A
		4b	pNH1 pNI2 pNG3 pNH4 pNG5 pNH6 pNH7 pNI8 pNH9		pFUS_B
Th3	1	5a	pNG1 pHD2 pHD3 pNN4 pHD5 pNN6 pHD7 pNN8 pHD9 pNN10	NI	pFUS_A
		5b	pHD1 pNN2 pHD3 pHD4 pNI5 pNG6 pHD7 pNG8 pNN9		pFUS_B
	2	6a	pNG1 pNN2 pNG3 pNN4 pNN5 pNI6 pNI7 pNN8 pNI9 pNG10	HD	pFUS_A
		6b	pNN1 pNG2 pNN3 pNN4 pNI5 pNN6 pHD7 pNG8 pNN9		pFUS_B
DAT1	1	7a	pNH1 pNG2 pHD3 pHD4 pNI5 pNG6 pHD7 pNI8 pNH9 pNI10	HD	pFUS_A
		7b	pNH1 pHD2 pNI3 pNG4 pHD5 pNG6 pHD7 pHD8 pNG9		pFUS_B
	2	8a	pNH1 pNI2 pHD3 pHD4 pNH5 pHD6 pHD7 pNH8 pNH9 pNG10	NI	pFUS_A
		8b	pHD1 pNG2 pNH3 pHD4 pHD5 pNG6 pHD7 pNG8 pHD9		pFUS_B
DAT2	1	9a	pNI2 pNH3 pNI4 pNH5 pHD6 pNI7 pNG8 pHD9 pNG10	NG	pFUS_A
		9b	pHD2 pNG3 pHD4 pHD5 pNG6 pHD7 pNI8 pNH9		pFUS_B
	2	10a	pNG2 pNH3 pNI4 pHD5 pHD6 pNH7 pHD8 pHD9 pNH10	NG	pFUS_A
		10b	pNH1 pNG2 pHD3 pNG4 pNH5 pHD6 pHD7 pNG8 pHD9		pFUS_B
DAT3	1	11a	pHD1 pHD2 pNG3 pHD4 pHD5 pNG6 pHD7 pNI8 pNN9 pNG10	NG	pFUS_A
		11b	pHD1 pNG2 pHD3 pHD4 pNG5 pNN6		pFUS_B
	2	12a	pNN1 pNI2 pHD3 pHD4 pNN5 pHD6 pHD7 pNN8 pNN9 pNG10	HD	pFUS_A
		12b	pHD1 pNG2 pNN3 pHD4 pHD5 pNG6		pFUS_B

Table 7: Golden Gate 1 Reactions

The reactions were then put in a thermocycler with the following cycle:

1. 37°C 5 min
2. 16°C 10 min
3. Reapeat 1-2 10x
4. 50°C 5 min
5. 80°C 5 min

To destroy all unligated linear dsDNA fragments including incomplete ligation products with lower number of repeats fused, 1 µL of Plasmid-Safe nuclease was added to the reactions together with 1 µL ATP (10mM). Following incubation for 1 hour at 37°C, 5 µL of the Golden Gate reactions were transformed into chemically competent cells (TOP10 or DH5α). The transformed cells were plated on Spec^{50} agar plates together with 40 µL X-gal (20mg/mL) and 40 µL IPTG (0.1M) for blue white selection. After incubation at 37°C over night, three white colonies were picked from each plate for a colony PCR. Primers pCR8_F1 and pCR8_R1 were used with the following program:

1. 95°C 2 min
2. 95°C 30 sec
3. 55°C 30 sec
4. 72°C 1 min 45 sec
5. Reapeat 2-4 34x
6. 72°C 5 min

The PCR products were run on 1% agarose gels together with a 1 kb DNA ladder. Correct clones containing the right plasmids should have a band at about 1.2 kb for vectors with

10 repeats and they should have a smearing above and a laddering effect under the 1.2 kb band. Overnight cultures in 3.5mL $Spec^{50}$ YT broth were started for one correct clone from each GG1 reaction. For plates without white colonies or with empty vectors, the GG1 reactions were repeated. The plasmids were extracted on the next day using DNA Plasmid Mini Kit (OMEGA), and concentrations were measured with NanoDrop.

Golden Gate 2 (GG2) Reaction Using the plasmids from the GG1 reactions, GG2 reactions were mixed accordingly:

150 ng	vector A (GG1 reaction a)
150 ng	vector B (GG1 reaction b)
150 ng	of last RVD vector (pLR)
75 ng	of final destination vector pRCIscript
1 µL	Esp3I restriction enzyme
1 µL	Quick Ligase
2 µL	10X T4 DNA ligase buffer
Up to 20 µL	H_2O

The thermocycler cycle was set as follows:

1. 37°C 5 min
2. 16°C 10 min
3. Reapeat 1-2 10x
4. 37°C 15 min
5. 80°C 5 min

5 µL of the GG2 reaction were transformed into chemically competent cells (TOP10 or 5HDα) and plated on Amp^{100} agar plates. Because of the high efficiency of the GG2 reaction no blue white selection was necessary. After incubation over night at 37°C, a colony PCR was again performed, this time using primers TAL_F1 and TAL_R2 and the following cycle:

1. 95°C 2 min
2. 95°C 30 sec
3. 55°C 30 sec
4. 72°C 3 min
5. Repeat 2-4 34x
6. 72°C 5 min

PCR products were analyzed on a 1% agarose gel where correct clones should have a band and or a smear around 3 kb and the laddering effect below. Correct clones were chosen for overnight culture in 3.5mL Amp^{100} YT broth and the next day the plasmids were isolated using the DNA Plasmid Mini Kit. The clones containing the full length TALENs were sent for sequencing with the primers SeqTALEN 5-1 and TAL_R2 and sequence results were analyzed using the online program TAL Plasmids Sequence Assembly Tool from Georgia Institute of Technology (`http://bit.ly/assembleTALsequences`). Pasting in the TALEN target sequence together with the forward and reverse sequencing results, the program rapidly shows if the cloned TALEN is correct. If this was not the case, the Golden Gate reactions were repeated.

3.4.3 Injection of TALEN mRNA and Screening of Injected Embryos

Between 5 and 10 µg of all TALEN plasmids were linearized with the restriction enzyme PvuI. After digestion at 37°C for two hours, 2 µL of the reactions were run on an 1% agarose gel to check for complete digestion. PvuI cuts the TALEN plasmids at two places, and a complete digest should have a band around 5200 bp and one around 1000 bp, varying slightly according to number of repeats in each TALEN. Digests were cleaned with the Cycle Pure Kit and the concentrations were measured with NanoDrop. mRNAs were transcribed using mMessage mMachine T3 Kit following the manufactures protocol of a capped transcription reaction assembly with a LiCl precipitation. 2 µL of each finished mRNA were run on a 1.5% agarose gel after being denatured at 75°C for 90 seconds. A correct mRNA should have a band between 1000 and 1500 bp. Finally, the mRNA concentrations were measured with NanoDrop and stored at -80°C.

ABTL wildtype zebrafish mating couples were set up with a separator for spawing in the afternoon on the day before injection. On the day of injection an injection mix for each TALEN pair was prepared in mpH_2O with 200 ng/µL of each TALEN mRNA and 2 µL phenol red to control injection volume and position.

Separators were removed from the crossing tanks and 1nL of the injection mix was injected into 1 cell stage embryos using a microinjector. At least 50 embryos were saved as uninjected controls. At 24 hours post fertilization (hpf) 10-15 embryos were pooled both from the control and injected embryos and gDNA was made as described above.

For the analysis of TALEN cutting efficiency, the targeted loci were amplified by PCR using 1 µL of gDNA and the primer mixes Th_TAL and DAT_TAL with annealing temperatures of 50 and 56°C respectively. 15 µL of the PCR products were used for the restriction digests with the enzymes EcoRI (for Th) and SphI (for DAT). Since the TALENs were designed to cut at a restriction site, this site would not be present if a cut and an in/del mutation had occurred there. Digests were ran on a 1.5% agarose gel and the efficiencies of the TALENs were decided by comparing the intensity of the cleaved bands to the intensity of the uncleaved bands in the injected samples. The uninjected controls should show complete cleavage of the PCR product. Embryos of injected clutches showing an in/del mutation were raised to adulthood and were named Generation 0 (G_0).

3.4.4 Creating a Stable Knockout

To identify mosaic G_0 TALEN mutant founders, fish of the G_0 generation were out-crossed to ABTL wildtype fish, and the resulting clutches were screened for the TALEN mutation using the same PCR and restriction enzyme digest as described above for screening of TALEN efficiency. Parents of clutches showing the TALEN mutation were identified as founders and their embryos, which were named filial generation 1 (F_1 generation), were raised to adulthood.

The F_1 generation was genotyped by fin-clipping and sequencing of individual fish. gDNA from the fin-clips was made by adding 50 µL lysis buffer to each fin-clip and heating for 10 minutes at 95°C. 10 µL proteinase K (10mg/mL) was then added and the fins were left to incubate at 55°C over night. PCRs and restriction digests were then performed on the gDNA from the fin-clips as earlier described. The digests were run on an 1% agarose

gel and fish containing a mutation were identified. To amplify only the mutated allele of the identified fish, the PCR was repeated using DNA picked with a pipette from the uncut band on the gel. Cleaned PCR products of the mutated bands were then sent for sequencing. Sequence products were aligned and analyzed using the software Sequencher (Gene Codes Corporation) and the ApE plasmid editor. Fish with frameshift mutations leading to a premature stop codon were separated according to their specific mutations and were named F_1 founders. F_1 founders could be out-crossed to ABTL wildtype fish to create a stable knockout line heterozygous for a specific mutation. For experiments where homozygous mutants were needed, F_1 founders identified with the same mutated allele were in-crossed yielding clutches where 1/4 of the embryos were homozygous for their TALEN mutation.

3.5 Immunohistochemistry

To ensure a complete gene knockout, immunohistochemistry (IHC) was performed using an antibody labelling the targeted gene. 3 days old zebrafish embryos were fixed in p-formaldehyde (4%) for 4 hours at room temperature or over night at 4°C. After washing twice in PBST, the embryos were dehydrated by washing in increasing concentrations of methanol in PBST (1:3, 1:1, 3:1 and 4:0). Embryos were stored at -20°C in 100% methanol.

IHC was carried out as reported by Holzschuh et al. (2003). Th F_2 larvae (11 and 7 bp deletion) were rehydrated with stepwise decreasing concentrations of methanol in PBST (75%, 50%, 25%) and then washed 3 times in PBST before they were incubated in proteinase K (10 mg/µL in PBST) for 45 minutes at room temperature. After washing with PBST the larvae were fixed again in p-formaldehyde (4%) for 20 minutes and washed again 5 times with PBST. Larvae were blocked in blocking solution (for IHC) for 1 hour and incubated over night with the primary antibody (rabbit anti-Th, dilution 1:400). The following day, the larvae were washed several times for 30 minutes in PBTD and then incubated over night with the secondary antibody (goat anti-rabbit Alexa 488, dilution 1:1000). On the third day, the larvae were washed four times for 15 minutes in PBTD and four times in PBST.

Genotyping of single embryos was done by cutting the tale of each larvae and making gDNA as described for making gDNA of fin-clips, however lysis buffer 2 was used. A PCR and digest was then performed, as described for screening of TALENs, to identify homozygous and hetereozygous mutants and wildtype siblings. The heads of the embryos were mounted dorsally in mounting medium for IHC and imaged with the Zeiss LSM510 fluorescent confocal microscope making z-stacks. The stacks were then projected into one final image using ZEN software (2012), and fluorescent signal between wildtype and homozygous mutant embryos were compared.

3.6 Whole Mount In Situ Hybridization

Probe Synthesis Antisense *th*, *dat*, *hb9* and *isl1* RNA probes were synthesized by in vitro transcription from plasmid DNA. 5 µg plasmid was first digested with the appropriate enzyme, XbaI for *dat* and *isl1*, XhoI for *th* and BamHI for *hb9*, over night. The DNA was then extracted with the Cycle Pure Kit and diluted in 15 µL mpH_2O. The in

vitro transcription reaction was then set up as follows:

1 µg	DNA
4 µL	Transcription buffer 5x
2 µL	NTPs (dig labelled)
2 µL	Enzyme (T7 for *dat* and *hb9* and T3 for *th* and *isl1*)
1 µL	RNAse inhibitor
up to 20 µL	H_2O

The reaction was incubated at 37°C for 2 hours. 2 µL DNAse was then added and after 15 minutes, the reaction was stopped by adding 1 µL EDTA. The RNA was precipitated by adding 2.5 µL LiCl (4M) and 75 µL pre chilled 100% EtOH (mixed thoroughly) and cooled at -20°C over night or at -80°C for at least 1 hour. Next, the mixture was centrifuged at full speed for 30 minutes at 4°C. The supernatant was decanted and the pellet washed in 70% EtOH by centrifuging again for 10 minutes. Finally, the supernatant was decanted, the pellet shortly dried and then re-suspended in 20 µL mpH_2O and 20 µL formamide. Probes were stored at -20°C.

Whole Mount In Situ Hybridization Two or three days old zebrafish embryos were fixed in p-formaldehyde (4%) as described for IHC. For whole mount in situ hybridization (WISH), the embryos were dehydrated with decreasing concentrations of methanol in PBST before they were incubated in proteinase K (10 mg/µL in PBST) at room temperature for 30 (2 dpf) or 45 (3dpf) minutes in a turning wheel. After washing twice in PBST, the embryos were re-fixed in p-formaldehyde (4%) for 20 minutes at room temperature and then washed 5 times in PBST again.

Hyb was used to pre-hybridize for 2 hours at 65°C in a water bath before it was replaced with Hyb+ containing either the *th*, *dat*, *hb9* or *isl1* probe (1:300), and the embryos were incubated over night at 65°C. The following morning, the embryos were washed accordingly at the hybridisation temperature:

20 min	Hyb
2x 20 min	50% Formamid in 2xSSCT
20 min	25% Formamid in 2xSSCT
2x 20 min	2xSSCT
3x 30 min	0.2xSSCT
5 min	PBST

After washing, blocking solution (2mg/mL BSA, 5% sheep serum in PBST) was added and incubated for 1 hour at room temperature. Embryos were then incubated over night with the primary antibody (anti-digoxigenin-AP) at 4°C. On the next morning the embryos were washed 7 x 15 minutes in PBST and 1 x 5 minutes in 0.1M TrisHCl (pH 9.5). Samples were then moved to 6 or 12 well-plates and equilibrated 3 x 5 minutes in freshly prepared NTMT buffer and left to stain in the dark. The staining time varied between probes and the embryos were monitored during the incubation time to obtain an optimal staining. For the *th* and *dat* probe 2 hours gave good results, while *isl1* and *hb9* were stained for 3 hours at room temperature, then over night at 4°C. The staining was stopped with stop-solution and the embryos were re-fixed in p-formaldehyde (4%) for one hour. Finally 80% glycerol in PBST with 1mM EDTA was added to the embryos

for storage at 4°C.

Single fish were genotyped as described for IHC, however for the fish labelled with *Isl1* or *HB9* the head of the embryo was used to make gDNA instead of the tail. Pictures were taken with the Zeizz Axioskop 2 of the head of *th* and *dat* stained embryos both dorsally and laterally, while the spine of *hb9* and *isl1* stained embryos was documented from the lateral side. The amount of positive cells were qualitatively compared between homozygous mutants and their wildtype siblings.

3.7 Prepulse Inhibition of the Startle Response

The assay of prepulse inhibition of the startle response was executed using acoustic stimuli elicited by a minishaker as described by Burgess and Granato (*Fig. 15*) (Burgess and Granato, 2007). The larvae were placed in a 6cm Petri dish containing a plastic grid (3x3) separating the larvae making identification of single larvae possible after the experiment. The Petri dish was then placed on a small vibration exciter controlled by a digital-analogue card (PCI-6221; National Instruments). The timing and waveforms of the startle stimuli were generated by custom software (DAQtimer) developed by Dr. Burgess. Three different startle stimuli were generated; a weak prepulse stimulus, a strong startle stimulus, and a sequence of prepulse and startle stimuli. Each larvae were exposed to each of the stimuli 20 times in a pseudorandom order. Films of the fish response to the acoustic stimuli were made by a high speed camera (FASTCAN-1024PCI, Photron Europe LTD) recording 1000 frames per second. All recordings were made in the dark, using infra-red lighting as a light source for the camera.

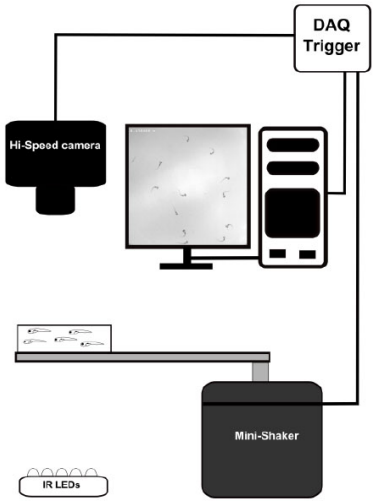

Figure 15: Schematic drawing of the apparatus for assaying the prepulse inhibition of the startle response using acoustic stimuli (Fernandes, 2013).

After the experiments, the fish were genotyped and the film data was analyzed by "Flote" software (Burgess and Granato, 2007). The % of SLC was calculated as the mean of the % response of each fish to each stimuli. Groups were then made based on stimulus and genotype and comparations of mean % SLC were made. The precentege of PPI was calculated as 100 X (percentage of fish responding to startle stimulus - percentage responding to prepulse and startle sequence)/(percentage responding to startle stimulus).

For the calculations of mean % PPI and mean % SLC, all larvae whose % SLC was greater than 90% or less than 10% were excluded. Startle responses were compared between mutants and wildtype using the Mann–Whitney U test.

3.8 CRISPRs

3.8.1 Making Cas9 mRNA

Cas9 plasmids were obtained from Addgene or from Filippo del Benes laboratory in Paris (*Table 4*). Plasmids were re-transformed into DH5α cells, plated on Amp^{50} agar plates, and left to incubate over night at 37°C. On the following day single colonies were picked to be grown in 50mL Amp YT broth at 37°C over night in a shaker. Plasmids were then isolated from the bacteria using the E.Z.N.A Plasmid Midi Kit (OMEGA)(Protocol 1) or Plasmid Midi Kit (Quiagen)(Protocol 2 and 2b).

5-10 µg of the Cas9 plasmids were linearized at 37°C for two hours (Protocol 1) or over night (Protocol 2 and 2b) with the restriction enzyme listed below:

pCS2-nCas9n	BssHII
Cas9 Paris	BbsI
MLM3613	PmeI
JDS246	PmeI/BbsI

2 µL of the digests were ran on a gel to check for complete digestion before the digests were cleaned using the Cycle Pure Kit and diluted in 40 µL 0.5x TE (Protocol 1) or 15 µL mpH_2O (Protocol 2 an 2b). The concentrations were measured with NanoDrop. The linearized Cas9 plasmids were transcribed in vitro using mMessage Machine T7 ULTRA kit following the manufactures protocol of a capped transcription reaction assembly and a poly(A) tailing procedure followed by a LiCl precipitation (Protocol 1 and 2). The finished mRNAs were diluted in 40 µL 0.5x TE (Protocol 1) or 60 µL mpH_2O (Protocol 2). 2 µL of the mRNAs were ran on an agarose gel as a control after being denatured for 90 seconds at 75°C (Protocol 1) or with formaldehyde loading dye (Protocol 2). The mRNA was finally stored at -80°C.

As a variant of protocol 2 (Protocol 2b), the linearized JDS246 Cas9 plasmid was transcribed at 20°C over night with a short poly A tale reaction of only 15 min. The mRNA was then recovered as in protocol 2.

3.8.2 guideRNAs

Three already published guide RNA (gRNA) plasmids, plus a gRNA expression vector backbone (DR274) were obtained (*Table 8*). In addition, four guide RNAs targeting the *th* and *dat* loci were designed using the online program ZiFiT (`http://zifit.partners.org/ZiFiT/`) or chopchop (`https://chopchop.rc.fas.harvard.edu/`) (*Table 9*) (Sander et al., 2007; Montague et al., 2014).

