UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA VEGETAL

Investigating the role of *foxj1a*$^{+}$ ependymal cells in zebrafish spinal cord regeneration using lineage tracing and cell ablation approaches

Ana Sofia Rocha Batalha

Mestrado em Biologia Molecular e Genética

Dissertação orientada por:
Professora Doutora Maria Gabriela Rodrigues
(Departamento de Biologia Animal, FCUL)

Doutora Ana Catarina Esteves Ribeiro
(Leonor Saúde Lab, Instituto de Medicina Molecular)

2017

Acknowledgements

Em primeiro lugar, tenho de agradecer muito às Doutoras Leonor Saúde e Ana Ribeiro por me terem proposto um projecto ambicioso como este e por todo o apoio e orientação. Ainda que o resultado final não tenha sido o esperado, foi devido a este ano que pude dar os primeiros passos como "mini-investigadora" num dos melhores institutos de ciência em Portugal. Tive também a oportunidade de conhecer pessoas maravilhosas que não me deixaram desanimar depois de meses de resultados falhados. Um enorme obrigado à Ana Margarida por toda a ajuda técnica e por me ter suportado como uma segunda sombra nos primeiros dias e ainda responder pacientemente às minhas dúvidas mais parvas no últimos, sempre com um sorriso. Também tenho muito que agradecer à Aida e à Lara por toda a ajuda com os peixes (especialmente com os mais teimosos) e pela energia positiva inesgotável − faz toda a diferença! Aproveito também para agradecer à Professora Gabriela pela orientação interna e pelo interesse neste trabalho.

Quem me conhece melhor sabe que esta viagem começou há muito tempo, quando uma menina de 8 anos descobriu um livro de genética na biblioteca da escola primária e ficou irreversivelmente fascinada. É muito giro dizer a toda a gente que não queremos ser médicos, ou professores, mas que queremos trabalhar em genética; é muito menos giro quando ninguém percebe bem o que queremos dizer ou, pior, quando dizem que nunca lá iremos chegar. As pessoas que marcam a diferença são as que não só nos encorajam como também conseguem dar-nos um bocadinho da sua paixão por estes mecanismos invisíveis que nos regulam e que nos unem com todos os outros seres numa única filogenia. Duas pessoas (R.G. e R.Z.) foram especialmente fulcrais para que eu voltasse a querer estudar bio molecular e genética e, consequentemente, para que chegasse hoje ao ponto de ter uma tese de mestrado terminada; terão sempre a minha gratidão.

Mas há tanto mais do que o que se passou dentro do laboratório. Há todos os momentos de distracção, de festa, de desabafo, de amizades que me acompanharam sempre durante este ano e, acima de tudo, há todas as pessoas que deles fizeram parte. Entre todos os que estiveram presentes quero agradecer especialmente: à Catarina, por ser uma excelente amiga e quase-psicóloga; ao Miguel, por todas as discussões e também por me mostrar que se podem fazer omeletas sem ovos; e ao Fábio, por ser um segundo irmão e ter sempre vinte histórias diferentes para distrair qualquer um.

"Os últimos são sempre os primeiros". À minha família, ao meu pai, à minha mãe, ao meu irmão: todo o apoio que me deram, desde sempre, toda a força, toda a confiança, são inestimáveis. Muito, muito, muito obrigada por estarem sempre do meu lado, sem me deixarem cair.

Um obrigado muito especial aos meus avós, José e Júlia, que mesmo sem o ver sempre souberam que este trabalho seria terminado.

Deus ao mar o perigo e o abismo deu,
Mas nele é que espelhou o céu.

Fernando Pessoa (in *Mensagem*)

Abstract

Spinal cord injury (SCI) is a disabling condition affecting millions of people worldwide. In order to improve functional recovery, new therapies are being devised to counteract the non-regenerative environment of the mammalian spinal cord. Ependymal cells (EC) of the spinal cord central canal have been proven to hold neural stem cell properties *in vitro*, suggesting that a pro-neurogenic fate could potentially be promoted *in vivo* after a lesion. The zebrafish *Danio rerio* has proved a valuable tool for developmental studies and, in recent years, for regenerative processes due to its remarkable ability to regenerate several organs and tissues. After SCI new neurons and glia are generated by the cells lining the central canal, which hold a strong resemblance to mammalian EC. Both mammalian and zebrafish EC have in common the expression of the transcription factor responsible for cilia formation: Foxj1/Foxj1a. However, no study has been performed to discover the progeny of zebrafish *foxj1a*$^{+}$ cells after a lesion or their functional role in regeneration. This work aimed at generating stable transgenic zebrafish lines that allowed the investigation of *foxj1a*$^{+}$ ependymal cell progeny after SCI, using lineage tracing, and also their specific ablation with a suicide gene.

A total of three constructs for lineage tracing lines and two constructs for cell ablation lines was injected with transposase mRNA into one-cell stage embryos but only one of the lineage tracing constructs resulted in clear germline transmission to F1 embryos. All the other four constructs proved very difficult to generate stable expression in injected embryos as well as an adequate number of positives, even after several optimization attempts. Reasons for this lack of transgenesis efficiency are not completely understood but may include inadequate coding sequence features, undetected problems in untranslated regions, or obstacles to transposition such as mutations in transposase recognition sequences.

Since the lineage tracing is based on the CreER^{T2}/LoxP technology, optimization of CreER^{T2} activation was also performed. It was found that recombination efficiency is directly influenced by 4-OHT concentration and enhanced by treatment with pronase. However, when using a lineage tracing line no recombination was detected in *foxj1a*$^{+}$ cells even after induced proliferation, prompting the need to validate CreER^{T2} function in this line.

Keywords: zebrafish, regeneration, spinal cord, ependymal cells, *foxj1a*

Resumo Alargado[1]

As lesões da espinal medula afectam actualmente milhões de pessoas em todo o mundo, que se deparam com um agravamento radical da sua qualidade de vida que, na maioria dos casos, não poderá ser recuperada. A Organização Mundial de Saúde define "lesão da espinal medula" como a perda total ou parcial de função neural provocada por um trauma ou uma patologia da espinal medula, resultando na diminuição do controlo motor abaixo do local da lesão assim como na perda de sensibilidade e regulação do sistema nervoso autónomo. A resposta clínica padrão actual tem como objectivo apenas tentar impedir o alastramento da lesão e consiste na realização de cirurgias para estabilizar a estrutura da coluna vertebral e descomprimir o local da lesão, seguidas de tratamento com metilprednisolona. No entanto, dado que o ambiente da lesão em mamíferos é extremamente inibitório para a ocorrência de regeneração, novos estudos têm tentado obter soluções terapêuticas que actuem na promoção de um ambiente pró-regenerativo, assim como na protecção dos tecidos que permanecerem intactos e funcionais após a lesão. A maior parte destas novas terapêuticas tem como objectivo prevenir o aparecimento de fenómenos associados à chamada lesão secundária ou de mitigar o seu efeito, dado que estes são em grande parte responsáveis pelo grau da lesão a longo prazo.