Creation of gRNAs by cloning gRNA ThEx2 was designed to cut at a NotI restriction enzyme site while the two gRNAs targeting the *dat* locus (DAT1 and DAT2) were designed to together cut out an ApaI restriction enzyme site. The 20 nt specific part of the gRNAs were ordered as 2 complimentary oligonucleotides from Sigma-Aldrich with

gRNA	Target site	Plasmid
th	GGATGCGCGTAAGGAGCGCG*AGG*	Zebrafish-gRNA-0007 (Hwang et al., 2013)
eGFP Paris	GGCGAGGGCGATGCCACCTA*CGG*	eGFP gRNA (Auer et al., 2013)
eGFP Addgene	*CCA*CCGGCAAGCTGCCCGTGCCC	pT7EGFPgRNA (Jao et al., 2013)

Table 8: Target sites for published gRNAs.

gRNA	Target Site	Oligos 5'-3'
ThEx2	GGAGCGCGAGGCGGCGGCCGCGG	*TAGG*AGCGCGAGGCGGCGGCCG *AAAC*CGGCCGCCGCCTCGCGCT
DAT1	GGCGCTGCTGGAGCCGCTGA*CGG*	*TAGG*CGCTGCTGGAGCCGCTGA *AAAC*TCAGCGGCTCCAGCAGCG
DAT2	GGTGGAGCTGGTGCTGGTGA*AGG*	*TAGG*TGGAGCTGGTGCTGGTGA *AAAC*TCACCAGCACCAGCTCCA
ThEx1	*CCG*AATTCAAGCAGCTCCACATC	

Table 9: gRNA target sites and oligonucleotides for their cloning into the DR274 backbone.

the appropriate overhangs for ligation into the gRNA expression vector backbone DR274 (*Table 9*). The gRNA oligonucleotides were diluted to 100 µM with mpH_2O and mixed 1:1 with with annealing buffer to a final concentration of 2 µM. The mix was heated for 5 minutes at 95°C on a heating block before the heating block was turned off and the samples were cooled down with the block for 2 hours. The annealed oligonucleotides, termed the linkers, were then phosphorylated with a T4 Polynucleotide Kinase (PNK) as follows:

Linker	16 µL (32pmol)
PNK	1 µL
10X PNK buffer	2 µL
10mM ATP	1 µL
Incubate at:	
37°C	30 min.
Inactivate kinase at:	
65°C	15 min.

5 µg of DR274 was digested over night at 37°C with BsaI. The open plasmid was then cleaned with the Cycle Pure Kit and a linker ligation reaction was set up:

Phosphorylated linker (ThEx2 / DAT1 / DAT2)	1.5 µL
DR274 (digested and cleaned)	14.5 µL
10x T4 ligase buffer	2 µL
T4 ligase	2 µL

The ligation was left over night at room temperature and on the following day, the whole ligation was transformed into DH5α cells and let to grow over night on Kan^{50} agar plates. 10 colonies from each plate were screened for the insertion of the linker using a colony PCR. 5 primers had been designed, two binding to the DR274 backbone (M13_fwd and DR274_rev) and three binding to each of the inserted linkers (ThEx2_gRNAtest / DAT1_gRNAtest / DAT2_gRNAtest). This would give one band in negative samples and two bands in positive samples containing the inserted linker. Annealing temperature was set to 52°C. The PCR was then run on a 1% agarose gel and three positive colonies from each ligation reaction were selected for over night culture in 3.5mL Kan^{50} YT broth.

After isolation with the DNA Plasmid Mini Kit, the plasmids were sent for sequencing with the primer M13_fwd. One positive sample for each gRNA was then selected for a 50mL Kan^{50} YT broth over night culture, and plasmids were isolated with the Plasmid Midi Kit (Quiagen) on the next day.

Finished gRNA plasmids and gRNA plasmids obtained from Addgene or Paris (7 µg) were digested with DraI over night and ran on a gel to confirm complete digestion. gRNAs were then transcribed using mMessage Machine T7 MEGA kit following the manufactures protocol (Transcription reaction assembly) using between 5 and 8 µg of digested and purified (Cycle Pure Kit) plasmid. The reactions were incubated for four hours, and the gRNAs were recovered using the RNeasy Mini Kit (Quiagen) diluting two consecutive times in 30 µL mpH_2O. 2 µL of each gRNA were ran on a gel to confirm successful transcription and the gRNAs were then stored at -80°C.

Cloning free generation of gRNAs gRNA ThEx1 was designed to cut at the Th TALEN target site at the EcoRI recognition site. To generate the gRNA, two oligonucleotides were ordered from Sigma-Aldrich with HPLC purification. One oligonucleotide was gene specific and contained the target sequence preceded by the SP6 promoter region and followed by an overlap region (*Table 10*). The second oligonucleotide contained the constant region of the gRNA including the overlap region.

Target Specific Oligo	ATTTAGGTGACACTATA*GATGTGGAGCTGCTTGAATT*GTTTTAGAGCTAGA AATAGCAAG
Constant Oligo	AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTA TTTTAACTTGCTATTTCTAGCTCTAAAAC

Table 10: Oligonucleotides for gRNA ThEx1.

The oligonucleotides were diluted in mpH_2O to 100 µM and 1 µL of each oligonucleotide was then mixed with 8 µL mpH_2O. The annealing took place in a thermocycler:

1. 95°C 5 min
2. 95-85°C -1.5°C/sec
3. 85-25°C -0.1°C/sec

To fill in the annealed oligonucleotides, a T4 polymerase reaction was set up:

2 µL	Tango buffer
1 µL	2.5nM dNTPs
0.5 µL	T4 DNA polymerase
6.5 µL	mpH_2O

The reaction was incubated at 11°C for 20 min and was then cleaned with the Cycle Pure Kit. The gRNA was then transcribed using the MEGAscript SP6 Transcription Kit (Ambion) following the manufacturers protocol and recovering the gRNA with the RNeasy Mini Kit, diluting twice in 20 µL mpH_2O.

3.8.3 Injection and screening of CRISPRs

1nL Cas9 mRNA and gRNA were injected into one cell stage zebrafish embryos in different amounts, ranging between 150-450 ng/µL for Cas9 and 12.5-300 ng/µL for the gRNAs.

Injections were performed as described in table 11. Phenol red was used as a marker for injection volume and location.

Inj.	gRNA	Cas9	Embryos
1	th (12.5 ng/µL)	pCS2-nCas9n[1] (300 ng/µL)	ABTL x ABTL
2	th (12.5 ng/µL)	Cas9 Paris[1] (300 ng/µL)	ABTL x ABTL
3	th (12.5 ng/µL)	MLM3613[1] (300 ng/µL)	ABTL x ABTL
4	eGFP Paris (12.5 ng/µL)	Cas9 Paris[1] (300 ng/µL)	otpb:gfp x ABTL
5	eGFP Paris & Addgene (12.5 ng/µL)	MLM3613[1] (150 ng/µL)	otpb:gfp x ABTL
6	eGFP Paris (12.5 ng/µL)	MLM3613[1] (150 ng/µL)	otpb:gfp x ABTL
7	eGFP Addgene (12.5 ng/µL)	MLM3613[2] (150 ng/µL)	otpb:gfp x ABTL
8	eGFP Paris (21 ng/µL)	MLM3613[2] (150 ng/µL)	otpb:gfp x ABTL
9	eGFP Paris (35 ng/µL)	MLM3613[2] (300 ng/µL)	otpb:gfp x ABTL
10	eGFP Paris (35 ng/µL)	JDS246[2] (300 ng/µL)	otpb:gfp x ABTL
11	eGFP Addgene (35 ng/µL)	JDS246[2] (300 ng/µL)	otpb:gfp x ABTL
12	ThEx2 (28 ng/µL)	JDS246[2] (300 ng/µL)	ABTL x ABTL
13	DAT1 & DAT2 (300 ng/µL)	JDS246[2] (655 ng/µL)	ABTL x ABTL
14	DAT1 & DAT2 (300 ng/µL)	JDS246[2b] (455 ng/µL)	ABTL x ABTL
15	ThEx1 (22 ng/µL)	JDS246[2b] (150 ng/µL)	ABTL x ABTL
16	eGFP Paris & Addgene (200 ng/µL)	JDS246[2b] (455 ng/µL)	otpb:gfp x ABTL

1: Cas9 Protocol 1, 2: Cas9 Protocol 2, 2b: Cas9 Protocol 2b

Table 11: List over injections preformed for the establishment of the CRISPR/Cas9 system.

The screening method for the th gRNA from Addgene was through a T7 endonuclease assay. The targeted locus was first amplified from injected and uninjected control gDNA by PCR using the primer mix CRISPR_thAG and an annealing temperature of 53°C. The PCR products were then cleaned with the Cycle Pure Kit and the concentrations were measured by NanoDrop. Between 200 and 500ng DNA was mixed with 2 µL NEBuffer 2.1 and mpH_2O up to 19 µL. The reactions were put in a thermocycler with the following cycle:

1. 95°C 5 min
2. -2°C/sec down to 85°C
3. -0.1°C/sec down to 25°C
4. 4°C

The samples were afterwards put on ice and 1 µL T7 endonuclease was added to the reactions which were left to incubate at 37°C for 15 minutes. The reactions were stopped by adding 2 µL 0.25M EDTA. The samples were finally loaded directly on a 1.5% agarose gel. Samples containing a mutation should have been cut by the endonuclease and have more than one band on the gel. Control samples should remain uncut.

Screening for knock-down of eGFP expression by the eGFP gRNAs was done using the Leica MZ16FA epifluorescent stereo microscope comparing the eGFP expression of injected embryos to uninjected controls 2 days post fertilization (dpf). Embryos were released from their chorions and put into a mounting solution containing tricain and methyl cellulose in a small Petri dish. Pictures were then taken, either ventrally of the brain, or laterally of the spine.

The screening method for the ThEx1, ThEx2 and gRNAs DAT1 and DAT2, was through restriction enzyme digests. Target loci were first amplified with a PCR using primer

mixes Th_TAL for the ThEx1 target site, DAT_TAL for the DAT1 and DAT2 target site and ThEx2 for the ThEx2 target site. The restriction digests were set up using 15 µL of the PCR products and the enzymes EcoRI (ThEx1), NotI (ThEx2) or ApaI (DAT1 and DAT2). The digests were then incubated for 2 hours at 37°C (EcoRI and NotI) or 25°C (ApaI) before they were ran on an 1.5% agarose gel. As for the screening of TALENs, the intensity of the uncut band compared to the cut bands in injected samples demonstrates the CRISPR cutting efficiency.

3.9 The Creation of Knock-in Lines

3.9.1 Design and Cloning of Donor Plasmids

Donor plasmids for knock-in experiments were designed to contain the wanted transgene for insertion flanked by Th or DAT TALEN target sites at each end (*Fig. 16*). 3' of the transgene, before the TALEN target site, a LoxP site was inserted to later be able to cut out potential concatemers using the cre recombinase. The plasmid pME-MCS was used as vector backbone. Together, three plasmids were planned for each locus, one with an eGFP ORF, one with a Gal4 ORF and one with a PhiC31 landing site (attP site).

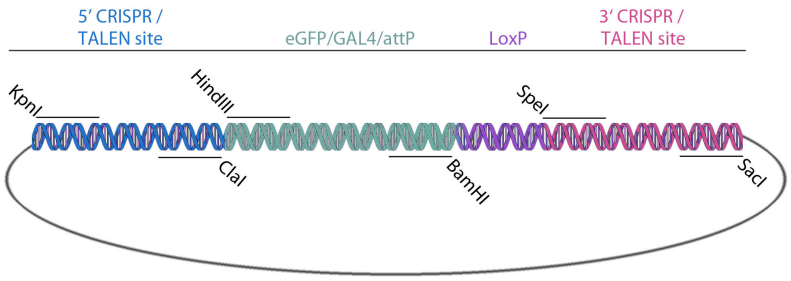

Figure 16: Design of donor plasmids for the targeted knock-in of transgenes into the *th* and *dat* loci.

Complementary LoxP site oligonucleotides were ordered from Sigma-Aldrich with BamHI and SpeI restriction site overhangs. The LoxP oligonucleotides were diluted, annealed and phosphorylated as described for gRNA oligonucleotides for cloning. 5 µg of the plasmid pME-MCS was digested for 2 hours at 37°C with BamHI-HF and SpeI. The open plasmid was then cleaned using the Cycle Pure Kit and a linker ligation reaction was set up:

Phosphorylated linker (LoxP)	1.5 µL
pME-MCS digested and cleaned	14.5 µL
10X T4 ligase buffer	2 µL
T4 ligase	2 µL

The ligation was left over night at room temperature and on the following day, the whole ligation was transformed into DH5α cells and let to grow over night on Kan^{50} agar plates. 7 colonies were picked, plasmids were isolated and they were sent for sequencing with the primer SeqA. One clone containing the linker was then grown in 50mL Kan^{50} YT broth and plasmids were isolated on the next day with the DNA Plasmid Midi Kit (OMEGA).

The TALEN target site sequences to be cloned 3' and 5' of the transgene were isolated

from genomic DNA. Primers were ordered with the overhangs KpnI and ClaI (Primer pairs Th1 and DAT1) for the 5' TALEN site and SpeI and SacI (Th2 and DAT2) for the 3' target site. For the Th primers the annealing temperature was set to 51°C and for DAT the annealing temperature was set to 56°C. Gal4, eGFP and the attP landing site were obtained from plasmids pME-Gal4VP16, pME-EGFP-CAAX and pKhspS2MOS respectively, using PCR and primer pairs eGFP (annealing temperature 56°C), Gal4 (annealing temperature 51°C) and attP (annealing temperature 56°C) with HindIII and BamHI overhangs for insertion into the pME-MCS-LoxP backbone. PCR products were digested with the appropriate enzymes for 2 hours at 37°C and cleaned with the Cycle Pure Kit while the backbone was digested over night, ran on a 0.7% agarose gel and cleaned through gel extraction with the kit NucleoSpin Gel and PCR Clean-up. All products were diluted in 40 µL mpH_2O. All inserts were phosphorylated before ligation as described earlier for phosphorylation of linkers.

The cloning was performed step by step starting with the 5' TALEN target site followed by the 3' TALEN target site and finishing with the Gal4, eGFP or attP inserts. The backbone was re-digested with enzymes depending on which insert was to be ligated and purified after each step. Ligation reactions were set up following the equation:

$$\frac{\text{60ng backbone} \times \text{bp insert} \times \text{ratio}}{\text{bp backbone}} = \text{ng insert}$$

For the shorter inserts (TALEN target sites, attP and Gal4) a ratio of 1:5 vector insert was chosen, while a ratio of 1:3 was used for the ligation of eGFP.

60ng backbone	Variable
Insert	Variable
10X T4 ligase buffer	2 µL
ATP (10mM)	2 µL
T4 ligase	2 µL
H_2O	Up to 20 µL

The ligations were left for 2 hours before 3 µL were transformed into DH5α cells and plated on Kan^{50} agar plates. Colony PCR was used to screen for positive clones on the following day using the same primers that were used for the isolation of the different pieces for the cloning. Positive clones were grown over night in 4.5mL Kan^{50} YT broth at 37°C, and plasmids were purified (Cycle Pure Kit) and sent for sequencing on the next day. The final donor plasmids containing a LoxP site, two TALEN target sites and a transgene were grown in 50mL Kan^{50} YT broth over night, and the plasmids were purified with the Plasmid Midi Kit (Quiagen).

3.9.2 TALEN Donor Plasmid Knock-in

1nL containing 200 ng/µL of each TALEN mRNA (Th3 or DAT3) and 20 ng/µL of the eGFP donor plasmid were injected into one cell stage ABTL zebrafish embryos. gDNA was made at 2 dpf and a PCR was preformed with one primer binding inside the insert (eGFPinsert_rev) and one binding in the genome (Th_TALfwd / DAT_TALfwd), yielding a PCR product only if the insertion was successful. Screening for eGFP insertion was also done by looking for green fluorescent cells under the Leica MZ16F epifluorescence stereo microscope 2 dpf. To make sure the TALENs used still could produce a DSB, a

PCR and restriction digest was performed on the targeted locus using EcoRI for Th and SphI for DAT.

3.9.3 TALEN Oligonucleotide Knock-in

As a second knock-in strategy, a single stranded attP oligonucleotide was designed for the insertion after a cut at the Th TALEN / ThEx1 gRNA target site. 50 core nucleotides of the attP landing site was flanked by 15 nucleotides homology arms surrounding the Th TALEN / ThEx1 gRNA cutting site (*Table 12*). The oligonucleotide, which was ordered from Sigma-Aldrich, was re-suspended in 10mM Tris.

attP-th donor oligonucleotide	*CATCTGAACATGCCG*GTAGTGCCCCAACTGGGGTAACCTTTGAGTTCTCT CAGTTGGGGGCGTAG*AATTCAAGCAGCTCC*

Table 12: Sequence of attP-th donor oligonucleotide. Homology arms are shown in italic.

3 or 6 µM attP-th donor oligonucleotide was injected together with 200 ng/µL of each Th3 TALEN. A PCR was preformed on the gDNA with the primers Th_TALfwd and Th_attPrev. They yield a product only after successful insertion. As a positive control of the primers, a PCR was also performed on the attP plasmid with the primers attP_rev and attP_plasFwd. Also in this case, the TALEN cutting efficiency was tested by a restriction enzyme digest using EcoRI.

3.9.4 CRISPR Donor Plasmid Knock-in

The first knock-in experiment using CRISPRs was a reproduction of the knock-in experiment described by Auer et al. (2013) where a GAL4 ORF was knocked into the eGFP locus of eGFP transgenic embryos with RFP expression under UAS control. Upon successful insertion, the before eGFP expressing cells would express RFP instead. For the experiment, a donor plasmid containing the Gal4 ORF (eGFP-GAL4) preceded by the gRNA eGFP Paris target site was received from Filippo Del Benes laboratory. The plasmid was re-transformed and grown on Kan^{50} agar plates over night and on the next day a single colony was selected to be grown over night in 50mL Kan^{50} YT broth. The plasmids were isolated using the Plasmid Midi Kit (Quiagen) the day before injection.

Inj.	**gRNA**	**Cas9**	**Donor Plasmid**	**Embryos**
KI-1	eGFP Paris (90 ng/µL)	JDS246^{2b} (500 ng/µL)	eGFP-Gal4 (70 ng/µL)	Tg(NeuroD:eGFP) x Tg(UAS:RFP, cry:eGFP)
KI-2	eGFP Paris (30 ng/µL)	JDS246^{2b} (150 ng/µL)	eGFP-Gal4 (35 ng/µL)	Tg(NeuroD:eGFP) x Tg(UAS:RFP, cry:eGFP)
KI-3	ThEx1 (30 ng/µL)	JDS246^{2b} (150 ng/µL)	Th-eGFP (35 ng/µL)	ABTL x ABTL
KI-4	ThEx1 (90 ng/µL)	JDS246^{2b} (300 ng/µL)	Th-eGFP (50 ng/µL)	ABTL x ABTL
KI-5	ThEx1, DAT1, DAT2 (30 ng/µL)	JDS246^{2b} (300 ng/µL)	Th-eGFP (35 ng/µL)	ABTL x ABTL
KI-6	ThEx1, eGFP Paris (30 ng/µL)	JDS246^{2b} (150 ng/µL)	eGFP-Gal4 (35 ng/µL)	ABTL x ABTL

2b: Cas9 Protocol 2b

Table 13: List over injections preformed for the establishment of a system for the targeted knock-in of transgenes.

1nL injection mix was injected into one cell stage Tg(NeuroD:eGFP) x Tg(UAS:RFP, cry:eGFP) embryos as described in table 13 (injections KI-1 and KI-2). At 2 dpf the embryos were analyzed with the Leica MZ16FA epifluorescent stereo microscope, looking for red expression in injected embryos. Positive embryos were dechorionated and placed in mounting medium together with uninjected control embryos. Pictures were taken of the head and upper part of the spine, both detecting green and red fluorescence. The images were later merged using the freeware ImageJ.

In the second knock-in experiment using CRISPRs, an eGFP ORF was intended knocked in at the *th* or *dat* locus of ABTL wildtype fish. The donor plasmid used was the Th-eGFP donor plasmid (originally designed for use with the Th TALENs) which contains an eGFP ORF flanked by Th TALEN / ThEx1 gRNA target sites. The plasmid was grown in 50mL Kan^{50} YT broth and was isolated using the Plasmid Midi Kit (Quiagen) the day before injection.

1nL injection mix was injected into one cell stage ABTL wildtype embryos as described in table 13 (injections KI-3, KI-4 and KI-5). 2 dpf the fish were analyzed with the Leica MZ16F epifluorescent stereo microscope looking for eGFP expression in the *th* or *dat* pattern. A PCR and restriction digest using Th_TAL primers and EcoRI or DAT_TAL primers and SphI was also performed to determine CRISPR cutting efficiency and to investigate if a large PCR product, indicating insertion, could be detected.

Last, a Gal4 ORF was intended knocked in at the *th* locus using the eGFP-GAL4 donor plasmid and the ThEx1 gRNA to create a double strand break in the genome and the eGFP Paris gRNA to linearize the plasmid. A PCR was performed on gDNA extracted from 24 hpf embryos using primers Th_TALfwd and Gal4_rev and an annealing temperature of 50°C. A PCR product would mean that an insertion had taken place.

4 Results

4.1 Targeted Mutagenesis

Sequencing of the *th* and *dat* TALEN target sites of ABTL wildtype fish showed that no single nucleotide polymorphisms were present which could interfere with TALEN binding (*Fig. 17* and *Appendix A2*). Digestion of the same loci with EcoRI (*th*) and SphI (*dat*) showed complete digestion of the PCR products.

Figure 17: Alignment of the refrence squence (ref) of the *th* and *dat* TALEN target site regions obtained from Ensamble Genome Browser (`http://www.ensembl.org/index.html`) and the sequencing result of the sequencing of wildtype ABTL fish from the fish facility (seq). No single nucleotide polymorphisms were identified which could interfere with TALEN binding.

All six TALEN pairs were successfully cloned according to sequencing results and their analysis. Injection of the TALEN mRNAs and the following restriction enzyme digest test showed that TALEN pairs 1 and 2, both for the *th* and *dat* loci, were unable to mediate an in/del mutation at their target sites (*Fig. 18*). Digests performed on gDNA from embryos injected with TALEN pairs Th3 or DAT3 had an undigested band, showing that the TALENs had caused an in/del mutation to occur at the target sites (*Fig. 18*).

Crossing out the adult G_0 generation to wildtype ABTL fish, 4 *th* mutant founders of 7 crosses were identified while 6 *dat* mutant founders were identified from 23 crosses. 70 Th F_1 fish were genotyped and 19 fish were classified as mutants carrying one of three mutated alleles (*Table 14*). 9 fish had an 11 bp deletion (allele no. m1404) resulting in a premature stop codon after 15 codons, 2 fish had a 1 bp insertion (allele no. m1405), leading to a premature stop codon after 19 amino acids, and 8 fish had a 7 bp deletion (allele no. m1403) removing the native ATG start codon (*Fig. 19*).

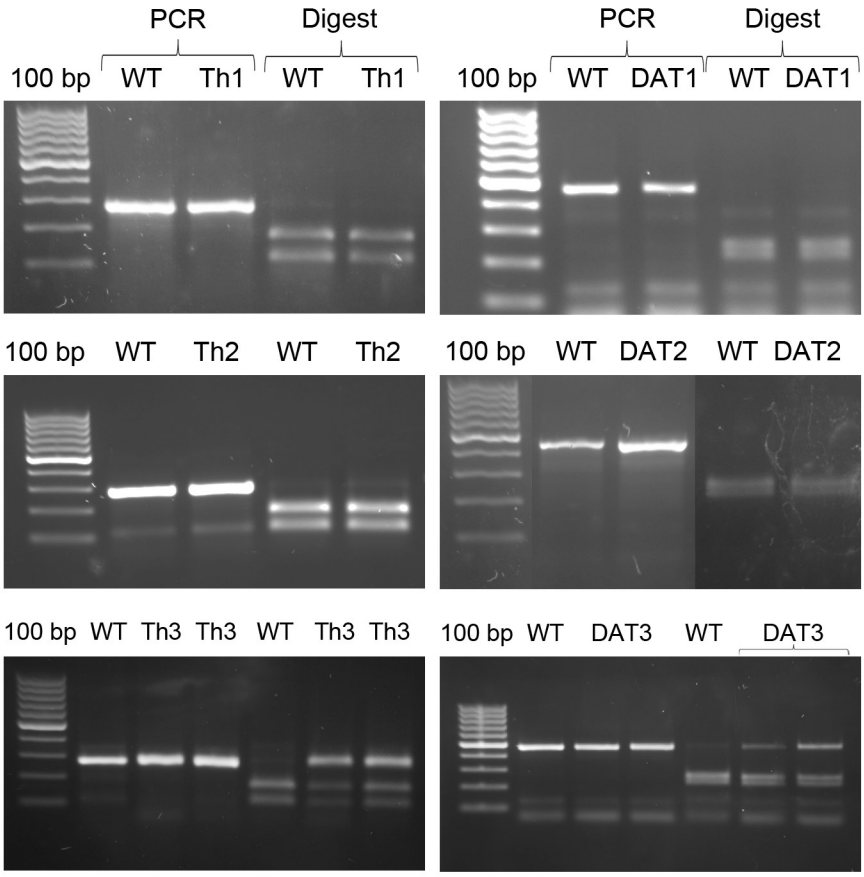

Figure 18: Agarose gel electrophoresis performed on the PCR products and digests of DNA from control embryos and embryos injected with TALEN mRNA. Digests of DNA from embryos injected with TALEN pairs 1 and 2 (both Th and DAT) showed a complete digest, while TALEN pairs 3 digests had undigested bands, indicating that an in/del mutation had occured.

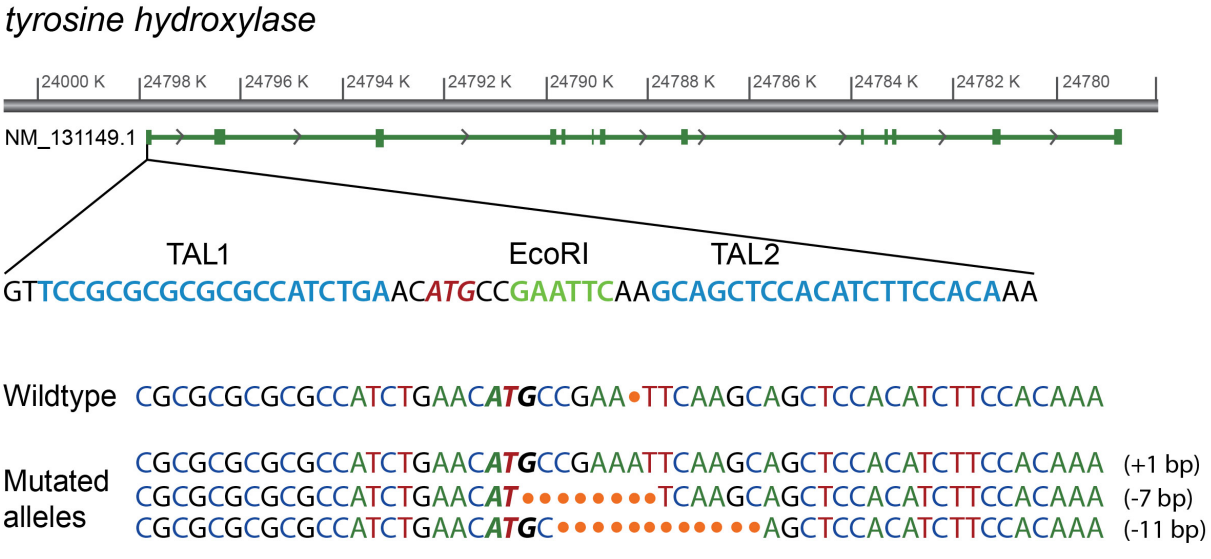

Figure 19: Genotyping of individual Th F_1 TALEN fish showed that three different in/del mutations had been induced by TALEN Th3. The native ATG is highlighted in bold/italics.

Allele	Name	Stock No.	Description
m1403	th	16181	TALEN induced mutation in tyrosine hydroxylase gene: 7 bp deletion at EcoR1 site in first exon.
m1404	th	16180, 16181	TALEN induced mutation in tyrosine hydroxylase gene: 11 bp deletion at EcoR1 site in first exon.
m1405	th	16181	TALEN induced mutation in tyrosine hydroxylase gene: 1 bp insertion at EcoR1 site in first exon.

Table 14: 3 different alleles were created by the injection of TALENs Th3.

Among 84 genotyped DAT F_1 fish, 17 mutants were identified with three different mutations (*Fig. 20*). 1 fish had a 10 bp deletion, 1 fish had a 7 bp deletion with a 1 bp insertion and 25 fish had a nine bp deletion, all of which removed the native ATG start codon.

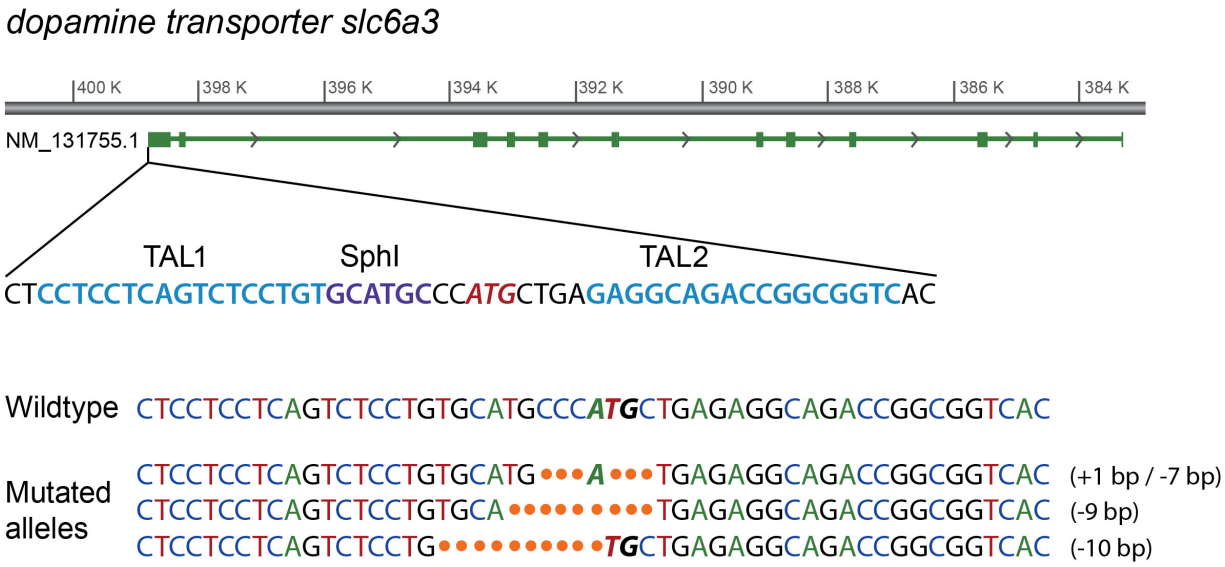

Figure 20: Genotyping of individual DAT F_1 TALEN fish showed that three different in/del mutations had been induced by TALEN DAT3. The native ATG is highlighted in bold/italics.

For the mutants where the native ATG had been deleted, the next in and out of frame ATGs were identified (*Fig. 21*). For *th* this was at bp 589 (amino acid (aa) 196) at the beginning of the proteins active aromatic amino acid hydroxylase domain. Transcription from this ATG would mean that 60% of the protein would be translated. The next out of frame ATG was found at bp 125, and a translation start here would lead to a stop codon after 18 aa.

For *dat*, the next in frame ATG was at bp 61 (aa 20). Transcription from this point would give 90% of the whole DAT protein. Additionally, a third in frame ATG was found 6 bp upstream of the native ATG.

4.1.1 Immunohistochemistry

IHC stainings using an anti- Th antibody are shown in figure 22. Homozygous mutants of the alleles th^{m1403} or th^{m1404} did not show fluorescent signal, while all WT and heterozygous siblings showed fluorescent signals in the *th* pattern, however to a variable degree.

Thus, a qualitative difference in Th expression between wildtype and homozygous *th* mutants was found.

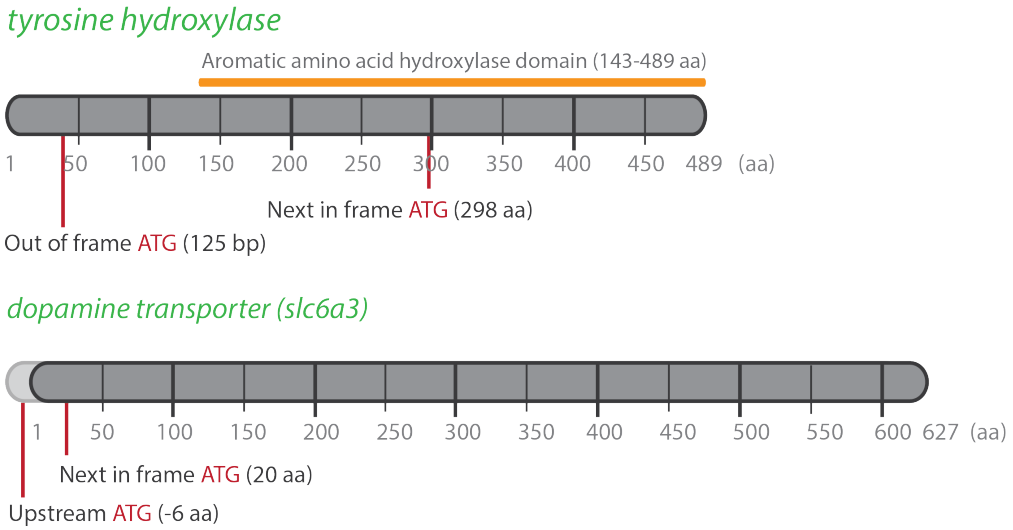

Figure 21: The next in frame ATG for the *th* and *dat* genes were found at bp 296 and 61 respectively. The *dat* gene has in addition an in frame ATG 6 bp upstream of the start codon.

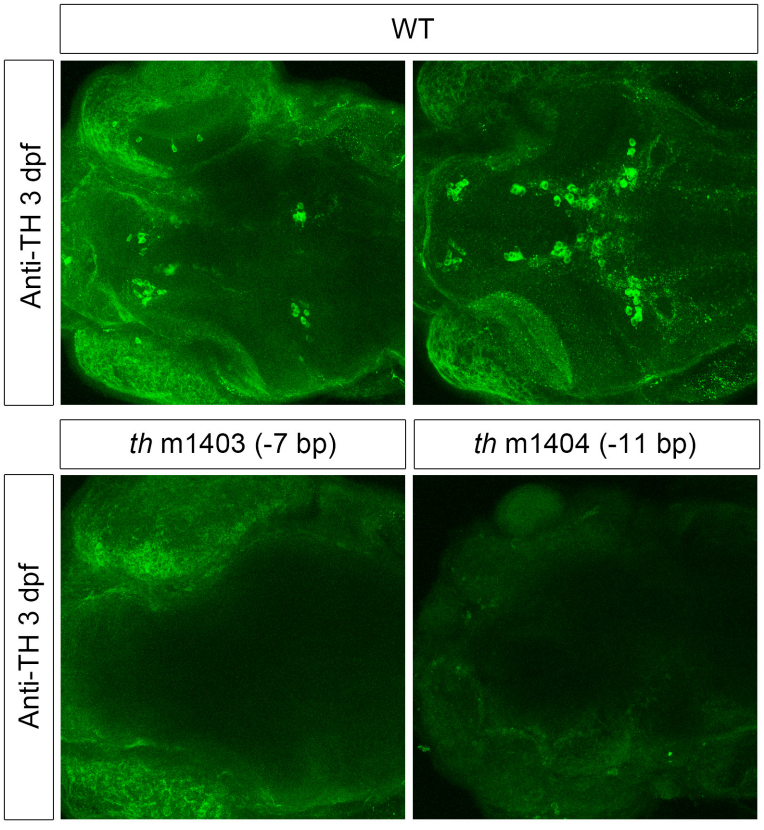

Figure 22: IHC of WT and *th* homozygous mutant zebrafish embryos, 3 dpf stained with Th antibody. Mutant embryos showed no fluorescent signal (n = 3 th^{m1403}, n = 2 th^{m1404}), while WT siblings (n = 13) showed fluorescent signal in the *th* pattern.

4.2 Establishment of the CRISPR/Cas9 System

The screening for mutations after injections 1, 2 and 3 (*Table 11*) was done with a T7 endonuclease assay which did not yield any clear or reproducible results. In most cases, both the control DNA and the DNA from the injected samples were cut by the T7 endonuclease. In other cases, too much DNA was lost during the assay, and no clear bands could be seen on the gel in neither of the samples. No conclusions regarding CRISPR efficiency could be drawn based on this assay.

Injections 4-11 were screened by looking at reductions of eGFP expression under a epifluorescent stereo microscope (Leica MZ16F or MZ16FA), focusing on the head region of the fish. No down-regulation was found in any of the injections. Images of injected and control embryos from injections 7 and 8 are shown in figure 23.

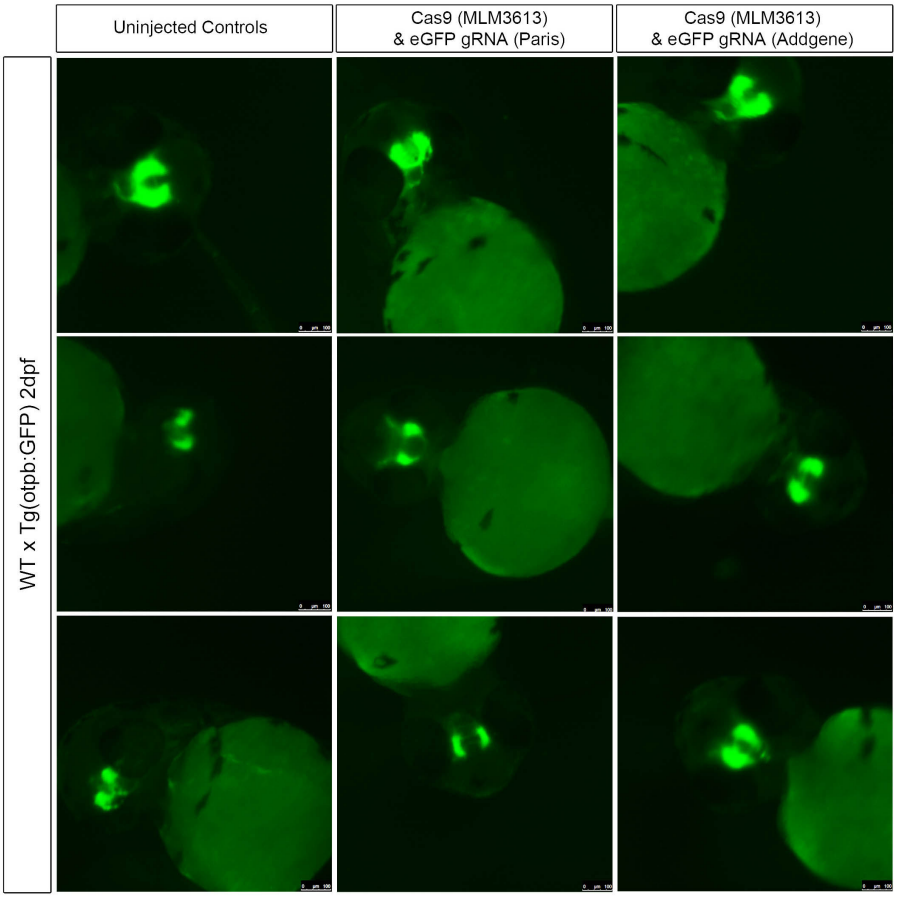

Figure 23: Otpb expression at 2 dpf of uninjected controls and embryos injected with MLM3613 Cas9 mRNA (protocol 2, 150 ng/µL) and gRNAs eGFP Paris or Addgene (12.5 ng/µL). No differences in eGFP expression could be detected.

Injections 12-15 were screened by restriction enzyme digests. For injection 12, with the ThEx2 gRNA, the assay was not conclusive due to incomplete digests also of the wild-type control samples. Injection 13 showed a negative result, while injection 14 and 15 had weak uncut bands in the digests of the injected samples indicating that a mutation had occurred at the target sites (*Fig. 24*).

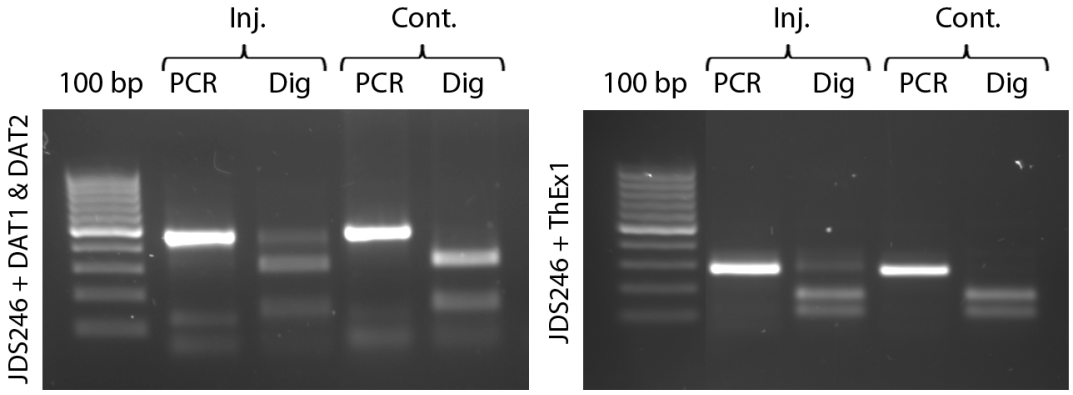

Figure 24: Agarose gel electrophoresis performed on the PCR products and digests of DNA from control embryos and embryos injected with JDS246 Cas9 mRNA (protocol 2b, 300 ng/µL) and ThEx1 gRNA or DAT1 and DAT2 gRNA (300 ng/µL). The injected samples show a weak undigested band, indicating the prescence of an in/del mutation.