Ao contrário do que se considerou durante décadas, investigações dos últimos 20 anos têm demonstrado que existem determinadas áreas no sistema nervoso central em que ocorre a formação de novos neurónios durante a vida adulta. Verificou-se também que existem células com capacidades estaminais neurais nestas e noutras zonas do sistema nervoso central, o que sugere que algumas células no adulto poderão ser estimuladas a originar novos neurónios e glia após uma lesão. As células ciliadas que revestem o interior do canal central na espinal medula, designadas células ependimárias, são um exemplo disso. Vários estudos observaram que estas células ependimárias respondem a uma lesão na espinal medula através da proliferação e migração para o local da lesão, onde originam astrócitos e oligodendrócitos; no entanto, são também capazes de originar neurónios quando cultivadas *in vitro*. O seu perfil de expressão inclui vários marcadores associados a células estaminais neurais, apoiando observações que indicam estas células como as únicas a possuir capacidades multipotentes no nicho do canal central. Pelo facto de serem as únicas células multiciliadas no canal central, as células ependimárias de mamífero podem ser identificadas pela expressão do factor de transcrição Foxj1, conhecido pelo seu papel como regulador da ciliogénese.

O peixe-zebra (*Danio rerio*) já é conhecido como modelo em biologia do desenvolvimento há mais de 30 anos, mas a sua utilização para o estudo da regeneração de tecidos e órgãos é bastante mais recente. As extraordinárias capacidades de regeneração deste organismo, aliadas à bateria de métodos genéticos e moleculares desenvolvidos e adaptados para este modelo, tornaram o peixe-zebra num aliado inestimável para entender processos regenerativos e compará-los com a situação dos mamíferos, cujas capacidades regenerativas são muito inferiores. As características biológicas do peixe-zebra são particularmente apelativas, por exemplo para geração de linhas transgénicas: são capazes de gerar um grande número de embriões – transparentes, que permitem a observação de fenótipos relativamente cedo – e atingem a maturidade aos 3 meses, acelerando o processo de geração da linha. Os métodos desenvolvidos para a geração de peixes transgénicos também têm demonstrado elevados níveis de sucesso, especialmente o sistema Tol2, que se baseia na injecção do DNA desejado juntamente com mRNA que codifica para uma transposase; esta reconhece sequências específicas que flanqueiam o

[1] Nota: Este texto não foi escrito ao abrigo do Acordo Ortográfico da Língua Portuguesa de 1990.

DNA injectado e insere-o no genoma do embrião. Entre as várias estruturas que o peixe-zebra consegue regenerar encontra-se a espinal medula, que obtém uma recuperação funcional quase completa um mês após a lesão. Este resultado deve-se tanto ao crescimento de axónios seccionados como à formação de novos neurónios no local da lesão, mecanismos promovidos por um ambiente pró-regenerativo. Os novos neurónios e células da glia são produzidos por células que revestem o canal central da espinal medula, tendo por isso função de epêndima, mas que também apresentam outras características morfológicas reminiscentes das células da glia radial, progenitores neurais durante o desenvolvimento (tanto em peixe-zebra como em mamífero). A utilização de marcadores moleculares para identificação destas células tem sido controversa, mas existe um marcador que poderá ser utilizado de forma clara. Devido à sua função ependimária, estas células possuem cílios e expressam por isso o factor de transcrição Foxj1a, um dos ortólogos do Foxj1 de mamífero. Será por isso importante desenvolver ferramentas que permitam estudar o papel que as células *foxj1a*$^{+}$ da espinal medula do peixe-zebra desempenham após uma lesão, que pode ser realizado através da identificação da sua descendência (experiência de *lineage tracing*) e da sua função real no contexto de uma lesão promovendo a sua ablação específica. Este trabalho apresentava assim como objectivo principal o estabelecimento de linhas transgénicas estáveis em peixe-zebra que permitissem efectuar de forma eficaz e separadamente: 1) a marcação permanente das células *foxj1a*$^{+}$ e da sua descendência; 2) a ablação específica das células *foxj1a*$^{+}$, em ambos os casos após uma lesão da espinal medula.

Para a geração da linha para *lineage tracing* foram testados três *constructs* separadamente, mas apenas com um deles foi possível observar passagem evidente do transgene para embriões F1. Ainda assim, o sucesso desta transgénese foi inferior ao reportado para o método utilizado. Os outros dois *constructs* apresentaram percentagens muito baixas de embriões positivos e de níveis de expressão dos transgenes após injecção, mesmo depois de tentativas de optimização, e de todos os que cresceram até atingir maturidade não foram detectados peixes fundadores de forma inequívoca. Para a geração da linha para ablação das células *foxj1a*$^{+}$ foram testados dois *constructs* com os quais não foi possível obter um número suficiente de embriões positivos para crescer. O primeiro foi abandonado quando se observou que várias características do plasmídeo em que estava inserido não seriam adequadas para manter um nível de expressão adequado à experiência, como a falta da sequência Kozak e do sinal de poliadenilação a flanquear a sequência codificante. O segundo foi desenhado de forma a optimizar todas as sequências para promover uma expressão eficiente do transgene, contendo sequências de reconhecimento da transposase e sequências regulatórias a 5' e 3' da sequência codificante vindas de um plasmídeo utilizado como controlo positivo para as injecções, assim como a sequência codificante de uma proteína fluorescente (como repórter) que apresentara bons níveis de expressão com o promotor *foxj1a*. Não obstante, este *construct* parece não ter sido integrado no genoma dos embriões em que foi injectado dado que praticamente toda a expressão observada às 24 horas-pós-fertilização tinha sido perdida 4 dias depois. As razões que levaram ao insucesso na obtenção de embriões positivos após a injecção são desconhecidas para a maioria dos *constructs* testados, mas poderão estar relacionadas com mutações não detectadas nas regiões regulatórias ou nas sequências de reconhecimento da transposase, que impediriam a correcta expressão do *construct* ou a sua inserção no genoma, respectivamente.

A experiência de *lineage tracing* concebida baseia-se na utilização do sistema de recombinação CreER^{T2}/LoxP, que utiliza uma versão modificada da enzima Cre a que foi acrescentado um receptor de estrogénio. Esta modificação adiciona um passo de controlo temporal à experiência visto que é necessário activar a CreER^{T2} com administração da droga 4-OHT para que a enzima possa efectuar a recombinação entre locais LoxP e assim promover a marcação de células *foxj1a*$^{+}$. O segundo objectivo deste trabalho foi então a optimização do protocolo de activação da CreER^{T2}, que foi conseguido pela primeira vez neste laboratório. Inicialmente foram utilizados embriões que expressavam a CreER^{T2} e a cassete repórter na maioria das células (expressão induzida pelo promotor de uma proteína *heat shock*),

tendo-se verificado que a eficácia da recombinação foi directamente influenciada pela concentração de 4-OHT administrada. A remoção do córion dos embriões por adição de pronase também promoveu este efeito ao permitir um acesso mais rápido da droga aos tecidos. Embriões e larvas F1 dos fundadores obtidos com a injecção de um dos *constructs* para *lineage tracing* foram também utilizados para testar a activação da $CreER^{T2}$, mas neste caso não se observou a ocorrência de recombinação em células *foxj1a*$^{+}$, mesmo após a sua proliferação ter sido induzida através da realização de uma lesão na espinal medula.