Fluorescent imaging of embryos from injection 16 was done focusing on both the head region and the spine region of the embryos, and there was an indication of a slight down-regulation of the eGFP expression in the spine of the injected embryos as compared to wildtype (*Fig. 25*).

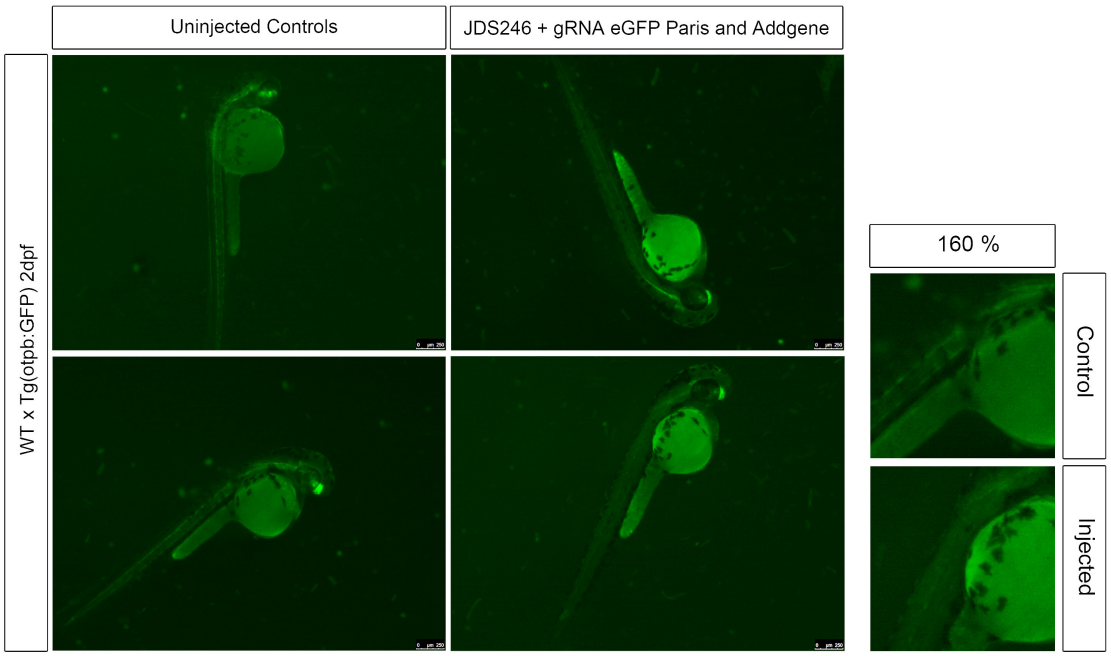

Figure 25: Otpb expression at 2 dpf of uninjected controls and embryos injected with JDS246 Cas9 mRNA (protocol 2b, 455 ng/µL) and gRNA targeting eGFP (Paris and Addgene) (200 ng/µL). A very slight downregulation of eGFP expression could be detected in the spine of the injected embryos.

4.3 The Creation of Knock-in Lines

4.3.1 Donor Plasmids

Six donor plasmids, three with the Th TALEN target site and three with the DAT TALEN target site, containing either an eGFP or Gal4 ORF or an attP landing site followed by a LoxP site, were successfully cloned, which was confirmed by sequencing (plasmid maps and sequence data is available in *Appendix A3*).

4.3.2 Knock-in Using TALENs

Screening of embryos injected with the Th-eGFP or DAT-eGFP donor plasmids together with TALENs Th3 or DAT3 with a epifluorescent stereo microscope (Leica MZ16FA) showed no fluorescently labelled neurons in the brain. Furthermore, for both Th and DAT, no positive PCR product was obtained when using primers where one primer bound inside the insert (eGFP) and one outside in the genome. Restriction enzyme digests confirmed that a double strand break and an in/del mutation was induced by the TALENs.

PCR performed on pooled embryos injected with the th-attP oligonucleotide together with the Th3 TALENs gave no PCR product for the primers where one bound inside the th-attP oligonucleotide sequence.

4.3.3 Knock-in Using CRISPRs

Injection KI-1 of the eGFP-GAL4 donor plasmid together with gRNA eGFP Paris in Tg(NeuroD:eGFP) x Tg(UAS:RFP, cry:eGFP) embryos gave no embryos with red fluorescent expression. Injection KI-2 however, gave embryos with red fluorescent expression in the eGFP pattern, confirming an insertion of the Gal4 transgene (*Fig. 26*). 1 of 6 injected embryos showed RFP expression.

Screening under the fluorescent microscope of embryos from injections KI-3, KI-4 and KI-5 did not reveal any green fluorescent cells in the *th* or *dat* expression pattern. PCR of embryos from injection KI-5 gave no bands when one primer was binding inside the Gal4 insert.

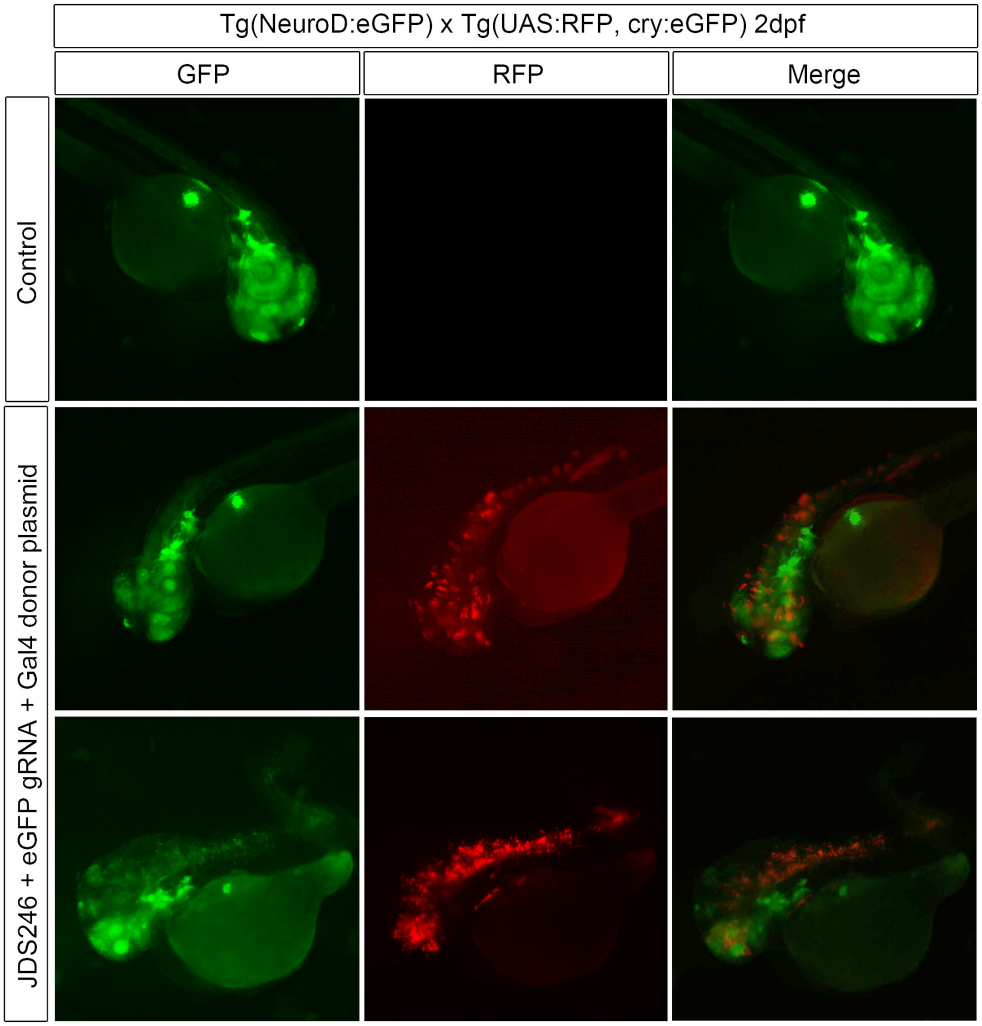

Figure 26: CRISPR/Cas9 mediated knock-in of GAL4 into the eGFP locus of Tg(neuroD:eGFP) x Tg(UAS:RFP, cry:eGFP) embryos injected with JDS246 Cas9 mRNA (protocol 2b, 150 ng/µL), gRNA eGFP (Paris) (30 ng/µL) and a GAL4 donor plasmid (40 ng/µL). The insertion of the GAL4 ORF is revealed by RFP expression in the eGFP pattern.

4.4 Dopamine and Neurogenesis

4.4.1 Whole Mount In Situ Hybridization

WISH stainings of spinal motor neurons in *th* homozygous mutants and wildtype siblings using *hb9* and *isl1* RNA antisense probes are shown in figure 27. The *hb9* probe stained the spinal motor neurons well, both in the 2 and 3 days old embryos. Qualitative analysis showed that a slightly stronger stain could be detected in the $th^{-/-}$ mutants as compared to wildtype siblings. The *isl1* probe did not stain the spinal motor neurons very well, and single stained cells were hard to identify making comparisons between wildtype and homozygous *th* mutants difficult.

WISH stainings of dopaminergic and noradrenergic neurons in *th* homozygous mutants and wildtype siblings using *th* and *dat* RNA antisense probes are shown in figure 28.

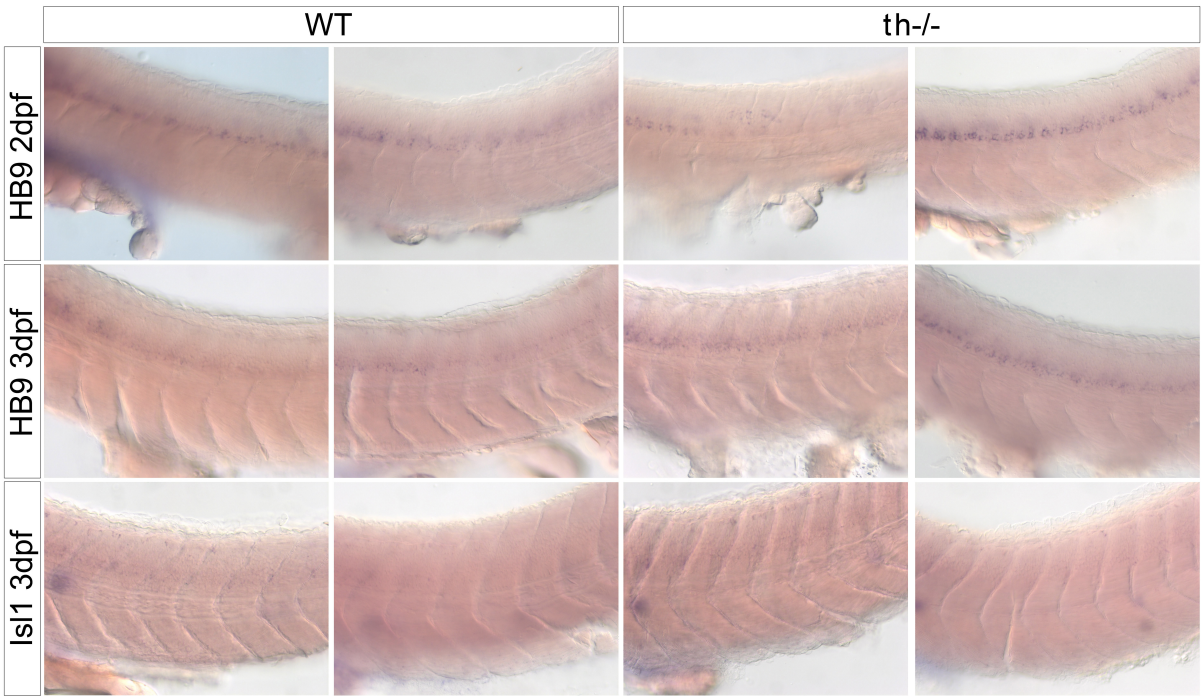

Figure 27: WISH of WT and $th^{-/-}$ mutant zebrafish embryos, 2 and 3 dpf stained with *hb9* and *isl1*. A slightly stronger *hb9* stain could be detected in the $th^{-/-}$ mutants as compared to wildtype siblings while no clear differences could be detected in the *isl1* stained embryos.

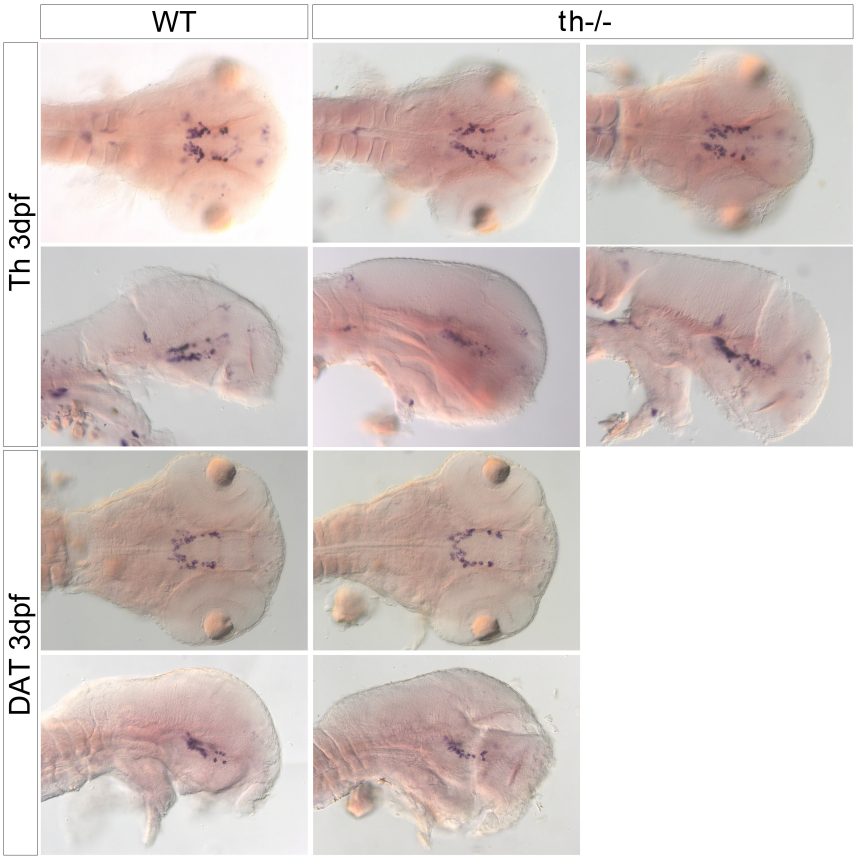

Figure 28: WISH of WT and $th^{-/-}$ mutant zebrafish embryos 3 dpf stained with *th* and *dat*. No major differences in cell number or position could be detected either in *th* or *dat* stained embryos.

Of embryos with the *th* stain, only one wildtype and two homozygous *th* mutants were identified while among the *dat* stained embryos only one wildtype and one homozygous mutant were identified. All major dopaminergic groups were present in both wildtype and homozygous mutants, and a first qualitative comparison between them did not show any major obvious differences in cell number or positions.

4.5 Dopamine and Prepulse Inhibition

Startle responses of single larvae to acoustic stimuli were filmed at 6 dpf by a high speed camera and tracked using the "Flote" software (Burgess and Granato, 2007). Tracking data showed that the tracking was not optimal as some larvae were not detected by the program and thus not tracked, and some movements were falsely tracked (*Fig. 29*).

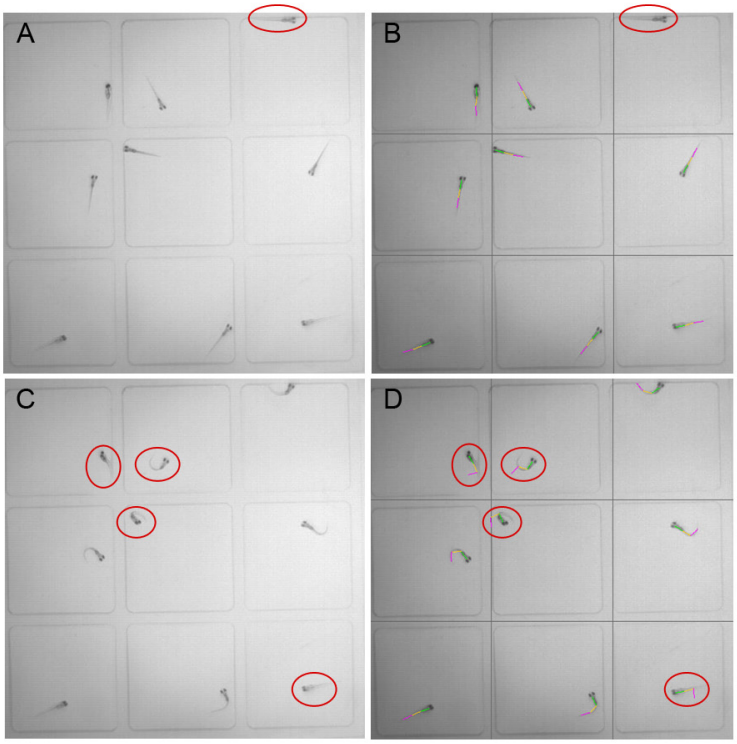

Figure 29: Detection and tracking of larvae movement. **A,B:** Non moving larvae with one fish which was not detected by the software (circeled in red).**C:** Short-latency C-start (SLC) startle responses where the larvae body form a C-bend before swimming. **D:** Software misinterpretation of larvae movements as can be seen by the coloured segments which do not follow the actual curvature of the larvae body (circeled in red).

Mean % SLC responses to weak (prepulse) and strong startle stimuli show that wildtype fish startle to a greater degree after a strong startle stimulus as compared to a weaker stimulus (30). SLC responses measured after a strong stimulus preceded by a weak prepulse (PPI) decreased the wildtype mean % SLC startle responses. Homozygous *th* mutants showed overall lower mean % SLC responses to the three different types of stimuli as compared to wildtype, however with the same trends (30). Due to low N numbers of experimental animals, the differences observed were not statistically significant.

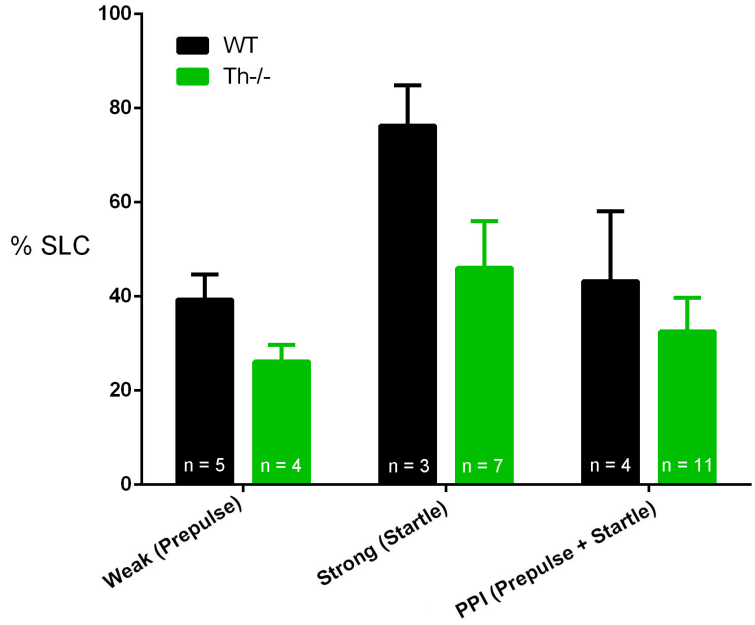

Figure 30: Mean % SLC responses of wildtype and *th* $^{-/-}$ larvae to weak, strong and prepulse and startle (PPI) sequence stimuli. *th* $^{-/-}$ larvae had overall less responses as compared to wildtype. All larvae showed a greater response to strong stimuli as compared to weak stimuli, and all larvae had less response to prepulse and startle sequence stimuli as compared to just the strong stimuli. p value > 0.05, error bars show standard errors of the mean.

Mean % PPI was calculated to be 61% for wildtype larvae and 44% for homozygous *th* mutants, indicating an impairment in PPI in *th* homozygous mutants. However due to the low N numbers of the experimental animals a statement of statistical significance could not be made.

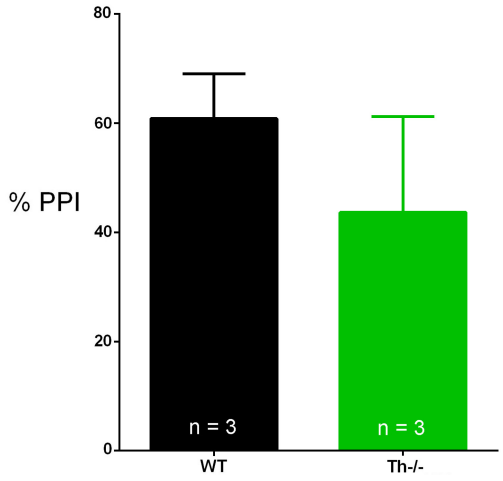

Figure 31: Mean % PPI was calculated to 61% for wildtype larvae and 44% for homozygous *th* mutants. p value > 0.05, error bars show standard error of the mean.

5 Discussion

5.1 Targeted Mutagenesis

5.1.1 TALEN Design

TALEN binding specificity, as earlier described, depends on the interaction of two amino acids, or the so called RVD, to a specific base of the DNA. Some RVDs bind strongly to one specific base, while others create weaker connections to their bases, and yet some RVDs can bind to more than one specific base (Cong et al., 2012). When designing a new TALEN pair, it is therefore suggested that at least every third or fourth RVD creates a strong connection to its base (Streubel et al., 2012). Several RVDs can be used to target guanine, and in the TALEN plasmids available for Golden Gate TALEN Assembly, the election was between RVDs NN and NH, both of which can also bind to adenine, however less frequently. In the Driever lab, only NN had been used before, and with success. The NH RVD was however shown to have both a slightly higher specificity for guanine and a stronger binding (Cong et al., 2012). For the first four TALEN pairs (Th1, Th2, DAT1 and DAT2) the NH RVD was therefore chosen. The TALEN target site finder Effector Nucleotide Targeter 2.0 used the NN RVD as a default, and this program was thus chosen for the design of the two first primer pairs (Doyle et al., 2012). For TALEN pairs number 3, the RVD NN and the TALEN targeter Mojo Hand was used (Bedell et al., 2012).

TALEN cutting efficiency, in addition to depending on the binding efficiency, also depends on the spacer length between the two TALENs of a pair (Schmid-Burgk et al., 2013). For a double strand break to occur, the FokI nucleases of the two TALENs have to meet in the middle of the spacer and form a dimer. Therefore, if the two TALENs bind either too close or too far away from each other, the nucleases might not dimererise and consequently no DSB will take place. Effector Nucleotide Targeter 2.0 did not have strict limits for the spacer length and suggested several TALEN pairs where the spacer length was up to 25 bases long. Mojo Hand however, had a maximum spacer length of 17 bases as default.

All TALEN pairs were targeted to the same locus around the native start ATG codon of either the *th* or *dat* locus. Thus, the target sites of the three TALEN pairs designed differed only slightly in the composition of RVDs and in spacer length. TALEN pairs Th1, DAT1 and DAT2 had spacers of 22, 25 and 21 bases respectively. These spacer lengths were probably too long, seeing as none of these TALENs succeeded in creating a DSB at the target site. The literature indicates that TALENs with a spacer length of 14 or 15 bases create DSBs most efficiently, and it is therefore likely to believe that with a spacer of over 20 bases, the FokI nucleases are not able to dimerise, and thus cannot create a DSB and consequently no in/del mutations are induced (Schmid-Burgk et al., 2013).

Additionally, TALEN Th2, with a spacer length of 17 bases failed to induce in/del mutations at the target site. However, in this case, a too long spacer length was most likely not the deciding factor since TALENs with a spacer length of 17 bases have repeatedly been proven to work in the lab (Driever lab, unpublished data). Possibly, the composition of RVDs was unfortunate and not enough specificity and or binding sufficiency was obtained in either of the TALENs of the pair, causing the TALENs not to bind properly to their

target sites.

TALEN pairs 3, with spacer lengths of 14 and 15 bases, induced mutations at the target site, as proven by the incomplete restriction enzyme digests. This indicates that a shorter spacer length is crucial for dimerization of the FokI nucleases and the creation of a DBS as well as the consequent induction of in/del mutations.

PCR and restriction enzyme digests of the gDNA from TALEN Th3 injected embryos showed that the uncut mutated band was only slightly stronger than the digested bands, arguing a quite low in/del inducing frequency. For the DAT3 TALENs, the mutated uncut band in the injected samples was generally weaker as the Th mutated bands, suggesting an even lower in/del inducing frequency, however considerable. This difference in efficiency between the Th and DAT TALENs could be derived from the different compositions of TALEN RVDs or also from differences in access to the gene in question within the genome.

5.1.2 TALEN Induced Mutations

F_1 generation TALEN fish were genotyped to identify single fish containing an in/del mutation and to determine the nature of this. Of the sequenced Th fish which originated from two different clutches, three different mutations were identified. Two of the mutations lead to a frameshift and a premature stop-codon (alleles m1404 and m1405). The third mutation (allele m1403) removed the native ATG, meaning that translation would no longer start here. A second in frame ATG, which could possibly be used by the translation machinery as a start codon, was found at bp 589. If translation would start here, about 60% of the protein would be translated. However, an out of frame ATG leading to a premature stop-codon was found at bp 125, meaning that transcription and translation would most likely begin here in the absence of the native ATG. As proven by IHC and staining of the Th protein, both mutations m1403 and m1404 lead to the kncokout of the *th* gene, signifying the creation of the first *th* knockout zebrafish line.

Among the genotyped DAT fish which originated from three different clutches, three different mutations were also identified, all of which removed the native ATG. However, in the *dat* gene, a second in frame ATG is present just two codons upstream of the native start ATG and this ATG was still present in the 7 bp deletion 1 bp insertion and in the by far most common 9 bp deletion. This meant that the *dat* gene most likely still would be transcribed from this upstream ATG, translating into a functional protein just lacking a few amino acids, something which does not result in the knockout of the gene. In the 10 bp deletion both the native and the upstream ATGs were removed, meaning that translation would not start here. The next in frame ATG though, was just 20 codons downstream from the original start codon. With a translational start there, 90% of the protein would be translated meaning that it most likely would be fully functional. Unfortunately, no out of frame ATG which could lead to a premature stop codon was identified before the next in frame ATG and thus no *dat* knockout zebrafish line could be started from the identified F_1 generation.

The analysis of the next in and out of frame ATGs in the *dat* gene demonstrated that the choice of TALEN target sites with the aim of creating a *dat* knockout was not optimal.

To obtain a knockout inducing in/del mutation at this target site the mutation would either have to not affect the native ATG and cause a frameshift, or it would have to take out both the native ATG and the second in frame ATG. Since TALENs often just induce short deletions (1-20 bps), the probability for them to induce a more than 60 bp deletion taking out the second ATG a well was close to none. For the TALENs not to affect the native ATG the target site should have been moved a few bases downstream, however this would mean that the restriction enzyme target site would not be affected by the in/del mutation and screening would have been made difficult. In summary, a different TALEN target site should have been selected to create a *dat* knockout mutant.

As an alternative, the fish from CRISPR injection no 14, where in/del mutations in the *dat* gene were induced could be used to obtain a *dat* knockout zebrafish line. The gRNAs DAT1 and DAT2 were targeting sites around bp 85 and bp 122 of the *dat* gene, meaning that the in/del mutations would occur without affecting the native ATG. If the mutations induced by the CRISPRs show to be out of frame leading to a premature stop codon, the fish could give rise to the *dat* knockout line. To achieve this, G_0 fish will be crossed out to ABTL wildtypes and the F_1 will be genotyped.

Analysis of the TALEN induced mutations showed a low variance among mutant alleles. This raised the question whether DNA repair after a DSB is biased, leading to a stereotyped repair in a genome specific fashion, thus reducing allele diversity. This was already claimed by Gagnon et al. (2014), who studied allele diversity after CRISPR induced DSBs and found that specific mutant alleles predominated in nearly all embryos. If the predominant allele does not lead to a frameshift, which was the case for the 9 bp deletion induced by the DAT3 TALENs, it makes the generation of knockout organisms more difficult. As a solution to this problem, Gagnon et al. (2014) suggest the knock-in of a stop codon oligonucleotide with a stop codon in all three reading frames ensuring a translation-terminating mutation regardless of DNA repair bias. This strategy, which was proven effective after a CRISPR induced DSB, should also work after a DSB created by TALENs.

5.1.3 Confirming the Knockout of *tyrosine hydroxylase*

To confirm the loss of Th protein in the identified *th* mutants, an IHC was performed on the F_2 generation of fish with allele numbers th^{m1403} and th^{m1404}. Fish identified with the allele th^{m1405} were not used for further experiments since they were only two individuals. Th antibody labelling showed that no Th protein is produced in *th* homozygous mutants of either of the relevant alleles, meaning that the TALEN induced mutations leading to premature stop-codons indeed also led to the knockout of the gene.

A confirmed *th* knockout organism should have no production of dopamine (DA) and thus no DA should be present in the body. However, as mentioned earlier, this was not the case for the *th* knockout mice, where relatively high levels of DA were found (Kobayashi et al., 1995). The origin of this DA is still not clear and experiments should be performed to determine if DA could be present in the *th* zebrafish knockout as well.

5.2 Establishment of the CRISPR/Cas9 System

5.2.1 Screening Methods

When working on the establishment of the CRISPR/Cas9 system, one of the first emerging challenges was to find a reliable and sensitive way for the screening of the CRISPR functionality and cutting efficiency. The recommended method for the screening of CRISPR cutting efficiency guided by the gRNA targeting the *th* locus, described by Hwang et al. (2013), was through a T7 endonuclease assay. Despite large efforts however, neither reproducible nor reliable results could be obtained using this assay. The problem was assessed to lay either in the denature and re-annealing process where strands might not be able to re-anneal properly creating mismatches also in wildtype DNA, or it is possible that the T7 endonuclease did not cut the mis-matched DNA strands specifically enough. When no optimization was achieved, it was decided that the T7 endonuclease assay could not be used for further screenings of CRISPR functionality and cutting efficiency.

An alternative method for the screening of CRISPR efficiency is to look at the knock-down of the expression of the gene targeted by the gRNA. Publications show, that when the targeted gene is a fluorescent reporter gene such as GFP, this can be done easily using an epifluorescent stereo microscope comparing the expression between injected and uninjected embryos (Jao et al., 2013; Auer et al., 2013). The GFP transgenic line Tg(otpb:gfp) used in this work has a strong GFP expression in the brain and weaker but clear expression in the spine. When looking for knock-down of GFP expression focusing on the brain, no changes could be detected in the Tg(otpb:gfp) x ABTL embryos injected with any Cas9 mRNA and any of the two already published GFP gRNAs used in this work. When focusing on the GFP expression in the spine on the other hand, signs of a very slight reduction could be observed after injection of Cas9 JDS246 (protocol 2b, methods section 3.7.1), and gRNA GFP Paris. However, the knock-down was far from as strong as obtained by Jao et al. (2013), and it was not strong enough to say with certainty that the CRISPR system was working and in fact was causing the faint reduction of GFP expression. To confirm that the CRISPR/Cas9 system was at all functioning, a more sensitive screening method was needed.

Consequently, it was decided to use a restriction enzyme digest, as used for the screening of TALENs, to also screen for CRISPR functionality and cutting efficiency. For this purpose, gRNAs targeting the *th* and *dat* loci were designed to cut at restriction enzyme recognition sites. Since gRNAs designed to be transcribed with the T7 promoter, which is the case for the gRNA backbone DR274 used in this work, need target sites starting with two guanine bases in addition to being followed by a PAM sequence, the possibilities were restricted (Hwang et al., 2013). For the *th* locus, this meant that for the gRNA target site additionally to contain a restriction enzyme target site, it had to be located in exon 2 of the gene. For the *dat* locus, no good target sites containing a restriction site could be found and two target sites close to each other flanking a restriction site were rather chosen. If the CRISPRs create a DSB at both target sites, the two DNA strands will be ligated back together without the piece containing the restriction enzyme site, and the loss of this will be visible in a restriction enzyme digest test.

Restriction enzyme digest analysis of the embryos injected with gRNA ThEx2 did not yield any clear results. An uncut band could always be seen in the control samples indicating an incomplete digest, something that made it near impossible to decide whether the uncut band in the injected samples were due to a mutation at the restriction site or an incomplete digest. Several measures were made to optimize the assay including cleaning all samples and leaving the digests over night, however non of this solved the problem.

The restriction digest of the samples from the injection of the two gRNAs targeting the *dat* locus gave a clear result confirming that the CRISPRs indeed had created DSBs at both target sites, eliminating the ApaI restriction site. The intensity of the uncut band was distinctly a lot weaker than the intensity of the digested bands indicating that the cutting efficiency of the CRISPRs was not very high. However, it is important to keep in mind that for a loss of restriction site, a DSB had to occur at both target sites in the same embryos and cells, giving reason to believe that the actual efficiency of each singe gRNA was higher than that shown by the restriction digest. To further assess the cutting efficiency of each DAT gRNA it would be possible to raise the injected embryos to adulthood, cross them out to wildtype and then sequence individual fish of the F_1 generation looking at which mutations actually have occurred and at which target site. However, since the DAT gRNAs were only created to establish the CRISPR/Cas9 system for site specific insertions, this was not a priority.

Due to the high sensitivity, restriction enzyme digests were finally elected as the preferred screening method for CRISPR cutting efficiency. This implicated that further gRNAs designed should somehow target a restriction enzyme recognition site.

5.2.2 Generating Cas9 mRNA

Establishing the CRISPR/Cas9 system, another major problem was to produce functional Cas9 mRNA. Cas9 mRNA is about 4 kb long, however the first transcription attempts using Cas9 protocol 1 (methods section 3.7.1) yielded mRNA that on an agarose gel had two or three separate bands around 1 kb, also after thorough denaturing of the sample before loading. Since single stranded mRNA can easily form secondary structures making it run longer than expected on an agarose gel, the Cas9 mRNA with several shorter bands was not immediately discarded as non-functional. When injecting this mRNA together with gRNAs that needed to be screened by the T7 endonuclease assay, it was impossible to know if the mRNA was functional or not, as the assay never yielded clear results.

Protocol 1 Cas9 mRNA was additionally injected together with gRNAs targeting GFP, as it was believed that this would create a clearly visible knock-down of GFP expression, and thus would give an answer to the question if the Cas9 mRNA was functional or not. As no knock-down could be easily observed in injected Tg(otpb:gfp) x ABTL wildtype embryos, a new protocol for the production of Cas9 mRNA was developed with guidance from Thomas Auer (Protocol 2, methods section 3.7.1).

When using protocol 2 Cas9 mRNA, a single band was obtained on the agarose gel around 2000-2500 kb something which was in accordance with the mRNA produced by Auer (Auer et al., 2013). The mRNA was first injected with gRNAs targeting GFP, still looking for a visible knock-down of GFP expression in the brain of Tg(otpb:gfp) x

ABTL wildtype embryos. Increasing concentrations of both gRNA and Cas9 mRNA was injected, as the literature does not yet agree on optimal injection amounts (Auer et al., 2013; Gagnon et al., 2014). Furthermore, two different plasmids were used for the Cas9 mRNA transcription and the two different eGFP gRNAs were used separately or together in the hope of creating a greater knock-down effect. In all cases, the screen was focused on GFP expression in the brain and no obvious effect could be seen.

The solution to the Cas9 mRNA issue came from Dr. Iain Drummond (Nephrology Division, Massachusetts General Hospital), who advised that the Cas9 mRNA should rather be transcribed at 20°C over night instead of the normal two hours at 37°C. This protocol 2b (methods section 3.7.1) Cas9 mRNA was injected with gRNAs DAT1 and DAT2 and ThEx1, where the cutting efficiency could be screened for by a restriction enzyme digest. In both cases, the digests showed that an in/del mutation had occurred, finally confirming a functional Cas9 mRNA. Possibly, due to the large size of the Cas9 mRNA, it benefits from a slower transcription reaction.

5.3 The Creation of Knock-in Lines

The first strategy for the targeted knock-in of transgenes was based on TALEN technology and non homologous end joining (NHEJ). The TALEN technique was well established in the lab at the moment, and it was repeatedly shown that TALENs can create DSBs inducing in/del mutations at almost any site in the genome (Zu et al., 2013; Xiao et al., 2013; Liu et al., 2014). The literature also shows that the TALENs can be used to linearize a vector or a plasmid upon co-injection into one cell stage zebrafish embryos (Zu et al., 2013). The transgene for insertion was therefore to be presented to the cell in a circular plasmid flanked by TALEN target sites. TALEN target sites at both sides of the transgene opens the possibility for insertion of just the transgene without the large vector backbone which could interfere with the expression of the gene.

Upon injection of the eGFP donor plasmid together with either TALEN pairs Th3 or DAT3, no green green fluorescent cells could be observed. Additionally, PCR of the loci did not indicate that any insertion had taken place. A successful knock-in would be dependent on DSBs created both in the genome and in the donor plasmid in the same cell. When the TALEN cutting efficiency is low, which was the case for both TALEN Th3 and DAT3, the chance for this to happen is reduced.

In an attempt to increase knock-in efficiency, a new strategy was made where the transgene to be inserted was presented to the cell as an already linear single stranded oligonucleotide. This way, the TALENs would only have to create a DSB in the genome and NHEJ would insert the oligonucleotide in the repair process. Several other groups have already reported successful insertions of similar short single stranded oligonucleotides in zebrafish using the CRISPR/Cas9 system (Hruscha et al., 2013; Hwang et al., 2013; Gagnon et al., 2014; Chang et al., 2013). Therefore, a short oligonucleotide containing the 50 nucleotide attP landing site sequence with short *th* homology arms was designed. Upon successful insertion of an attP landing site, other transgenes could later be knocked in at this locus using the PhiC31 integrase, making it a promising strategy (Bischof et al., 2007; Mosimann et al., 2013). Regrettably, no insertion event of the attP oligonucleotide could be detected by PCR meaning that NHEJ had not inserted the oligonucleotide. Possibly, the

amounts of each component injected was not optimal to favour an insertion.

Based on published data it was concluded that for the knock-in to succeed, DSBs should be induced with a higher frequency (Auer et al., 2013). Creating new potentially more efficient TALENs would be quite a time consuming process, and so the CRISPR/Cas9 system, where you can quickly generate new gRNAs to target different sites in the gene sequence, was chosen to further develop a method for the targeted knock-in of transgenes. As a first step in this process, the knock-in experiment based on NHEJ and co-injection of a circular donor plasmid published by Auer et al. (2013) was reproduced successfully, meaning that this was a promising strategy to follow. With the aim of inserting transgenes at the *th* locus, a gRNA targeting the TALEN target sites already present in the Th donor plasmids was designed. For its generation, an alternative cloning free method was used showing that gRNAs can be designed and produced yet more efficiently than earlier presumed as tedious cloning steps can be completely skipped (Gagnon et al., 2014). Although injections of this ThEx1 gRNA showed that its in/del inducing frequency was rather low an attempt was done to use this gRNA to induce the insertion of an eGFP ORF at the *th* locus in ABTL wildtype fish. Regrettably, the efficiency of the ThEx1 gRNA was too low to induce the knock-in.

As a last attempt at increasing DSB efficiency for a targeted knock-in, two different gRNAs were injected together with a donor-plasmid where one gRNA was targeting the genome and one gRNA was targeting the donor plasmid, hopefully making more linearized donor plasmid available for insertion, a strategy which was already proved efficient(Auer et al., 2013). The eGFP-GAL4 donor plasmid was injected with gRNAs ThEx1 and eGFP Paris in ABTL wildtype embryos. Since these embryos have no eGFP locus, the eGFP gRNA would only guide a cut in the donor plasmid and the ThEx1 gRNA would only guide cuts in the genome. This should have increased both the amounts of DBS in the genome and linearized plasmid ready for insertion, however as the results show, not enough to induce an insertion of the GAL4 ORF.

The potential power of an efficient and reliable method for the targeted knock-in of transgenes is great. Efforts are continuously being made on optimization and further development of such a method and coming to the very end of this work, a new publication described an optimized knock-in method in zebrafish using TALENs ensuring germline transmission exceeding 10% (Shin et al., 2014). The strategy was based on homologous recombination rather than NHEJ, and the results showed that the size of the homology arms and the way of linearization of the donor construct both have impact on knock-in frequencies. Interestingly, the donor constructs used were linearized by restriction enzymes before co-injections with TALENs into zebrafish embryos, a method which was earlier proven less efficient (Auer et al., 2013). In the light of this publication however, it would be compelling to create new donor constructs and intent TALEN-mediated knock-in of transgenes based on homologous recombination.