Devido ao insucesso em gerar linhas transgénicas funcionais neste trabalho, será necessário utilizar novos métodos e/ou novos *constructs* para obter ferramentas biológicas que permitam no futuro compreender a função das células *foxj1a*$^{+}$ durante a regeneração da espinal medula no peixe-zebra.

Palavras-chave: peixe-zebra, regeneração, espinal medula, células ependimárias, *foxj1a*

Table of Contents

List of Figures

List of Tables

Abbreviations

4-OHT: 4-hydroxytamoxifen

aa.: Aminoacid[2]

***actb1*:** *β-actin1* gene

BLAST: Basic Local Alignment Search Tool

BLBP: Brain Lipid Binding Protein

bp: Base pairs (multiple: kbp, kilo base pairs, 10^3 bp)

BSA: Bovine serum albumin

(c)DNA: (complementary) Deoxyribonucleic acid

CNS: Central nervous system

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

CU: Codon usage

CUPS: Codon usage preference score

DMSO: Dimethyl sulfoxide

dpi: Days-post-injury

dpf: Days-post-fertilization

EC: Ependymal cell(s)

(E)GFP: (Enhanced) Green Fluorescent Protein

ERG: Ependymo-radial glial cell(s)

EtOH: Ethanol

GFAP: Glial Fibrillary Acid Protein

GLAST: Glutamate Aspartate Transporter

h: Hours

HDAC5: Histone Deacetylase 5

hpf: Hours-post-fertilization

IMM: Instituto de Medicina Molecular

MeOH: Methanol

min: Minutes

[2] Individual aminoacids are referred throughout the text by the standard 3-letter abbreviations.

(m)RNA: (messenger) Ribonucleic acid

mTOR: Mammalian Target of Rapamycin

M&M: Materials and Methods

NCBI: National Center for Biotechnology Information

NEB: New England Biolabs

NSC: Neural stem cell

nt: Nucleotides

OMIM: Online Mendelian Inheritance in Man

ON: Overnight

ORF: Open reading frame

PBS: Phosphate buffered saline

PCR: Polymerase Chain Reaction

PFA: Paraformaldehyde

PNS: Peripheral nervous system

PTU: 1-phenyl-2-thiourea

rpm: Revolutions per minute

RT: Room temperature

RT-PCR: Reverse Transcription Polymerase Chain Reaction

s: Seconds

SC: Spinal cord

SCI: Spinal cord injury

SGZ: Sub-granular zone (hippocampus)

STAT3: Signal Transducer and Activator of Transcription 3

SVZ: Sub-ventricular zone

TAE: Tris/Acetate/EDTA buffer

TALEN: Transcription Activator-Like Effector Nuclease

TK: Thymidine kinase enzyme (from Herpes Simplex Virus)

Tm: Melting temperature

Tu: Tübingen zebrafish line

WHO: World Health Organization

WISH: Whole-mount *in situ* hybridization

Introduction

1. Spinal cord injury

Spinal cord injury (SCI) is defined by the World Health Organization (WHO) as damage inflicted to the spinal cord by means of a traumatic event or an underlying pathology that results in partial or complete loss of neural function posteriorly to the lesion site [1]. Immediate consequences of SCI span not only impairment of motor control but also of sensory input and autonomic regulation, leading to several other clinical conditions, like loss of bowel and bladder control, even in mild cases. The level and permanency of the disabilities caused by SCI call for an immediate joint action from the scientific and medical communities in order to find better ways to ameliorate, and preferably eliminate, these patients' symptoms.

1.1 The uninjured and injured spinal cord of mammals

The spinal cord is the component of the central nervous system (CNS) responsible for mediating communication between the encephalon and the peripheral nervous system (PNS) [2, 3]. Anatomically, the spinal cord consists of a cylinder of soft tissue encased inside the vertebral column and surrounded by three layers of connective tissue (meninges): dura mater, arachnoid mater and pia mater. The centre of the spinal cord is composed of a lumen, the central canal, lined by an epithelium of ciliated ependymal cells (EC) that promote cerebrospinal fluid flow. In a transverse section of the cord it is possible to distinguish two major areas: the grey matter, located centrally and containing interneurons, the cell bodies and dendrites of efferent neurons, afferent fibres and glial cells; and the white matter, comprising the surrounding myelinated fibres of interneurons that run longitudinally. The spinal cord proper ends at the level of L1 vertebra (in human; L3 in rats) [4] and is followed posteriorly by descending spinal roots composing the terminal filament and *cauda equina* [3].

SCI comprises damage affecting any part of the spinal cord down to *cauda equina* and starts with an immediate mechanic lesion to the cord tissue, designated primary injury. During the first two hours post trauma [5] there is disruption of neural tissue [6] and cord vasculature, causing defective blood flow to the lesion site and haemorrhage. The extent of the primary injury can be a strong indication of prognosis [6], but it is also known that the severity of these lesions is mostly dependent on damage derived from secondary injury [5], which comprises all the secondary phenomena that occur from the first minutes to the first weeks post trauma [4]. These include lesion extension and consequences of blood flow disruption, like ischemia, leucocyte infiltration and consequent enhanced inflammation, as well as extravasation of excitatory aminoacids [4-8]. All these factors potentiate cytotoxicity, leading to oligodendrocyte and neuron death by necrosis and apoptosis and release of factors inhibiting regeneration. The chronic phase involves further irreversible damage to the spinal cord, such as white matter demyelination, progression of apoptosis in both orthograde and retrograde directions and reactive astrocyte activation [4-8], which leads to the formation of a glial scar preventing axon regrowth [4]. Permanent excitability of some neurons may also develop, leading to chronic pain and even mood-related disorders.

The current clinical approach to SCI consists in decompression surgery and administration of methylprednisolone, which has proved ineffective in promoting substantial motor recovery [4, 9-11]. Current investigations focus on finding novel treatment strategies associating bioactive molecules with new delivery systems. These can be grouped according to their main goal: 1) modulation of extracellular matrix composition and endogenous cell response to prevent secondary injury and the onset of a non-

regenerative environment; 2) stimulation of intrinsic regeneration capabilities in cells at the lesion site (including axon regrowth and cell differentiation); 3) transplantation of cells with regenerative abilities to the lesion site [4, 7, 12-14].