5.4 Dopamine and Neurogenesis

The analysis of the effect of DA on neurogenesis was done by evaluating the number of spinal motor neurons and the number of catecholaminergic neurons in *th* homozygous mutants compared to their wildtype siblings. The *isl1* stain was of such a poor quality

that no conclusions could be made based on this stain regarding the effect of dopamine (DA) on the neurogenesis of spinal motor neurons. To obtain a better *isl1* stain it would be possible to do in situ hybridizations on sections of embryos instead of the whole embryo to better reach the cells and to obtain pictures where the *is1*-stained cells would be clearer.

The *hb9* stain clearly labelled the motor neurons in the spine. Because of time constrains it was not possible to complete a thorough cell count, however, if anything the stain looked stronger in the homozygous *th* mutants as compared to wildtype siblings. This contradicts the findings done by Reimer et al. (2013) where the number of spinal motor neurons were reduced in different DA deficient zebrafish models. However, the DA deficient zebrafish models used in this study were, as before mentioned, not ideal. Morpholinos may elicit non-specific effects and the neurotoxin 6-hydroxydopamine (6-OHDA) used to kill dopaminergic cells could easily have acted off-target as it is shown that 6-OHDA is only partially selective for catecholaminergic neurons (Storch and Collins, 2013; Javoy, 1975). Because no clean genetic procedure was used to eliminate all DA, the reduction in motor neurons observed by Reimer et al. (2013) is not guaranteed to be caused by the reduction of DA alone. Nevertheless, to trust the contradicting results produced in the current work, a larger number of mutants should be analyzed.

Regarding the effect of DA on neurogenesis in the brain, no major effect could be detected on the few *th* mutants and wildtypes evaluated. This is in concordance with the data obtained from the $Th^{-/-}$ mouse model research where no major morphological changes could be detected when looking at AADC positive dopaminergic cells in $Th^{-/-}$ mice (Zhou and Palmiter, 1995; Zhou et al., 1994). However, to draw a confirmatory conclusion, thorough cell counts should be performed and also here a greater number of mutants and wildtype siblings should be evaluated.

Still, as a pure knockout line is a much more reliable and clean animal model than those created by morpholinos or pharmaceuticals, together with the fact that the findings obtained with this mutant model are in agreement with the $Th^{-/-}$ mouse model conclusions, the results obtained in this work could in fact question the importance of DA in neurogenesis as claimed by Reimer et al. (2013).

5.5 Dopamine and Prepulse Inhibition

Earlier prepulse inhibition (PPI) research has shown that different forms of overstimulation of DA receptors is linked to impairments in PPI (Burgess and Granato, 2007; Braff and Geyer, 1990; Zhang et al., 2000). Such an impairment means that startle responses to startle stimuli after a prepulse are not modulated and are thus about the same as the startle responses to just the startle stimulus without the prepulse. To support the hypothesis of DA as an important modulator of PPI, experiments were done to see how PPI is affected if no dopamine is present in the brain at all, assuming that this is the case for the $th^{-/-}$ zebrafish line. Preliminary results show that overall startle responses seem to be decreased in larvae lacking DA as compared to wildtype. This could indicate that DA is of importance also for normal startle responses. This assumption is additionally supported by research of the startle response in rats where the startle response is shown to be enhanced by DA agonists (Meloni and Davis, 2000, 1999).

Calculations of mean % PPI showed that this behaviour seemed to be reduced in $th^{-/-}$ larvae as compared to wildtype, reinforcing the statement that DA is important for the modulation of PPI and that both too much and too little DA can cause impairments in this behaviour. However, due to high biological variabilities between experimental larvae and low N-numbers, statistically significant differences were not found between groups neither regarding % SLC to different stimuli nor regarding % PPI. Thus no final conclusion could be drawn based on these preliminary data. Improvements should be made by increasing the N-numbers for each experiment and to do several experiments over different days confirming the reproducibility of the results. Additionally, work should be done to optimize tracking parameters. Poor image contrast in combination with shadows created by the grid separating the fish made it hard for the tracking algorithm to detect all fish and to correctly track their movements.

6 Conclusions

The TALEN technology proved to be an effective and reliable approach to obtain targeted mutations within the zebrafish genome, however, precautions should be taken upon their design, especially considering spacer length which should be between 15 and 17 bases, and on the composition of RVDs which can interfere with TALEN binding specificity and efficiency. When aiming at knocking out a gene by targeted mutagenesis, it is important that the induced mutation actually disrupts the gene enough for the if translated protein to be non-functional. A documented and assumed DNA-repair bias can make this difficult, if the same in frame mutation is always induced. As a solution to this problem, it has been suggested to co-inject a stop-codon oligonucleotide with stop codons in all three readings frames and that upon insertion will ensure a knockout of the gene.

The CRISPR/Cas9 system resulted more complicated to establish than estimated due to the lack of a good protocol for the transcription of Cas9 mRNA. Finally, it was found beneficial to use a longer and slower transcription reaction at a lower temperature than suggested in the literature. In the search for a good screening method for the CRISPR cutting efficiency, it was concluded that a restriction enzyme digest is by far the most sensitive and conclusive first assessment. With a functioning Cas9 and reliable screening method, the CRISPR/Cas9 system proved effective and very versatile, seeing as gRNAs can be rapidly produced and injected alone or in combination. Experience however, demonstrated that it can be a challenge to find good gRNAs inducing high in/del frequencies. Consequently, time should be invested into searching and screening for efficient gRNAs before starting any larger CRISPR/Cas9 experiments. Fortunately, as time passes, more and more papers are published on the topic and advice for the optimization of the method come with them.

According to the literature, both TALENs and CRISPRs can be used for the targeted knock-in of transgenes. Conversely, neither TALENs nor CRISPRs could be applied to successfully create targeted knock-ins at the *th* or *dat* locus in this work.The fact that CRISPR knock-ins of previously published work could be reproduced however, demonstrate that the CRISPR/Cas9 knock-in system functions in principle, but that fine tunings of the system are necessary. Work must be invested into finding good effective gRNAs with high DSB inducing frequencies, and optimal injection volumes of the different components of the system must be decided to ensure both high knock-in frequencies together with low embryonic mortality rates. Regarding TALENs, a strategy based on homologous recombination could prove to be a solution.

Based on the results obtained in this work, the loss of DA does not reduce the generation of spinal motor neurons, and it does not affect neurogenesis of catecholaminergic neurons in the brain. This is both in agreement and disagreement with previous results on DA and neurogenesis, however more thorough research using the $th^{-/-}$ zebrafish line should be performed before drawing any confirmatory conclusions.

Preliminary results from prepulse inhibition (PPI) experiments indicated that DA is needed both for a normal SLC startle response and for the modulation of PPI, however more experiments must be completed before a final conclusion can be made.

References

Allodi, I. and Hedlund, E. Directed midbrain and spinal cord neurogenesis from pluripotent stem cells to model development and disease in a dish. *Frontiers in Neuroscience*, 8(109) (2014).

Alvarez-Buylla, A. and Garcia-Verdugo, J. M. Neurogenesis in Adult Subventricular Zone. *The Journal of Neuroscience*, 22(3):629–634 (2002).

Andersson, E., Tryggvason, U., Deng, Q., Friling, S., Alekseenko, Z., Robert, B., Perlmann, T. and Ericson, J. Identification of intrinsic determinants of midbrain dopamine neurons. *Cell*, 124:393–405 (2006).

Ang, S. L. Transcriptional control of midbrain dopaminergic neuron development. *Development*, 133:3499–3506 (2006).

Arber, S., Han, B., Mendelsohn, M., Smith, M., Jessell, T. and Sockanathan, S. Requirement for the Homeobox Gene Hb9 in the Consolidation of Motor Neuron Identity. *Neuron*, 23:659–674 (1999).

Auer, T., Duroure, K., De Cian, A., Concordet, J. and Del Bene, F. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Research*, 24:142–153 (2013).

Baker, S., Baker, K. and Hagg, T. Dopaminergic nigrostriatal projections regulate neural precursor proliferation in the adult mouse subventricular zone. *European Journal of Neuroscience*, 20:575–579 (2004).

Bartlett, J. and Stirling, D. *A Short History of the Polymerase Chain Reaction*, volume 226 of *Methods in Molecular Biology*. Humana Press (2003).

Bedell, V. M., Wang, Y., Campbell, J. M., Poshusta, T. L., Starker, C. G., Krug, R. G., Tan, W., Penheiter, S. G., Ma, A. C., Leung, A. Y. H., Fahrenkrug, S. C., Carlson, D. F., Voytas, D. F., Clark, K. J., Essner, J. J. and Ekker, S. C. In vivo genome editing using a high-efficiency TALEN system. *Nature*, 491(7422):114–118 (2012).

Bedell, V. M., Westcot, S. E. and Ekker, S. C. Lessons from morpholino-based screening in zebrafish. *Briefings in Functional Genomics Volume*, 10(4):181–188 (2011).

Benoit-Marand, M., Jaber, M. and Gonon, F. Release and elimination of dopamine in vivo in mice lacking the dopamine transporter: functional consequences. *European Journal of Neuroscience*, 12:2985–2992 (2000).

Bischof, J., Maeda, R., Hediger, M., Karch, F. and Basler, K. An optimized transgenesis system for Drosophila using germ-line-specific PhiC31 integrases. *PNAS*, 104:3312–3317 (2007).

Bjorklund, L. M., Sanchez-Pernaute, R., Chung, S., Andersson, T., Chen, I. Y. C., McNaught, K. S. P., Brownell, A.-L., Jenkins, B. G., Wahlestedt, C., Kim, K.-S., and Isacson, O. Embryonic stem cells develop into functional dopaminergic neurons after transplantation in a Parkinson rat model. *PNAS*, (4):2344–2349 (2002).

Blinkov, S. M. and Glezer, I. I. *The Human Brain in Figures and Tables. A Quantitative Handbook*. Plenum Press (1968).

Braff, D. and Geyer, M. Sensorimotor gating and schizophrenia. Human and animal model studies. *Archives of General Psychiatry*, 47:181–188 (1990).

Braff, D., Grillon, C. and Geyer, M. Gating and habituation of the startle reflex in schizophrenic patients. *Archives of General Psychiatry*, 49:206–215 (1992).

Brand, A. H. and Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*, 118:401–415 (1993).

Burgess, H. and Granato, M. Sensorimotor gating in larval zebrafish. *The Journal of Neuroscience*, 27:4984–4994 (2007).

Cermak, T., Doyle, E., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J., Somia, N., Bogdanove, A. and Voytas, D. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research*, 39:7879 (2011).

Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J. and Xi, J. Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Research*, 23:465–472 (2013).

Cong, L., Zhou, R., Kuo, Y., Cunniff, M. and Zhang, F. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nature Communications*, 3:968 (2012).

Devoto, P. and Flore, G. On the origin of cortical dopamine: is it a co-transmitter in noradrenergic neurons? *Current Neuropharmacology*, 6:115–125 (2006).

Devoto, P., Flore, G., Pani, L. and Gessa, G. Evidence for co-release of noradrenaline and dopamine from noradrenergic neurons in the cerebral cortex. *Molecular Psychiatry*, 6:657–664 (2001).

Diaz, J., Ridray, S., Mignon, V., Griffon, N., Schwartz, J.-C. and Sokoloff, P. Selective Expression of Dopamine D3 Receptor mRNA in Proliferative Zones during Embryonic Development of the Rat Brain. *The Journal of Neuroscience*, 17(11):4282–4292 (1997).

Doyle, E., Booher, N., Standage, D., Voytas, D., Brendel, V., VanDyk, J. and Bogdanove, A. TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Research* (2012).

Engler, C., Gruetzner, R., Kandzia, R. and Marillonnet, S. Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. *PLoS ONE*, 4(5):e5553 (2009).

Eriksson, P. S., Perfilieva, E., Bjork-Eriksson, T., Alborn, A.-M., Nordborg, C., Peterson, D. A. and Gage, F. H. Neurogenesis in the adult human hippocampus. *Nature Medicine*, 4:1313–1317 (1998).

Ernst, A., Alkass, K., Bernard, S., Salehpour, M., Perl, S., Tisdale, J., Possnert, G., Druid, H. and Frisen, J. Neurogenesis in the Striatum of the Adult Human Brain. *Cell*, 156(5):1072–1083 (2014).

Feng, G., Mellor, R. H., Bernstein, M., Keller-Peck, C., Nguyen, Q. T., Wallace, M., Nerbonne, J. M., Lichtman, J. W. and Sanes, J. R. Imaging Neuronal Subsets in Transgenic Mice Expressing Multiple Spectral Variants of GFP. *Neuron*, 28(1):41–51 (2000).

Fernandes, A. M., Fero, K., Arrenberg, A., Bergeron, S., Driever, W. and Burgess, H. Deep Brain Photoreceptors Control Light-Seeking Behavior in Zebrafish Larvae. *Current Biology*, 22(21):2042–2047 (2012).

Fernandes, M. *Genetic analysis of development and behavioral roles of neurons specified by the Orthopedia transcription factor*. Ph.D. thesis, Albert-Ludwigs-Universitat, Freiburg im Breisgau (2013).

Flames, N. and Hobert, O. Transcriptional Control of the Terminal Fate of Monoaminergic Neurons. *Annual Review of Neuroscience*, 34(1):153–184 (2011).

Gagnon, J. A., Valen, E., Thyme, S. B., Huang, P., Ahkmetova, L., Pauli, A., Montague, T. G., Zimmerman, S., Richter, C. and Schier, A. F. Efficient Mutagenesis by Cas9 Protein-Mediated Oligonucleotide Insertion and Large-Scale Assessment of Single-Guide RNAs. *PLoS ONE*, 9(5):e98186 (2014).

Giros, B., Jaber, M., Jones, S., Wightman, R. and Caron, M. Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter. *Nature*, 379:606–612 (1996).

Groth, A., Olivares, E., Thyagarajan, B. and Calos, M. A phage integrase directs efficient site-specific integration in human cells. *PNAS*, 97:5995–6000 (2000).

Heidenreich, E., Novotny, R., Kneidinger, B., Holzmann, V. and Wintersberger, U. Non-homologous end joining as an important mutagenic process in cell cycle-arrested cells. *EMBO*, 22:2274–2283 (2003).

Hoffman, H. and Ison, J. Reflex modification in the domain of startle: Some empirical findings and their implications for how the nervous system processes sensory input. *Psychological Review*, 87:175–189 (1980).

Hoglinger, G., Arias-Carri, O., Ipach, B. and Oertel, W. Origin of the dopaminergic innervation of adult neurogenic areas. *Journal of Comparative Neurology*, 522:2336–2348 (2014).

Hoglinger, G., Rizk, P., Muriel, M., Duyckaerts, C., Oertel, W., Caille, I. and Hirsch, E. Dopamine depletion impairs precursor cell proliferation in Parkinson disease. *Nature Neuroscience*, 7:726–735 (2004).

Holzschuh, J., Barrallo-Gimeno, A., Ettl, A.-K., D?rr, K., Knapik, E. W. and Driever, W. Noradrenergic neurons in the zebrafish hindbrain are induced by retinoic acid and require tfap2a for expression of the neurotransmitter phenotype. *Development*, 130:5741–5754 (2003).

Holzschuh, J., Ryu, S., Aberger, F. and Driever, W. Dopamine transporter expression distinguishes dopaminergic neurons from other catecholaminergic neurons in the developing zebrafish embryo. *Mechanisms of Development*, 101:237–243 (2001).

Horvath, P. and Barrangou, R. CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science*, 327:575–579 (2010).

Hruscha, A., Krawitz, P., Rechenberg, A., Heinrich, V., Hecht, J., Haass, C. and Schmid, B. Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development*, 24:4982–4987 (2013).

Hwang, W., Fu, Y., Reyon, D., Maeder, M., Tsai, S., Sander, J., Peterson, R., Yeh, J. and Joung, K. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature Biotechnology*, 31:227–229 (2013).

Inoue, A., Takahashi, M., Hatta, K., Hotta, Y. and Okamoto, H. Developmental regulation of Islet-1 mRNA expression during neuronal differentiation in embryonic zebrafish. *Developmental Dynamics*, 199(1):1–11 (1994).

Jao, L.-E., Wente, S. R. and Chen, W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *PNAS*, 110(34):13904–13909 (2013).

Javoy, F. Partial selectivity of 6-OHDA induced neuronal degeneration after intraissular injection in the brain with special reference to the nigro-striatal dopamine system. *The Journal of Neural Transmission*, 37(3):219–227 (1975).

Joung, K. and Sander, J. TALENs: a widely applicable technology for targeted genome editing. *Nature Reviews Molecular Cell Biology*, 14:49–55 (2013).

Kaplan, M. S. and Hinds, J. W. Neurogenesis in the adult rat: electron microscopic analysis of light radioautographs. *Science*, 197:1092–1094 (1977).

Kele, J., Simplicio, N., Ferri, A., Mira, H., Guillemot, F., Arenas, E. and Ang, S. Neurogenin 2 is required for the development of ventral midbrain dopaminergic neurons. *Development*, 133:495–505 (2006).

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. and Schilling, T. F. Stages of embryonic development of the zebrafish. *Developmental Dynamics*, 203(3):253–310 (1995).

Kizil, C., Kaslin, J., Kroehne, V. and Brand, M. Adult neurogenesis and brain regeneration in zebrafish. *Developmental Neurobiology*, 72(3):429–461 (2012).

Kobayashi, K., Morita, S., Sawada, H., Mizuguchi, T., Yamada, K., Nagatsu, I., Hata, T., Watanabe, Y., Fujita, K. and Nagatsu, T. Targeted Disruption of the Tyrosine Hydroxylase Locus Results in Severe Catecholamine Depletion and Perinatal Lethality in Mice. *The Journal of Biological Chemistry*, 270:27235–27243 (1995).

Kumar, C. M., Sathisha, U., Dharmesh, S., Rao, A. A. and Singh, S. A. Interaction of sesamol (3,4-methylenedioxyphenol) with tyrosinase and its effect on melanin synthesis. *Biochimie*, 93(3):562 – 569 (2011).

Kwan, K. M., Fujimoto, E., Grabher, C., Mangum, B. D., Hardy, M. E., Campbell, D. S., Parant, J. M., Yost, H. J., Kanki, J. P. and Chien, C.-B. The Tol2kit: A multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Developmental Dynamics*, 236(11):3088–3099 (2007).

Leo, D. and Gainetdinov, R. Transgenic mouse models for ADHD. *Cell and Tissue Research*, 354:259–271 (2013).

Liu, Y., Luo, D., Lei, Y., Hu, W., Zhao, H. and Cheng, C. A highly effective TALEN-mediated approach for targeted gene disruption in Xenopus tropicalis and zebrafish. *Methods*, 69:58–66 (2014).

Madras, B. K., Miller, G. M. and Fischman, A. J. The Dopamine Transporter and Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry*, 57(11):1397–1409 (2005).

Mahler, A., Filippi, A. and Driever, W. DeltaA/DeltaD Regulate Multiple and Temporally Distinct Phases of Notch Signaling during Dopaminergic Neurogenesis in Zebrafish. *The Journal of Neuroscience*, 30:16621–16635 (2010).

Mali, P., Esvelt, K. and Church, G. Cas9 as a versatile tool for engineering biology. *Nature Methods*, 10:957–963 (2013).

Mao, Z., Bozzella, M., Seluanov, A. and Gorbunova, V. Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair*, 7(10):1765–1771 (2008).

Meloni, E. G. and Davis, M. Enhancement of the acoustic startle response in rats by the dopamine D1 receptor agonist SKF 82958. *Psychopharmacology*, 144(4):373–380 (1999).

Meloni, E. G. and Davis, M. Enhancement of the acoustic startle response by dopamine agonists after 6-hydroxydopamine lesions of the substantia nigra pars compacta: corresponding changes in c-Fos expression in the caudate-putamen. *Brain Research*, 879(1-2):93 – 104 (2000).

Miller, J., Tan, S., Qiao, G., Barlow, K., Wang, J., Xia, D., Meng, X., Paschon, D., Leung, E., Hinkley, S., Dulay, G., Hua, K., Ankoudinova, I., Cost, G., Urnov, F., Zhang, H., Holmes, M., Zhang, L., Gregory, P. and Rebar, E. A TALE nuclease architecture for efficient genome editing. *Nature Biotechnology*, 29:143–148 (2010).

Montague, T. G., Cruz, J. M., Gagnon, J. A., Church, G. M. and Valen, E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Research*, 42:W401–W407 (2014).

Mosimann, C., Puller, A. C., Lawson, K. L., Tschopp, P., Amsterdam, A. and Zon, L. I. Site-directed zebrafish transgenesis into single landing sites with the PhiC31 integrase system. *Developmental Dynamics*, 242(8):949–963 (2013).

Nilsson, M., Perfilieva, E., Johansson, U., Orwar, O. and Eriksson, P. Enriched environment increases neurogenesis in the adult rat dentate gyrus and improves spatial memory. *The Journal of Neurobiology*, 39:569–578 (1999).

Puelles, E., Annino, A., Tuorto, F., Usiello, A., Acampora, D., Czerny, T., Brodski, C., Ang, S., Wurst, W. and Simeone, A. Otx2 regulates the extent, identity and fate of neuronal progenitor midbrain. *Development*, 131:2037–2048 (2004).

Reimer, M., Norris, A., Ohnmacht, J., Patani, R., Zhong, Z., Dias, T., Kuscha, V., Scott, A., Chen, Y., Rozov, S., Frazer, S., Wyatt, C., Higashijima, S., Patton, E., Panula, P., Chandran, S., Becker, T. and Becker, C. Dopamine from the brain promotes spinal motor neuron generation during development and adult regeneration. *Developmental Cell*, 25:478–491 (2013).

Reyon, D., Tsai, S., Khayter, C., Foden, J., Sander, J. and Joung, K. FLASH assembly of TALENs for high-throughput genome editing. *Nature Biotechnology*, 30:460–465 (2012).

Ryu, S., Mahler, J., Acampora, D., Holzschuh, J., Erhardt, S., Omodei, D., Simeone, A. and Driever, W. Orthopedia Homeodomain Protein Is Essential for Diencephalic Dopaminergic Neuron Development. *Current Biology*, 17(10):873 – 880 (2007).

Sambrook, J. and Russel, D. W. *Molecular Cloning: A Laboratory Manual, 3rd Ed.* Cold Spring Harbor Laboratory Press (1968).

Sander, J. D., Zaback, P., Joung, J. K., Voytas, D. F. and Dobbs, D. Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool. *Nucleic Acids Research*, 35:W599–605 (2007).

Schmid-Burgk, J. L., Schmidt, T., Kaiser, V., Honing, K. and Hornung, V. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. *Nature Biotechnology*, 31(1):76–81 (2013).

Seeman, P. Dopamine and schizophrenia. *Scholarpedia*, 2(10):3634 (2007). (revision No 91208).

Shin, J., Chen, J. and Solnica-Krezel, L. TALEN-mediated precise gene modification by homologous recombination in zebrafish. *Development*, 14(19):3807–3818 (2014).

Shulman, J. M., De Jager, P. L. and Feany, M. B. Parkinson's Disease: Genetics and Pathogenesis. *Annual Review of Pathology: Mechanisms of Disease*, 6(1):193–222 (2011).

Simon, H. H., Saueressig, H., Wurst, W., Goulding, M. and Oeary, D. D. Fate of midbrain dopaminergic neurons controlled by the engrailed genes. *The Journal of Neuroscience*, 21:3126–34 (2001).

Storch, A. and Collins, M. A. *Neurotoxic Factors in Parkinson's Disease and Related Disorders*. Springer (2013).

Streubel, J., Blucher, C., Landgraf, A. and Boch, J. TAL effector RVD specificities and efficiencies. *Nature Biotechnology*, 30(7):593–595 (2012).

Swistowski, A., Peng, J., Liu, Q., Mali, P., Rao, M. S., Cheng, L. and Zeng, X. Efficient Generation of Functional Dopaminergic Neurons from Human Induced Pluripotent Stem Cells Under Defined Conditions. *STEM CELLS*, 28(10):1893–1904 (2010).

Takamura, N., Nakagawa, S., Masuda, T., Boku, S., Kato, A., Song, N., An, Y., Kitaichi, Y., Inoue, T. and Koyama, I. The effect of dopamine on adult hippocampal neurogenesis. *Progress in Neuropsychopharmacology and Biological Psychiatry*, 50:116–124 (2014).

Tay, T. L., Ronneberger, O., Ryu, S., Nitschke, R. and Driever, W. Comprehensive catecholaminergic projectome analysis reveals single-neuron integration of zebrafish ascending and descending dopaminergic systems. *Nature Communications*, 2:171 (2011).

Todd, R. Neural development is regulated by classical neurotransmitters: dopamine D2 receptor stimulation enhances neurite outgrowth. *Biological Psychiatry*, 31:794–807 (1992).

Volkow, N., Wang, G., Newcorn, J., Kollins, S., Wigal, T., Telang, F., Fowler, J., Goldstein, R., Klein, N., Logan, J., Wong, C. and Swanson, J. Motivation deficit in ADHD is associated with dysfunction of the dopamine reward pathway. *Molecular Psychiatry*, 11:1147–54 (2011).

Wah, D., Bitinaite, J., Schildkraut, I. and Aggarwal, A. Structure of FokI has implications for DNA cleavage. *PNAS*, 95:10564–10569 (1998).

Wang, M., Jin, P., Bumcrot, D., Marigo, V., McMahon, A., Wang, E., Woolf, T. and Pang, K. Induction of dopaminergic neuron phenotype in the midbrain by Sonic hedgehog protein. *Nature Medicine*, 1:1184–1188 (1995).

Weber, E., Gruetzner, R., Werner, S., Engler, C. and Marillonnet, S. Assembly of Designer TAL Effectors by Golden Gate Cloning. *PLoS ONE*, 6(5):e19722 (2011).

Wolpert, L. and Tickle, C. *Principles of Development*. Oxford University Press (2010).

Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z., Yang, Z., Zu, Y., Li, W., Huang, P., Tong, X., Zhu, Z., Lin, S. and Zhang, B. Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Research*, 41:e141 (2013).

Zhang, J., Forkstam, C., Engel, J. and Svensson, L. Role of dopamine in prepulse inhibition of acoustic startle. *Psychopharmacology*, 149:181–188 (2000).

Zhou, Q. and Palmiter, R. Dopamine-deficient mice are severely hypoactive, adipsic, and aphagic. *Cell*, 83:1197–1209 (1995).

Zhou, Q., Quaife, C. and Palmiter, R. Targeted disruption of the tyrosine hydroxylase gene reveals that catecholamines are required for mouse fetal development. *Nature*, 374:640–643 (1994).

Zu, Y., Tong, X., Wang, Z., Liu, D., Pan, R., Li, Z., Hu, Y., Luo, Z., Huang, P., Wu, Q., Zhu, Z., Zhang, B. and Lin, S. TALEN-mediated precise gene modification by homologous recombination in zebrafish. *Nature Methods*, 10:329–331 (2013).

A Appendix

A.1 List of primers

Name	Sequence	Tm
Th_TALfwd	AAGATGTTCAATGACGGAG	54
Th_TALrev	GTTGAAGTGGACAATGTAG	54
DAT_TALfwd	ACTGGCCACAAAATCTAAAC	56
DAT_TALrev	TCTTCTTGCCCCAGGTCTCA	58
pCR8_F1	TTGATGCCTGGCAGTTCCCT	60
pCR8_R1	CGAACCGAACAGGCTTATGT	56
TAL_F1	TTGGCGTCGGCAAACAGTGG	62
TAL_R2	GGCGACGAGGTGGTCGTTGG	64
SeqTALEN_5-1	CATCGCGCAATGCACTGAC	58
Th1fwd_KpnI	AGCTGGTACCTAGTTCCATCTGTTCCGCG	55
Th1rev_ClaI	AGCTATCGATGCTCTTTGTGGAAGATGTGG	
Th2fwd_SpeI	AGCTACTAGTTAGTTCCATCTGTTCCGCG	55
Th2rev_SacI	AGCTGAGCTCGCTCTTTGTGGAAGATGTGG	55
DAT1fwd_KpnI	AGCTGGTACCTGTGTGTCCATCAGAGCATCTCC	60
DAT1rev_ClaI	AGCTATCGATGCGTGTGTGTGTGACCGC	60
DAT2fwd_SpeI	AGCTACTAGTTGTGTGTCCATCAGAGCATCTCC	60
DAT2rev_SacI	AGCTGAGCTCGCGTGTGTGTGTGACCGC	60
eGFPfwd_HindIII	AGCTAAGCTTCTGGACCATGGTGAGCAAGG	59
eGFPrev_BamHI	AGCTGGATCCAAGCTGGGTCTCAGGAGAGC	60
Gal4fwd_HindIII	AGCTAAGCTTGATCCGATGAAGCTACTGTC	53
Gal4rev_BamHI	AGCTGGATCCACTAGTAGATCCGCTACATATC	52
attPfwd_HindIII	AGCTAAGCTTGTACTGACGGACACACCGAAGC	61
attPrev_BamHI	AGCTGGATCCCGCGACTGACGGTCGTAAGC	60
LoxPfwd_BamHI	GATCCATAACTTCGTATAGGCTACCGTATACGAAGTTATA	
LoxPrev_SpeI	CTAGTATAACTTCGTATACGGTAGCCTATACGAAGTTATG	
SeqA	ACGGCCAGTCTTAAGCTCG	58
CRISPR_thAGfwd	ACGACAGGTTTACATAGACTGC	56
CRISPR_thAGrev	GGGTCTAATAATTCAGGGAGGC	56
ThEx2_fwd	ATTTTCGCACAGCGTTCAGC	58
ThEx2_rev	GGGGTCTAATAATTCAGGGAGGC	58
eGFP-DATins5'_fwd	TCCCAGACACTAGATGGCG	57
eGFP-DATins5'_rev	GGTGCTCAGGTAGTGGTTGT	58
eGFP-DATins3'_fwd	GAAGTTCGAGGGCGACACC	59
eGFP-DATins3'_rev	AACTGAACCCCATTCTGCTCC	58
M13_fwd	TGTAAAACGACGGCCAGT	55
DR274_rev	GAGGCATAAATTCCGTCAGC	55
ThEx2_gRNAtest	TTCTAGCTCTAAAACCGGCC	55
DAT1_gRNAtest	TCTAGCTCTAAAACTCAGCGG	55
DAT2_gRNAtest	CTCTAAAACTCACCAGCACC	54
Th_attPrev	AACTGAGAGAACTCAAAGG	50
attP_plasFwd	TGAAGATTGTACGAGAAACC	51
Gal4_rev	GAAGACTCTTCGCTTGAGG	53

Table 15: List of primers used in this work.

A.2 Gene Sequence Data

A.2.1 Th

```
>Th TALEN / ThEx1 gRNA target site sequence [\textit{th} Exon 1]
AAGATGTTCAATGACGGAGAAGGAGGGGGTCGGTTTGACGTCAGTCAAGAGAAACGATCT
CTATAAATAAGGGGGAAATAAGGCTGCATGAAAGTAAGTGTCCAGCGCTCCCAAGAGACC
ACAATATAGTTCCATCTGTTCCGCGCGCGCGCCATCTGAACATGCCGAATTCAAGCAGCT
CCACATCTTCCACAAAGAGCATCCGGAGAGCTGCGTCTGAACTCGAGCGCTCAGATTCTA
TTACGGTAGGAATTTTTGACTACATTGTCCACTTCAAC

>ThEx2 gRNA target site [\textit{th} Exon 2]
TCTCAGAAGTTTGTTGGGAGGCGGCAGAGTTTGATCGAGGATGCGCGTAAGGAGCGCGAG
GCGGCGGCCGCGGCGGCGGAGGCTGCAGGACTGAGCGAGCAGATCGTGTTTGAGGAGAGC
GATGGAAAAGCATTGATCAGCCTCTTCTTCACTCTCAGGTGCTCTAAAAGCCCTGCGCTG
TCCCGCACCTTGAAGGTTTTCGAG
```

A.2.2 DAT

```
>DAT TALEN / DAT1 and DAT2 gRNA target site [\textit{dat} Exon 1]
ACTGGCCACAAAATCTAAACATTAGACACCTAACAGTGTGTTTTGAGCTGTAAAGGATTA
TAAACTCTTTAATTGAATGCAAAGAAACAGAAACAGCTGATGGAGTATACACTAACCTGC
ACATTTCCCAGACACTAGATGGCGCCAAACTAAACTAATTTTAATTCTATTTTTGTGTGT
GTGTCCATCAGAGCATCTCCTCCTCAGTCTCCTGTGCATGCCCATGCTGAGAGGCAGACC
GGCGGTCACACACACACGCACACGCACACACACACACACATGTCCTCCGTCAGCGGCTCCAG
CAGCGCCGCCGGGCCCCGAGAGGTGGAGCTGGTGCTGGTGAAGGAGCAGAATGGGGTTCA
GTTCACCTCCTCCAGCCTGCGGAGCCCCGGTGCGCACTCACACACACACACCCACTCTCA
CACACACCCGTCCGGACAGCAGCGTGAGACCTGGGGCAAGAAGA
```

A.3 Plasmid Maps & Sequences

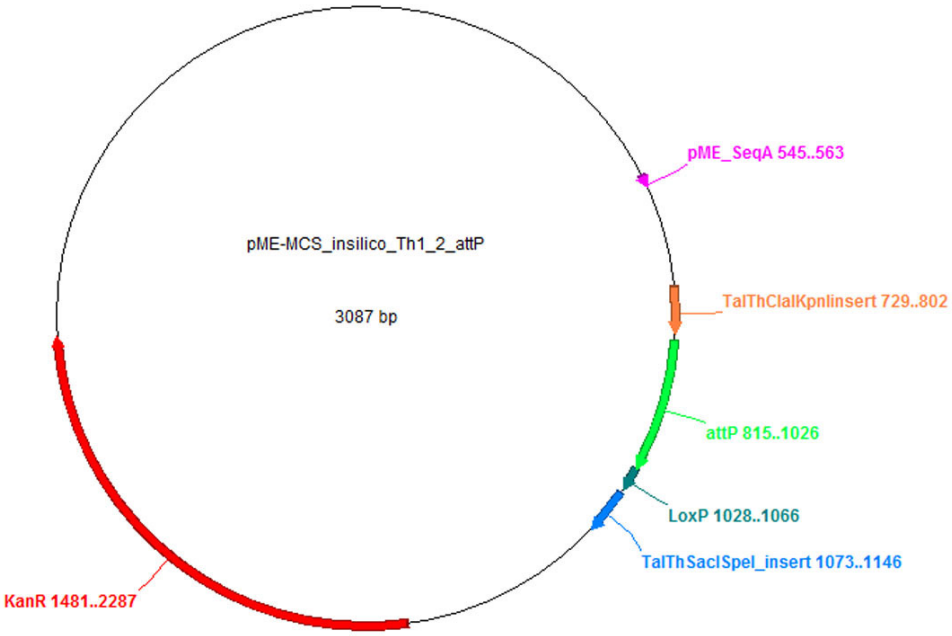

Figure A.1: Plasmid map of donor plasmid Th-attP.

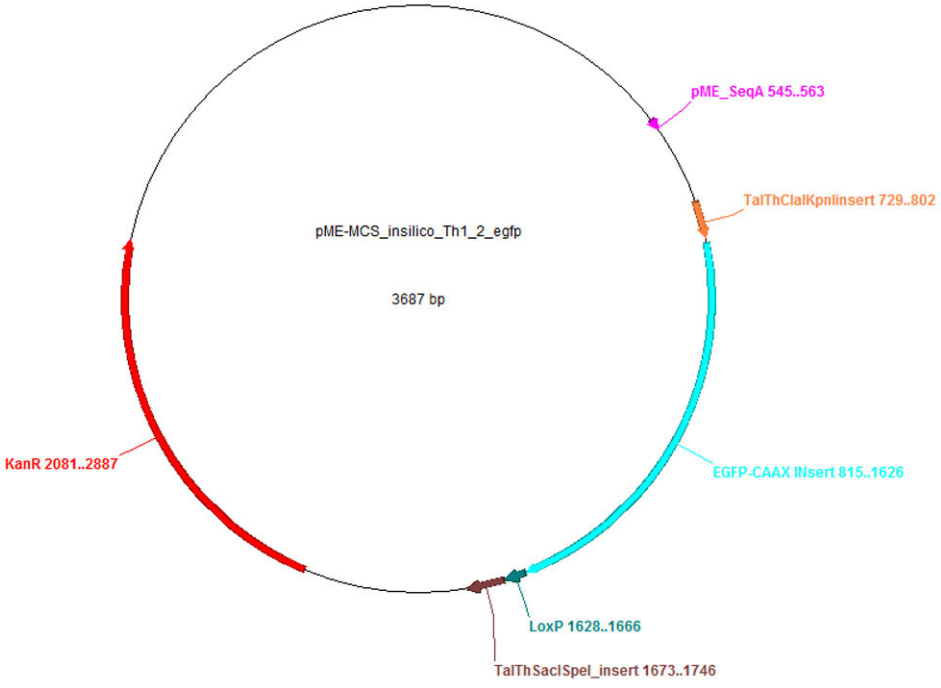

Figure A.2: Plasmid map of donor plasmid Th-eGFP.

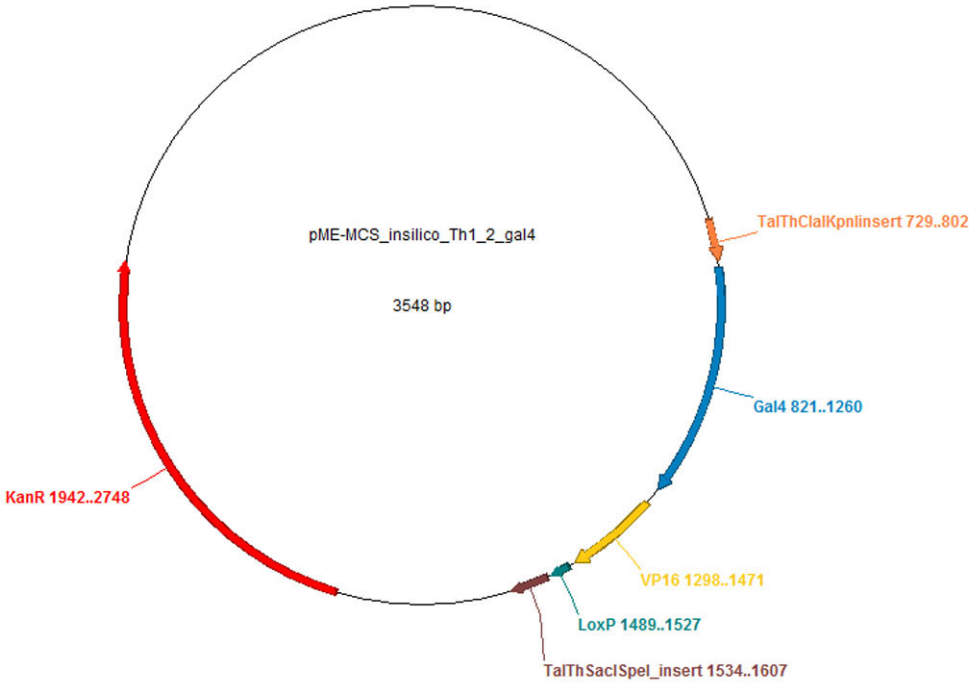

Figure A.3: Plasmid map of donor plasmid Th-GAL4.

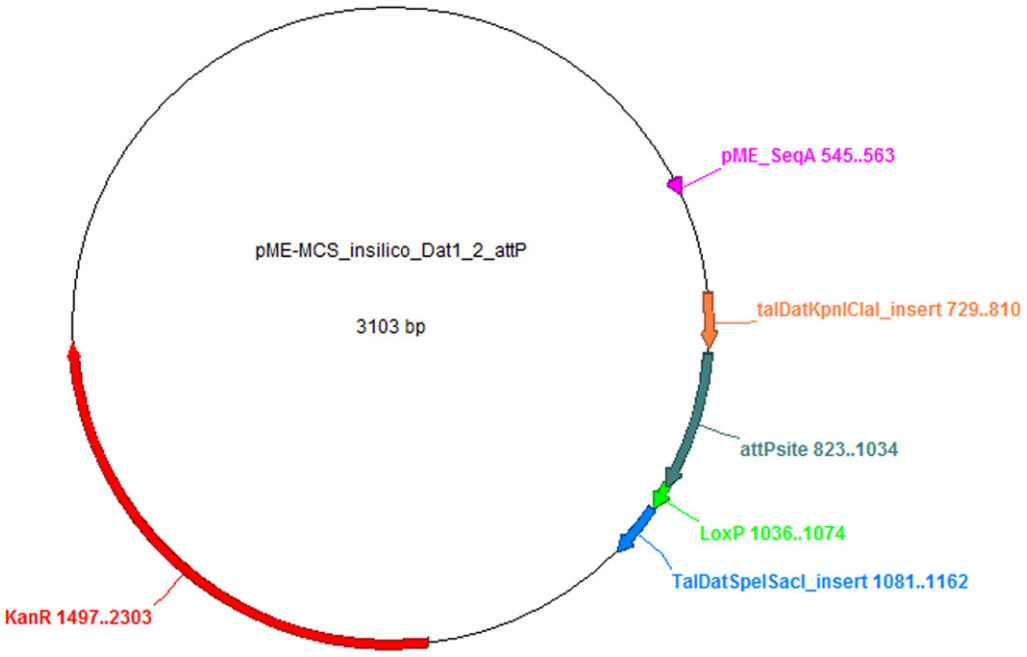

Figure A.4: Plasmid map of donor plasmid DAT-attP.