1.1.1 Obstacles to mammalian regeneration in the CNS

After a SCI, the injury site is invaded with a multitude of chemical signals that modulate cellular responses and contribute to a non-regenerative environment [4, 13]. Oligodendrocyte death causes myelin debris to aggregate around neurons, increasing concentrations of myelin-associated inhibitors like Nogo-A protein and myelin-associated glycoprotein, which act synergistically to inhibit axon regrowth [4, 15]. Secretion of chondroitin sulphate proteoglycans (CSPGs) by reactive astrocytes further exacerbates this effect due to the inhibitory action of these molecules, expressed during neural development as negative cues for neurite growth [13]. Other molecules expressed in astrocytes or fibroblasts, such as Ephrin-B2 and Semaphorin3, are known to also inhibit regeneration by inducing the formation of the glial scar cellular pattern and preventing growing neurites from entering the glial scar [13]. CNS neurons may also lack some intrinsic features capable of supporting effective axon regrowth after an injury. Activation of the mTOR and STAT3 pathways, both proved to contribute to peripheral axon regrowth, is absent in CNS neurons after an injury. Likewise, chromatin remodelling via HDAC5 does not show the same level of activation as in injured PNS neurons, where it promotes expression of regeneration-associated genes [16].

1.2 Neurogenesis in the adult CNS

For many decades it was believed that the CNS was not capable of regeneration [17], but that view has been changing as new areas of the adult brain are found to possess cells with neural stem cell (NSC) properties both *in vitro* and *in vivo* [18]. NSC are adult multipotent stem cells, meaning that they are found in adult tissues and exhibit the capacity for proliferation and self-renewal as well as for generating progeny of all three neural lineages: neurons, oligodendrocytes and astroglia [19]. The first brain areas where adult neurogenesis was confirmed were the subventricular zone (SVZ) of the forebrain and the subgranular zone (SGZ) of the hippocampus [18-21], both in mammalian models and in humans, but more recently neural progenitor activity has also been found in other brain areas, such as neocortex, cerebellum, amygdala, hypothalamus and the meninges [18, 21].

1.2.1 Neural progenitors in the spinal cord

In addition to the encephalon, the spinal cord has also been reported to harbour cells with NSC properties. Initial reports in the late 1990's showed the existence of cells in the ventricular layer of the central canal capable of self-renewal and differentiation into the three neural lineages when cultured *in vitro* [22-24]. These correspond to the epithelium-forming ependymal cells that can be found lining the central canal lumen and have since been associated with other NSC features. Several studies characterized EC as being a quiescent population of cells with a very slow division rate [24-26] and a symmetric division plane [26], indicating that homeostatic division serves mainly to maintain the population size. In these conditions, all EC are found to express Vimentin and neural stemness markers such as Sox2 [27], Sox9 [28] and Musashi1 [25, 26, 29, 30]. A subpopulation of these cells was also found to express Nestin, another protein associated with NSC [31]. Regarding expression of the astrocyte marker GFAP [32], some authors claim to have observed its expression in a fraction of dorsally located EC [26], while most studies report all EC as $GFAP^-$ [25, 30, 33]. Nevertheless, EC located at the dorsal tip of the central canal exhibit some differences in morphology from the more common cuboidal EC or tanycytes [34], presenting radial processes that extend dorsally and can reach the pial surface [26, 30]. These features show clear similarities both with type B cells, which are considered

NSC residing in the forebrain SVZ, and with radial glial cells, which functioned as NSC during prenatal neurogenesis before giving rise to EC [35]. Interestingly, most of these features have also been found in adult human ependymal cells extracted from the spinal cord, which have also been proved to hold multilineage differentiation potential *in vitro* [36].

More importantly, EC are known to develop an acute response to lesions in the spinal cord, within which three main hallmarks can be defined: proliferation, migration and increased multipotent capacity. Immediately after a SCI is inflicted, ependymal cells start to proliferate [24, 25, 30, 37-39] at an increasing rate that peaks at 3 days-post-injury (dpi) and then declines slowly [37]. Contrary to their quiescent state, mitotic divisions after an injury are mostly asymmetrical, indicating EC are not only self-renewing but also generating new cells to be differentiated [24]. During this time, EC also show a higher capacity to differentiate into the three neural lineages *in vitro* [25, 30, 40], which can be due in part to the strong upregulation of Nestin expression seen in the first 24 hours post injury [33, 37, 39]. Starting from 3 dpi, EC are observed to migrate from the central canal to the injury centre where they differentiate into astrocytes and contribute to the inner portion of the glial scar [24, 25, 30, 37, 38], stabilizing the lesion core and preventing cavity formation and injury spreading [41]. A minor fraction of EC migrate to the spinal cord parenchyma to give rise to oligodendrocytes [25, 30], but so far there has been no proof of neuron generation *in vivo*. Transplantation experiments have also shown that ependymal cells' response as NSC may be greatly dependent on the environment in which they are placed. For instance, spinal cord EC transplanted to the dentate gyrus of the hippocampus were able to generate neurons, recapitulating the endogenous neurogenic activity of that site [42].

Moreover, EC were confirmed to be the only cell type in the central canal niche to hold NSC properties [25, 30, 43], rejecting former hypothesis that $GFAP^+$ astrocytes or oligodendrocyte precursors in the subventricular zone could also hold similar potential. Ependymal cells can be precisely identified by the expression of the forkhead domain-containing Foxj1 transcription factor (formerly known as HFH-4; OMIM entry *602291). Foxj1 is considered a major regulator of motile cilia formation [44-47] found necessary for EC differentiation [48]; since EC are the only multiciliated cells in the central canal niche [3], expression of this transcription factor identifies exclusively ependymal cells in this context.

2. The zebrafish

Danio rerio (F. Hamilton, 1822), commonly known as zebrafish, is a small freshwater teleost fish of the Cypriniformes order natural of shallow streams or pools in North-eastern India [49, 50] that had been kept as an aquarium fish for many years before being recognized as an excellent model organism for biological studies [51]. In laboratory, adults can reach up to 4-5 cm in length, growing continuously during the average lifespan of 2-3 years [49]. Reproductive maturity, however, can be reached at 90 days-post-fertilization (dpf) in optimal growing conditions, granting this species a short generation time that greatly facilitates genetic screens or procedures like transgenesis. Like many other fish species, fertilization is external, which allows control over the amount of embryos generated by managing the time for which male and female fish are kept together. Reproductively active couples can spawn up to 300 embryos per week [52], each about 0.7 mm in diameter [53], visible to the naked eye and easily examined under a stereo microscope. Embryo development is also external which, adding to the natural transparency maintained until 24 hours-post-fertilization (hpf), makes the zebrafish an ideal model to study early developmental processes (for description of developmental stages up to 72 hpf please refer to [53]). For all these reasons, and the relative low cost of maintaining a fish facility (compared to a mouse facility with the same number of animals) [54], the zebrafish has risen as one of the major tools for studying developmental biology in the last three decades [55, 56].

The growing significance of this vertebrate model also led to the development of a battery of genetic and molecular techniques, as well as genomic data. Forward and reverse genetic methods were optimized and today it is possible to precisely mutate a gene using zinc-finger nucleases [57], TALENs [58] or the CRISPR/Cas9 system [59], or to knock down its expression with antisense morpholino oligonucleotides [60]. Transgenesis can be carried out with various genome integration systems, from transposons [61] to phage integrases [62] and meganucleases [63], and diverse strategies exist for conditional [64, 65] and inducible [66-68] expression of transgenes. The sequencing and ongoing annotation of the zebrafish genome [69] was also a major contribution to the field, bringing to light the relevant similarities between zebrafish systems and genes and those of humans and highlighting *D. rerio* as an ideal model to study human disease [54, 70].

So far, the zebrafish has been used to uncover disease mechanisms ranging from metabolic [71], heart [72] and hematopoietic [73] disorders to muscular dystrophies [74], neurodegenerative conditions [75] and several cancer types (including solid tumours) [76]. More recently, it has also emerged as a very efficient platform for drug discovery [77] allowing large scale toxicity and efficacy analysis to be performed on disease models relatable to the human condition.

2.1 Lessons in regeneration

Post embryonic regeneration is a trait heterogeneously dispersed throughout the animal kingdom [78]. It seems to follow a pattern in which simpler, more ancient organisms are provided with higher regenerative abilities, that weaken during the course of evolution. For instance, hydras are well known for their ability to regenerate their whole body [79] while some amphibians like the newt are able to replace a missing limb [80]; on the opposite side of the spectrum lie mammals, whose regenerative abilities are far less extensive.

Some teleost fish, and specifically the zebrafish, have also proved to hold amazing regeneration abilities. The fins were the first structures observed to undergo successful regeneration [81] and the more recent ascension of *D. rerio* as a model organism brought a deeper understanding of the underlying mechanisms. But zebrafish regeneration can also be used as an approach to human disease. Several injury methods have been used to mimic the effects of a myocardial infarction by cardiomyocyte loss in the zebrafish and all resulted in replacement of the lost tissue with new cardiac muscle cells. In fact, it was found that spared cardiomyocytes responded to injury by proliferating, dedifferentiating and re-differentiating in a permissive environment to replace the missing tissue [82, 83]. Additionally the zebrafish has also been studied for its ability to regenerate the kidneys (pro- and mesonephros) [84], pancreas [85], liver [86], lateral line (hair cells) [87] and other body parts.

Regarding CNS injuries, zebrafish exhibit both axonal regeneration (regrowth of axons after transection) and neuronal regeneration (generation of new neurons and glia) [88], which have been extensively studied due to the relevant similarities with the human CNS anatomy [54]. Particularly in what concerns the spinal cord, it was found that zebrafish regenerate the majority of severed axons after SCI [89], exhibiting an almost complete functional recovery one month after a crush injury [90]. Axon regrowth occurs in the presence of myelin [91] (considered inhibitory for mammalian axon regrowth) and other molecules that constitute a pro-regenerative environment, with gliosis and inflammation being transient events [92]. The source of new neurons and glial cells after a lesion has been identified as ependymo-radial glial cells (ERG), so called for the presence of both ependymal and radial glial features [91]: while composing the ependymal layer of the spinal cord, they also extend projections that touch the pial surface, strongly resembling mammalian radial glia and type B cells of the adult brain. After an injury, ERG exhibit strong proliferation [90, 91, 93-97] and upregulate NSC markers such as Oct4, Sox2, Nestin, Sox11b and Vimentin [93-95, 97]. From 3 dpi, ERG migrate [90, 94, 95] and some differentiate

into GFAP$^+$ glial cells that will form the glial bridge, a structure connecting both sides of the transected spinal cord that works as a scaffold for regenerating axons traversing the lesion site [94]. Other ERG give rise to new neurons while migrating, as was observed by the co-localization of ERG and proliferation markers with HuC/D [90, 91, 95], a marker of early neurogenesis [98]. Moreover, different neuron types are generated according to the dorsal-ventral position of the source ERG, mirroring the neural marker patterning observed during embryonic neurogenesis [96, 99].

Identification criteria for ERG have not been consensual [100]. While most studies recognized these cells solely based on their morphology and localization, others have used molecular markers like GLAST, BLBP [93], Olig2 (for motor neuron generating ERG) [91, 99, 101] or GFAP [94], which is downregulated upon injury [90, 94, 97]. Alternatively, given that these cells compose the ventricular layer of the spinal cord and have ependymal function, a different marker could be used to label ERG. The forkhead box transcription factor Foxj1a (one of the zebrafish orthologs of mouse Foxj1) is known to control motile ciliogenesis in zebrafish [46, 47, 102], thus being expressed in cells possessing motile cilia such as in Kupffer's vesicle, pronephric ducts and also ependymal cells [103, 104]. One report has in fact observed an upregulation of *foxj1a* expression in ependymal cells after a SCI [103], probably associated with the ERG proliferative response, promoting further use of this transcription factor to identify ERG in a context of zebrafish spinal cord lesion.

3. Objectives

The zebrafish as a model organism presenting remarkable regenerating abilities can provide a privileged insight into the mechanisms underlying functional repair after SCI. Moreover, given the structural similarities between zebrafish and human SC, understanding how this fish regenerates can point to existing features in mammalian systems that can be exploited or stimulated to promote a more efficient recovery, and even regeneration, after a lesion. One such similarity is the presence and nature of ependymal cells in the SC central canal, which have been proved to hold NSC properties in both systems (*in vivo* in the zebrafish, *in vitro* in mammals). These EC are also comparable in morphology and are thought to originate from the same embryonic cell type (radial glia), even though *D. rerio* EC remain more similar to their progenitor during adulthood. In mouse models these cells are easily identified by the expression of Foxj1 transcription factor, which is the ortholog of zebrafish Foxj1a. However, Foxj1a has only been used once as a spinal cord EC marker and no specific analysis was made into the NSC abilities of *foxj1a*$^+$ ependymal cells or their functional role during SC regeneration.

To bridge that gap, two different approaches were conceived: 1) a lineage tracing experiment in which *foxj1a*$^+$ cells are labelled right before a SCI and their progeny can be traced at several time points after the lesion due to permanent expression of a reporter protein; 2) a functional ablation experiment in which *foxj1a*$^+$ cells are selectively eliminated right after a SCI so that their effective role in regeneration can be assessed. Both approaches depend on the previous establishment of stable transgenic lines possessing: 1) a reporter construct with coding sequences for two fluorescent proteins in tandem, in which the first is flanked by equal LoxP sites, and a Cre driver construct to promote recombination only in *foxj1a*$^+$ cells (lineage tracing); 2) the "suicide gene" *thymidine kinase (TK)* from Herpes Simplex Virus driven by the *foxj1a* promoter, whose enzyme catalyses the conversion of the harmless substrate ganciclovir into a toxic product that selectively kills proliferating cells (functional ablation) [105].

The first objective of this work is then to generate the transgenic lines with which both approaches can be performed. Given previous difficulties in generating a functional transgenic line for lineage tracing, three different lines will be established to improve the likelihood that one will be effective. Additionally, since this version of Cre requires activation with an oestrogen analogue [106, 107], the protocol for Cre activation must also be optimized.