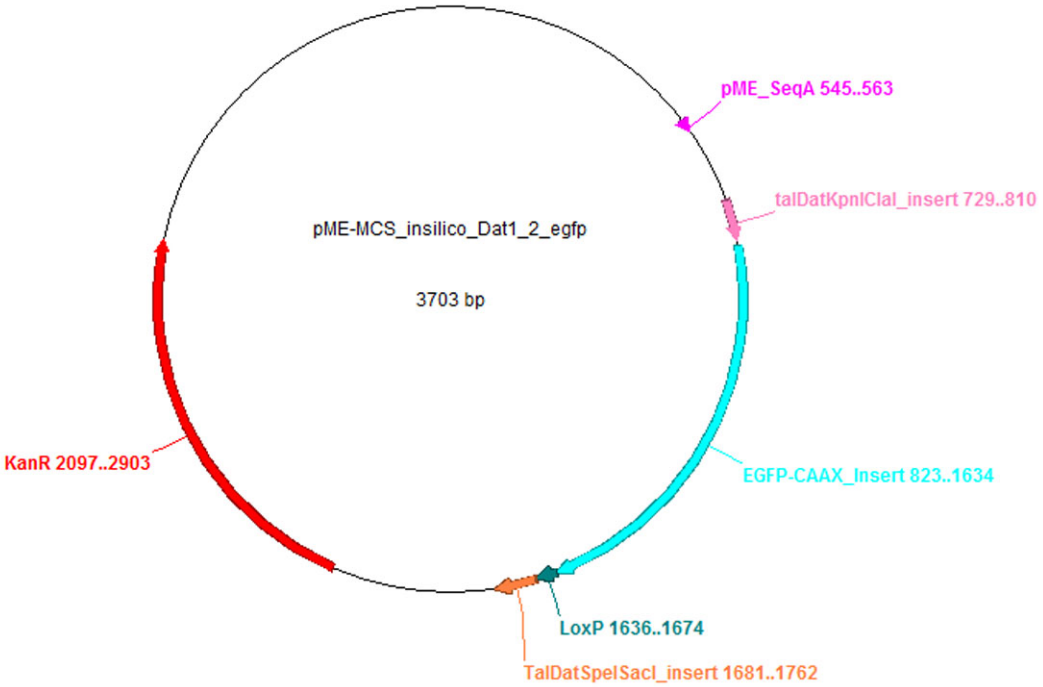

Figure A.5: Plasmid map of donor plasmid DAT-eGFP.

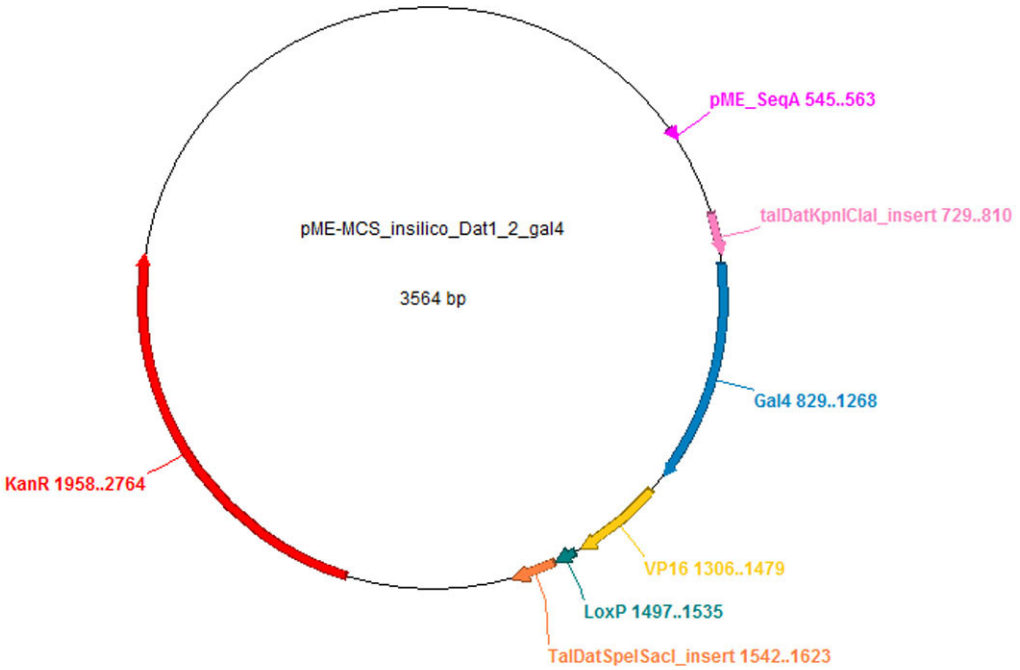

Figure A.6: Plasmid map of donor plasmid DAT-GAL4.

>DAT-attP donor plasmid sequence
CTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA
TACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGA
GCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCA
CGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATACGCGTACCGC
TAGCCAGGAAGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTA
GTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGGGCCGTTGCTTC
ACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAA
CAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGCCTG
GCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAA
AACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGAC
CTGTTCGTTGCAACACATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAA
AGCAGGCTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAAT
TGGGTACCTGTGTGTCCATCAGAGCATCTCCTCCTCAGTCTCCTGTGCATGCCCATGCTG
AGAGGCAGACCGGCGGTCACACACACACGCATCGATAAGCTTGTACTGACGGACACACCG
AAGCCCCGGCGGCAACCCTCAGCGGATGCCCCGGGGCTTCACGTTTTCCCAGGTCAGAAG
CGGTTTTCGGGAGTAGTGCCCCAACTGGGGTAACCTTTGAGTTCTCTCAGTTGGGGGCGT
AGGGTCGCCGACATGACACAAGGGGTTGTGACCGGGGTGGACACGTACGCGGGTGCTTAC
GACCGTCAGTCGCGGGATCCATAACTTCGTATAGGCTACCGTATACGAAGTTATACTAGT
TGTGTGTCCATCAGAGCATCTCCTCCTCAGTCTCCTGTGCATGCCCATGCTGAGAGGCAG
ACCGGCGGTCACACACACACGCGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTG
CGCGCTTGGCGTAATCATGGTCATAGCACCCAGCTTTCTTGTACAAAGTTGGCATTATAA
GAAAGCATTGCTTATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCA
TTATTTGCCATCCAGCTGATATCCCCTATAGTGAGTCGTATTACATGGTCATAGCTGTTT
CCTGGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAATAA
TATCATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGA
GCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTAT
ATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGCTTGT
ATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATG
ATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCA
TCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGAA
AAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGC
TGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCG
ATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA
GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATA
AACTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACC
TTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAG
ACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTAC
AGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTC
ATTTGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAATTGGTTGTAACACTGGCAGA
GCATTACGCTGACTTGACGGGACGGCGCAAGCTCATGACCAAAATCCCTTAACGTGAGTT
ACGCGTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATC
CTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGG
TTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAG
CGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACT
CTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTG
GCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGC
GGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCG
AACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGG
CGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAG
GGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTC
GATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCT
TTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTT

>DAT-eGFP donor plasmid sequence
CTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA
TACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGA
GCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCA
CGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATACGCGTACCGC
TAGCCAGGAAGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTA
GTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGGGCCGTTGCTTC

```
ACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAA
CAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGCCTG
GCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAA
AACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGAC
CTGTTCGTTGCAACACATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAA
AGCAGGCTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAAT
TGGGTACCTGTGTGTCCATCAGAGCATCTCCTCCTCAGTCTCCTGTGCATGCCCATGCTG
AGAGGCAGACCGGCGGTCACACACACACGCATCGATAAGCTTCTGGACCATGGTGAGCAA
GGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAA
CGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGAC
CCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCAC
CCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTT
CTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGA
CGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCAT
CGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTA
CAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGT
GAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCA
GCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCAC
CCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTT
CGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGAGGAGGAAGATC
TAAGCTGAACCCTCCTGATGAGAGTGGCCCCGGCTGCATGAGCTGCAAGTGTGTGCTCTC
CTGAGACCCAGCTTGGATCCATAACTTCGTATAGGCTACCGTATACGAAGTTATACTAGT
TGTGTGTCCATCAGAGCATCTCCTCCTCAGTCTCCTGTGCATGCCCATGCTGAGAGGCAG
ACCGGCGGTCACACACACACGCGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTG
CGCGCTTGGCGTAATCATGGTCATAGCACCCAGCTTTCTTGTACAAAGTTGGCATTATAA
GAAAGCATTGCTTATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCA
TTATTTGCCATCCAGCTGATATCCCCTATAGTGAGTCGTATTACATGGTCATAGCTGTTT
CCTGGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAATAA
TATCATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGA
GCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTAT
ATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGCTTGT
ATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATG
ATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCA
TCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGAA
AAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGC
TGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCG
ATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA
GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATA
AACTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACC
TTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAG
ACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTAC
AGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTC
ATTTGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAATTGGTTGTAACACTGGCAGA
GCATTACGCTGACTTGACGGGACGGCGCAAGCTCATGACCAAAATCCCTTAACGTGAGTT
ACGCGTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATC
CTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGG
TTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAG
CGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACT
CTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTG
GCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGC
GGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCG
AACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGG
CGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAG
GGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTC
GATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCT
TTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTT
```

```
>DAT-GAL4 donor plasmid sequence
CTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA
TACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGA
GCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCA
```

CGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATACGCGTACCGC
TAGCCAGGAAGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTA
GTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGGGCCGTTGCTTC
ACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAA
CAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGCCTG
GCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAA
AACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGAC
CTGTTCGTTGCAACACATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAA
AGCAGGCTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAAT
TGGGTACCTGTGTGTCCATCAGAGCATCTCCTCCTCAGTCTCCTGTGCATGCCCATGCTG
AGAGGCAGACCGGCGGTCACACACACACGCATCGATAAGCTTGATCCGATGAAGCTACTG
TCTTCTATCGAACAAGCATGCGATATTTGCCGACTTAAAAAGCTCAAGTGCTCCAAAGAA
AAACCGAAGTGCGCCAAGTGTCTGAAGAACAACTGGGAGTGTCGCTACTCTCCCAAAACC
AAAAGGTCTCCGCTGACTAGGGCACATCTGACAGAAGTGGAATCAAGGCTAGAAAGACTG
GAACAGCTATTTCTACTGATTTTTCCTCGAGAAGACCTTGACATGATTTTGAAAATGGAT
TCTTTACAGGATATAAAAGCATTGTTAACAGGATTATTTGTACAAGATAATGTGAATAAA
GATGCCGTCACAGATAGATTGGCTTCAGTGGAGACTGATATGCCTCTAACATTGAGACAG
CATAGAATAAGTGCGACATCATCATCGGAAGAGAGTAGTAACAAAGGTCAAAGACAGTTG
ACTGTATCGTCGAGGTCGACCCCGGGAATTCAGATCTCTCGAGCCGCCCCCCCGACCGAT
GTCAGCCTGGGGGACGAGCTCCACTTAGACGGCGAGGACGTGGCGATGGCGCATGCCGAC
GCGCTAGACGATTTCGATCTGGACATGTTGGGGGACGGGGATTCCCCGGGTCCGGGATTT
ACCCCCCACGACTCCGCCCCCTACGGCGCTCTGGATATGTAGCGGATCTACTAGTGGATC
CATAACTTCGTATAGGCTACCGTATACGAAGTTATACTAGTTGTGTGTCCATCAGAGCAT
CTCCTCCTCAGTCTCCTGTGCATGCCCATGCTGAGAGGCAGACCGGCGGTCACACACACA
CGCGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATG
GTCATAGCACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAAT
TTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGA
TATCCCCTATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGT
GTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAATAATATCATCATGAACAATAAA
ACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAAC
GTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCG
CGATAATGTCGGGCAATCAGGTGCGACAATCTATCGCTTGTATGGGAAGCCCGATGCGCC
AGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGT
CAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTAC
TCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGAAAAACAGCATTCCAGGTATT
AGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCG
GTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGC
TCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCG
TAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCTCACC
GGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAA
ATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGC
CATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAA
ATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTT
TTTCTAATCAGAATTGGTTAATTGGTTGTAACACTGGCAGAGCATTACGCTGACTTGACG
GGACGGCGCAAGCTCATGACCAAAATCCCTTAACGTGAGTTACGCGTCGTTCCACTGAGC
GTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAAT
CTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGA
GCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGT
TCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATA
CCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTAC
CGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGG
TTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCG
TGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAG
CGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCT
TTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTC
AGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTT
TTGCTGGCCTTTTGCTCACATGTT

>Th-attP donor plasmid sequence
CTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA
TACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGA

GCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCA
CGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATACGCGTACCGC
TAGCCAGGAAGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTA
GTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGGGCCGTTGCTTC
ACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAA
CAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGCCTG
GCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAA
AACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGAC
CTGTTCGTTGCAACACATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAA
AGCAGGCTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAAT
TGGGTACCTAGTTCCATCTGTTCCGCGCGCGCGCCATCTGAACATGCCGAATTCAAGCAG
CTCCACATCTTCCACAAAGAGCATCGATAAGCTTGTACTGACGGACACACCGAAGCCCCG
GCGGCAACCCTCAGCGGATGCCCCGGGGCTTCACGTTTTCCCAGGTCAGAAGCGGTTTTC
GGGAGTAGTGCCCCAACTGGGGTAACCTTTGAGTTCTCTCAGTTGGGGGCGTAGGGTCGC
CGACATGACACAAGGGGTTGTGACCGGGGTGGACACGTACGCGGGTGCTTACGACCGTCA
GTCGCGGGATCCATAACTTCGTATAGGCTACCGTATACGAAGTTATACTAGTTAGTTCCA
TCTGTTCCGCGCGCGCCATCTGAACATGCCGAATTCAAGCAGCTCCACATCTTCCACA
AAGAGCGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATC
ATGGTCATAGCACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATC
AATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGC
TGATATCCCCTATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCC
CGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAATAATATCATCATGAACAAT
AAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGA
AACGTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGC
TCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGCTTGTATGGGAAGCCCGATGC
GCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGAT
GGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG
TACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGAAAAACAGCATTCCAGGT
ATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCG
CCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCT
CGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGA
GCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCTC
ACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGG
GAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCT
TGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCA
AAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGA
GTTTTTCTAATCAGAATTGGTTAATTGGTTGTAACACTGGCAGAGCATTACGCTGACTTG
ACGGGACGGCGCAAGCTCATGACCAAAATCCCTTAACGTGAGTTACGCGTCGTTCCACTG
AGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGT
AATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCA
AGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATAC
TGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTAC
ATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCT
TACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGG
GGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACA
GCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGT
AAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTA
TCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTC
GTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGC
CTTTTGCTGGCCTTTTGCTCACATGTT

>Th-eGFP donor plasmid sequence
CTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA
TACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGA
GCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCA
CGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATACGCGTACCGC
TAGCCAGGAAGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTA
GTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGGGCCGTTGCTTC
ACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAA
CAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGCCTG
GCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAA

AACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGAC
CTGTTCGTTGCAACACATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAA
AGCAGGCTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAAT
TGGGTACCTAGTTCCATCTGTTCCGCGCGCGCGCCATCTGAACATGCCGAATTCAAGCAG
CTCCACATCTTCCACAAAGAGCATCGATAAGCTTCTGGACCATGGTGAGCAAGGGCGAGG
AGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACA
AGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGT
TCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCT
ACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGT
CCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACT
ACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGA
AGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACA
ACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCA
AGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACA
CCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCG
CCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCG
CCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGAGGAGGAAGATCTAAGCTGA
ACCCTCCTGATGAGAGTGGCCCCGGCTGCATGAGCTGCAAGTGTGTGCTCTCCTGAGACC
CAGCTTGGATCCATAACTTCGTATAGGCTACCGTATACGAAGTTATACTAGTTAGTTCCA
TCTGTTCCGCGCGCGCGCCATCTGAACATGCCGAATTCAAGCAGCTCCACATCTTCCACA
AAGAGCGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATC
ATGGTCATAGCACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATC
AATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGC
TGATATCCCCTATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCC
CGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAATAATATCATCATGAACAAT
AAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGA
AACGTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGC
TCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGCTTGTATGGGAAGCCCGATGC
GCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGAT
GGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG
TACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGAAAAACAGCATTCCAGGT
ATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCG
CCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCT
CGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGA
GCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCTC
ACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGG
GAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCT
TGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCA
AAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGA
GTTTTTCTAATCAGAATTGGTTAATTGGTTGTAACACTGGCAGAGCATTACGCTGACTTG
ACGGGACGGCGCAAGCTCATGACCAAAATCCCTTAACGTGAGTTACGCGTCGTTCCACTG
AGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGT
AATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCA
AGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATAC
TGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTAC
ATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCT
TACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGG
GGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACA
GCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGT
AAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTA
TCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTC
GTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGC
CTTTTGCTGGCCTTTTGCTCACATGTT

>Th-GAL4 donor plasmid sequence
CTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA
TACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGA
GCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCA
CGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATACGCGTACCGC
TAGCCAGGAAGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTA
GTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGGGCCGTTGCTTC

ACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAA
CAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGCCTG
GCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAA
AACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGAC
CTGTTCGTTGCAACACATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAA
AGCAGGCTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAAT
TGGGTACCTAGTTCCATCTGTTCCGCGCGCGCGCCATCTGAACATGCCGAATTCAAGCAG
CTCCACATCTTCCACAAAGAGCATCGATAAGCTTGATCCGATGAAGCTACTGTCTTCTAT
CGAACAAGCATGCGATATTTGCCGACTTAAAAAGCTCAAGTGCTCCAAAGAAAAACCGAA
GTGCGCCAAGTGTCTGAAGAACAACTGGGAGTGTCGCTACTCTCCCAAAACCAAAAGGTC
TCCGCTGACTAGGGCACATCTGACAGAAGTGGAATCAAGGCTAGAAAGACTGGAACAGCT
ATTTCTACTGATTTTTCCTCGAGAAGACCTTGACATGATTTTGAAAATGGATTCTTTACA
GGATATAAAAGCATTGTTAACAGGATTATTTGTACAAGATAATGTGAATAAAGATGCCGT
CACAGATAGATTGGCTTCAGTGGAGACTGATATGCCTCTAACATTGAGACAGCATAGAAT
AAGTGCGACATCATCATCGGAAGAGAGTAGTAACAAAGGTCAAAGACAGTTGACTGTATC
GTCGAGGTCGACCCCGGGAATTCAGATCTCTCGAGCCGCCCCCCCGACCGATGTCAGCCT
GGGGGACGAGCTCCACTTAGACGGCGAGGACGTGGCGATGGCGCATGCCGACGCGCTAGA
CGATTTCGATCTGGACATGTTGGGGGACGGGGATTCCCCGGGTCCGGGATTTACCCCCCA
CGACTCCGCCCCCTACGGCGCTCTGGATATGTAGCGGATCTACTAGTGGATCCATAACTT
CGTATAGGCTACCGTATACGAAGTTATACTAGTTAGTTCCATCTGTTCCGCGCGCGCGCC
ATCTGAACATGCCGAATTCAAGCAGCTCCACATCTTCCACAAAGAGCGAGCTCCAGCTTT
TGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCACCCAGCT
TTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTTGCAACGAACAG
GTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATATCCCCTATAGTGAG
TCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATCTCTGA
TGTTACATTGCACAAGATAAAATAATATCATCATGAACAATAAAACTGTCTGCTTACATA
AACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGATTA
AATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAA
TCAGGTGCGACAATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAA
CATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTG
ACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGG
TTACTCACCACTGCGATCCCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATCCTGAT
TCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCT
GTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGA
ATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTT
GAACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCTCACCGGATTCAGTCGTCACT
CATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATT
GATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGC
CTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAAT
CCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAGAATTG
GTTAATTGGTTGTAACACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGCAAGCTCA
TGACCAAAATCCCTTAACGTGAGTTACGCGTCGTTCCACTGAGCGTCAGACCCCGTAGAA
AAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACA
AAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT
CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCG
TAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATC
CTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGA
CGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCC
AGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGC
GCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACA
GGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGG
TTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA
TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCT
CACATGTT