Materials and Methods

Zebrafish lines and husbandry

Fish were kept at the IMM Fish Facility at controlled standard conditions [52] including temperature (28ºC) and light/dark cycle (14h/10h). Embryos were raised in embryo medium-containing Petri dishes (embryo medium: 5mM NaCl, 160µM KCl, 340µM CaCl2.2H2O, 340µM MgSO4.7H2O, 0.3mg.L^{-1} methylene blue) placed in an incubator at the same temperature and light/dark cycle until 5 or 6 days-post-fertilization (dpf) and were then transferred to main system tanks with approximately 60 larvae per tank. Developmental staging of embryos and larvae was performed according to Kimmel *et al* [53]. Animals were considered fit for mating at 3 months of age or after reaching 2-3 cm in length [49]. Two zebrafish wild-type lines were used, AB [55] and Tübingen [108]. The following transgenic lines already available at the facility were used: *Tg(hsp70l:mCherry-T2a-CreER^{T2})$^{\#12}$* (referred as *hs:mCh-Cre*) [109]; *Tg(–3.5ubi:EGFP(floxed)-mCherry*) (referred as *ubi:Switch*) [110]; *Tg(hsp70l:DsRed2(floxed)-EGFP*) (referred as *hs:Red2Green*) [111]; *Tg(0.6foxj1a:gfp)* (referred as *foxj1a:EGFP*) [112]; *Tg(fli1:CreER^{T2}; gCrystallin:EGFP; ubi:Switch*) (referred as *fli:Cre; ubi:Switch*).

All experiments and manipulations were approved by IMM internal ethics committee, in accordance with standards issued by Direcção-Geral de Alimentação e Veterinária, in order to minimize animal suffering and the number of fish and embryos used.

Generation of new transgenic lines

Plasmid cloning

To generate the construct pTol(*foxj1a:mCherry-T2a-CreER^{T2})* the promoter sequence of *D. rerio foxj1a* gene was first PCR amplified (primers: A1, A2) using the plasmid pTol(*foxj1a:EGFP*) [112] as template and adding recognition sites for ApaI and FseI restriction enzymes next to the 5' and 3' borders of the promoter, respectively. This sequence was then cloned into pTol(*her4.1:mCherry-T2a-CreER^{T2}*) [111] using blunt ligation after hydrolysis of the plasmid with ApaI and FseI to remove the *her4.1* promoter followed by Klenow fragment (Invitrogen)-mediated cohesive end filling.

To generate the construct pTol*(foxj1a:DsRed2(floxed)-EGFP)* the *EGFP* coding sequence was first PCR amplified (primers: A3, A4) using pTol(*foxj1a:EGFP*) as template and adding recognition sites for SmaI and AscI restriction enzymes next to the 5' and 3' borders of the amplicon, respectively. This fragment was ligated to pTol(*ef1α:DsRed2(floxed)-ntr-EGFP*) (a gift from Michael Brand) after hydrolysis of the plasmid with the same enzymes for removal of the (*ntr-EGFP)* ORF. The second part of the cloning procedure involved amplifying the promoter sequence of *D. rerio foxj1a* gene by PCR (primers: A5, A2) using the plasmid pTol(*foxj1a:EGFP*) as template and adding recognition sites for XhoI and FseI restriction enzymes next to the 5' and 3' borders of the promoter, respectively. This sequence was then switched for the *ef1α* promoter in the plasmid pTol(*ef1α:DsRed2(floxed)-EGFP*) using cohesive 5'/blunt 3' ligation after hydrolysis of the plasmid with FseI followed by Klenow fragment-mediated cohesive end filling and hydrolysis of both the plasmid and the promoter with XhoI.

To generate the construct pTol*(foxj1a:TK-GFP)* the promoter sequence of *D. rerio foxj1a* gene was first amplified by PCR (primers: A6, A7) using the plasmid pTol(*foxj1a:EGFP*) as template and adding recognition sites for HindIII and BamHI restriction enzymes next to the 5' and 3' borders of the

promoter, respectively. This sequence was then ligated to the plasmid pG1(*ins:TK-GFP*) [113] after hydrolysis with both restriction enzymes and removal of the *ins* promoter. The expression cassette (*foxj1a:TK-GFP)* was then amplified by PCR (first primer set: A5, A4; second primer set: A5, A8) from the newly made pG1(*foxj1a:TK-GFP*) adding restriction sites for XhoI and AscI restriction enzymes next to the 5' and 3' borders of the cassette, respectively. Plasmid pTol(*ef1α:DsRed2(floxed)-ntr-EGFP*) was hydrolysed with the same enzymes to remove its expression cassette and allow ligation of the (*foxj1a:TK-GFP)* cassette, thus inserting the desired construct in a Tol2-based vector.

To generate the construct pTol*(foxj1a:TK-T2a-EGFP)* the coding sequences for *TK*, T2a peptide and *EGFP* were first amplified and fused in frame using fusion PCR (Supp. Fig. S.4) [114, 115]. Briefly, each coding sequence was first individually amplified by PCR using primers containing 20nt complementary to the adjoining coding sequence in the final assembly (*TK*: primers A9, A10; template: pG1(*ins:TK-GFP*). T2a peptide: primers A11, A12; template: pTol(*her4.1:mCherry-T2a-CreER^{T2}*). *EGFP*: primers A13, A14; template: pTol(*foxj1a:EGFP*)). Then the sequences for *TK* and T2a peptide were fused in a second round of PCR (primers: A9, A12) and this product was finally fused to the *EGFP* sequence in the final PCR step (primers: A9, A14). Plasmid pTol(*foxj1a:EGFP*) was hydrolysed with BamHI and ClaI, for removal of *EGFP* coding sequence, prior to incubation with Klenow fragment for fill-in of 3' recessive ends; blunt ligation was promoted between the remaining part of pTol(*foxj1a:EGFP*) and intact PCR product *TK-T2a-EGFP*.

All expression cassettes sequences were confirmed by sequencing prior to microinjection: pTol(*foxj1a:mCherry-T2a-CreER^{T2}*), primers A1 and A2; pTol*(foxj1a:DsRed2(floxed)-EGFP*), primers A3 and A4 (*EGFP* insertion), A5 and A2 (promoter switch); pTol*(foxj1a:TK-GFP*), primers A6 and A7 (promoter switch), A4, A7 and B2 (expression cassette ligation); pTol*(foxj1a:TK-T2a-EGFP*), primers A6 and A3. pG1(*ins:TK-GFP*) was partially sequenced with primers B1 and B2.

All PCR reactions were conducted using Phusion DNA Polymerase (Thermo Scientific) according to the manufacturer's instructions; annealing temperatures were calculated using the online tool NEB Tm Calculator (available at http://tmcalculator.neb.com).

Restriction enzyme recognition sites, incubation temperatures and commercial sources are detailed in Appendix I. Primer sequences, special features (anything beyond template-complementary sequence) and melting temperature (Tm; obtained by the formula Tm=(#A + #T)*2 + (#G + #C)*4) are detailed in Appendix III.

Microinjection and selection of founders

The efficient Tol2 transposon system has been widely used in the zebrafish community to generate transgenics [107, 116] and was chosen to promote stable integration of the desired constructs in this work into the zebrafish genome. This system uses a transposase naturally found in the genome of the medaka fish (but not in *D. rerio*) [117] that recognizes two precise sequences at the borders of the construct [118] and transposes any DNA fragment up to 66kbp between those sequences to the host genome [119].

Embryos from natural, healthy (>70 eggs) [52] spawning from incross of the lines mentioned in Appendix II were injected at one-cell stage with ~1.4nL of a mixture containing plasmid DNA and Tol2 transposase mRNA in DNase/RNase-free water (Gibco) at the concentrations mentioned in Appendix II.

Briefly, a capillary-like glass needle was filled with up to 3µL of injection mixture, mounted on the microinjection apparatus and calibrated using a graticule [120, 121]; microinjection was performed by inserting the needle through the yolk and pumping one single shot of the mixture close to the egg cell.

Embryos were kept at a density of approximately 30 per Petri dish until 5 dpf and the medium was changed when necessary to avoid contamination. Injected embryos grown to adulthood were crossed to either AB or Tübingen (Tu) wild-type fishes to screen for the presence of the transgenes in their progeny and thus identify founder fishes and calculate germline transmission rates as the percentage of embryos from the founder fish that carry the desired construct.

Alternative screening techniques for *Tg(foxj1a:mCherry-T2a-CreER^{T2})*

Genomic PCR

24 hpf embryos from incross of potential founder fish (adult *Tg(hs:Red2Green)* fish that had been injected with (*foxj1a:mCherry-T2a-CreER^{T2}*) construct as zygotes) were collected for DNA extraction according to an established method [122]. Briefly, every 10 embryos were placed in microcentrifuge tubes, immersed in 100µL NaOH 50mM and incubated at 95ºC for 10min, after which they were vortexed. After cooling down to 4ºC, 10µL Tris-HCl pH 8.0 were added and the tubes were centrifuged for 6min at 10000rpm. The supernatant was used immediately for the PCR, as suggested [122]. Plasmid pTol(*her4.1:mCherry-T2a-CreER^{T2}*) was used as a positive control of the reaction. Amplification was conducted using DreamTaq DNA Polymerase (Fermentas), primers A15 and A17 and the following specific parameters: annealing temperature, 55ºC; extension time, 1min.

Whole-mount *in situ* hybridization (WISH)

24 hpf and 48 hpf embryos from incross of potential founder fish as well as AB and *Tg(fli:Cre; ubi:Switch)* lines were quickly dechorionated using 1mL of pronase 30mg.mL^{-1} (Sigma) and processed for WISH with methods adapted from *The Zebrafish Book* [123]. Significant protocol changes are: all wash steps made with PBST (PBS/0.5%Tween20); an additional permeabilization step with 2% H_2O_2 in 100% MeOH; proteinase K (Roche) digestion time of 10min and 15min (24 hpf and 48 hpf embryos, respectively); no use of acetic anhydride; pre-hybridization with Hyb$^+$ at 60ºC for at least 3h; probe denaturation in Hyb$^+$ solution for 5min at 80ºC; Fab-AP incubation at 4ºC ON; use of BM Purple (Roche) for color reaction and after it is stopped embryos are fixed in PFA 4% for 20min, washed in PBST and stored in 10% glycerol in PBST.

RNA probes used: *foxj1a* [124], diluted 1:100; *CreER^{T2}* (construct generated by Ana Pereira from plasmid pCAG-CreERT2 [125]), diluted 1:500.

Reverse Transcription PCR (RT-PCR)

Embryos from incross of potential founder fish and AB fish were harvested and kept in standard conditions. At 24 hpf RNA extraction was performed using TRIzol Reagent (Life Technologies) according to manufacturer's instructions (in this case: 1mL TRIzol for every 50 embryos), followed by incubation with RQ1 RNase-free DNase (Promega). Approximately 1µg RNA was used for the reverse transcription reaction, carried out with DyNAmo cDNA Synthesis Kit (Thermo Scientific) according to manufacturer's instructions and using random hexamer primers for amplification. The PCR step was

conducted with Phusion DNA Polymerase using A16 and A17 primers for *CreER^{T2}* transcript detection and A18 and A19 for *β-actin1* transcript detection as a positive control for the RT step. Specific reaction parameters were: annealing temperature, 69ºC; extension time: 26s.

Embryos from incross of *Tg(hs:mCh-Cre*) fish were harvested and kept in standard conditions. At 24hpf they were subjected to heat shock treatment (37ºC for 30min in embryo medium pre-heated to 42ºC) to promote expression of *CreER^{T2}*. RNA extraction occurred 1h30min, 2h and 3h after heat shock ended using the same procedure detailed above, followed by the RT-PCR reaction.

Agarose gel electrophoresis

Results from PCR and hydrolysis reactions were visualized by electrophoresis using 0.8%-1% agarose (Lonza) gels (for fragments over 400bp) or 2% gels (for fragments under 400bp) in TAE buffer. DNA staining was performed with RedSafe (iNtRON Biotechnology). In all agarose gel pictures the leftmost lane depicts 5uL of 1Kb Plus DNA Ladder (Invitrogen), with relevant fragment dimensions (in bp) mentioned to the side of the picture.

4-hydroxytamoxifen and heat shock treatment

4-hydroxytamoxifen (4-OHT, Sigma) was diluted to 10mM in absolute EtOH and single use aliquots were kept at -20ºC in the dark to prevent drug breakdown.

Double transgenic embryos from mating of *Tg(hs:mC-Cre*) and *Tg(hs:Red2Green*) fish were subjected at mid-gastrulation stage (shield to 75% epiboly) to heat shock according to published method [109] (37ºC for 30min in embryo medium pre-heated to 42ºC), after which they were returned to non-heated medium at the 28ºC incubator. At 24 hpf embryos were treated with 4μM, 8μM, 8μM (plus 14μg.mL^{-1} pronase) or 10μM 4-OHT and kept at 33ºC in glass Petri dishes in the dark. Six hours after 4-OHT addition embryos were subjected to a second identical heat shock to promote reporter expression. At 48 hpf embryos were observed and processed for immunochemical detection of EGFP (see below).

Double transgenic embryos from incross of *Tg(ubi:Switch)x(foxj1a:CreER^{T2}; cmlc2:EGFP*) founders were treated with 10μM 4-OHT and 14μg.mL^{-1} pronase at 24 hpf, and kept at 33ºC ON in glass Petri dishes in the dark.

Separately, larvae from the same mating were subject to spinal cord injury (SCI) at 5 dpf (see below) and immediately treated with 10μM 4-OHT and kept in glass Petri dishes in the dark at 33ºC for 4h30min, after which they were returned to 28ºC. 24h later a second dose of 4-OHT was added to the medium.

To facilitate infiltration of the drug into tissues, 1%DMSO (Sigma) was added at the same time as 4-OHT to all experiments including negative controls. To prevent embryo pigmentation and improve visualization, a 0.003% PTU (Sigma) solution in 10% Hank's saline was added to all embryos from 24 hpf onwards. In all experiments, negative control was performed with embryos from the same lines by replicating all conditions but replacing 4-OHT with absolute EtOH (8μL).

Lesion assay

5 dpf larvae were anesthetized in 1x Tricaine solution diluted in embryo medium and remained there during the procedure. When a larva no longer showed touch response, it was placed laterally on the floor of the Petri dish and the tip of a 20G syringe needle was used to make an incision through the dorsal lateral face at the level of the anal pore [126], transecting the whole width of the spinal cord but not cutting beyond the ventral edge of the notochord. Injured larvae were transferred to dishes with fresh embryo medium and checked daily to remove dead larvae.

Tricaine solution (25x): 4g.L^{-1} Tricaine S (Western Chemical, Inc), 20mM Tris (pH 9), adjust to pH 7.

Immunohistochemistry

48 hpf embryos were dechorionated with 1mL pronase 30mg.mL^{-1} and fixed in PFA 4% for 4h at RT followed by one wash step in PBST, one wash step in PBST/0.5% TritonX100/100mM glycine for 1h and one wash step in PBST for 10min. Embryos were incubated in blocking buffer (PBST/ 1%BSA/ 0.5%TritonX100/ 1%DMSO) for 3h and then incubated with primary antibody at 4ºC ON. The second day started with four washes of 30min in blocking buffer followed by incubation with secondary antibody at 4ºC in the dark ON. Embryos were then washed four times for 30min in PBST, fixed in PFA 4% for 30min and stored in PBST for observation until the following day. Primary antibody: IgY chicken anti-GFP (Aves Labs, GFP-1020), diluted 1:500 in blocking buffer. Secondary antibody: IgG goat anti-chicken AlexaFluor 488 (Invitrogen, A11039), diluted 1:1000 in blocking buffer.

Image acquisition and processing

Fluorescence and stereo microscopy: All darkfield and fluorescence images were acquired with Zeiss AxioZoom V16 fluorescence stereo microscope equipped with AxioCam MRm, Zeiss filter sets FS38HE (for EGFP and mmGFP signal) and FS63HE (for mCherry and DsRed2 signal) and manufacturer's software ZEN 2012 Blue Edition. All brightfield images were acquired with Leica Z6 APO macroscope equipped with Leica DFC490 camera, and software Adobe Photoshop CS3.

Image processing (cropping, rotating, pseudocolouring and exposure correction for gel pictures) was conducted in ImageJ and Adobe Photoshop CS2.

Bioinformatics and statistical analysis

Full sequences of pTol(*her4.1:mCherry-T2a-CreER^{T2}*) and pTol(*ef1α:DsRed2(floxed)-ntr-EGFP*) and partial sequence of pTol(*foxj1a:EGFP*) (including the complete expression cassette) were already available at the lab. *TK*, *mGFP5* and *D. rerio β-actin1* coding sequences were retrieved from NCBI database (Accession: *HSV1 Thymidine Kinase*, V00470; synthetic construct *mGFP5*, U87973; *D. rerio actb1*, BC165331). *mmGFP* sequence was constructed from *mGFP5* by replicating the method described in [127]. Sequence manipulation and analysis was performed with ApE software. DNA and protein alignments were obtained with T-Coffee [128] by combining the output of several common aligners (M-Coffee package) for pairwise methods and multiple methods, respectively.

Codon usage data and comparisons

Codon usage (CU) data for *TK*, *mCherry*, *DsRed2*, *EGFP*, *mmGFP* and *actb1* was obtained with Sequence Manipulation Suite [129] (available online at http://www.bioinformatics.org/sms2/codon_usage.html). CU data for *Danio rerio* was obtained from the Codon Usage Database [130] (available online at http://www.kazusa.or.jp/codon/). Synonymous CU preference for each of the aforementioned genes was compared with *D. rerio* CU with a two-tailed Fisher's Exact Test for 2xn contingency tables, computed in MATLAB software (MathWorks) using *myfisher* function [131]; this function returns an exact p-value for 2x2, 2x3 or 2x4 tables and a Monte Carlo approximation for larger ones. The absolute frequency of each codon in each of the two coding sequences being compared at a time (one of the genes and *D. rerio* overall CU) was listed in a table, grouping together synonymous codons for each amino acid (except Met and Trp, for which no synonymous codons are available). This created 18 contingency tables (one for each amino acid) with 2 lines (one for each sequence) and *n* rows (one for each possible codon). For example, all tables for Ala are 2x4 tables given that either GCG, GCA, GCT and GCC can code for alanine:

Alanine:	**GCG**	**GCA**	**GCT**	**GCC**	**p-value**
D. rerio*	13	25	32	30	**6E-6**
TK	12	0	8	26	

*-Given the higher difficulty in computing this test's results with large observed values, the values used for *D. rerio* observed codon frequency correspond to: (absolute count for that codon)/(sum of counts for all synonymous codons on that aa.)*100.

Ultimately, a p-value was obtained for each amino acid, and the amount of aa. with no statistically significant discrepancy (α=0.05) between the two compared sequences was taken as a measure of the CU similarity between those sequences, designated as codon usage preference score (CUPS).

Relative codon usage for *D. rerio* and for each of the aforementioned genes was calculated, for each codon, as: (absolute count for that codon)/(sum of counts for all synonymous codons on that aa.).

Data presentation

For figures 1.2, 1.4, 1.5 and 1.6 survival rates of injected embryos were normalized to survival rates of non-injected embryos from the same batch; that is, non-injected survival was considered 100% and injected survival rate was adjusted proportionally. In the rare occasions when non-injected embryos were not kept until 24hpf, injected survival rates were normalized to the average survival of all non-injected embryos of the same line. Percentage of positive embryos was calculated relative to the number of injected embryos alive at 24hpf. Survival and positive rates were calculated using Microsoft Excel. All plots were generated with Prism 6 software (GraphPad).

Average values mentioned throughout the text are represented as average±SD.

Results

1. Generation of zebrafish transgenic lines

Given all the recent findings concerning the potential ability of spinal cord ependymal cells to aid in the regeneration process after a spinal cord injury in mammals, it was necessary to understand exactly how important these cells are during the same process in zebrafish. To that end, two major approaches can be taken: promoting genetic labelling of *foxj1a*$^+$ cells before a spinal cord injury (SCI) in order to identify their progeny during the regeneration process (lineage tracing experiment), and inducing specific ablation of ependymal cells to uncover their functional role after SCI. This chapter details the attempts at generating several zebrafish transgenic lines that could be used to investigate the role of *foxj1a*$^+$ cells in spinal cord regeneration using both approaches mentioned.

1.1 *Tg(foxj1a:CreER^{T2}; cmlc2:EGFP)*

Specific labelling of *foxj1a*$^+$ cells can be achieved by inserting a transgene for a fluorescent protein under the control of the *foxj1a* promoter. However, a lineage tracing experiment requires this labelling to be performed at a specific time (in this case, right before a spinal cord injury) and to persist not only in these cells but also in their progeny, regardless of their *foxj1a* expression status. Expressing the *CreER^{T2}* coding sequence under the control of the *foxj1a* promoter ensures recombination will occur only in *foxj1a*$^+$ cells, while a ubiquitously expressed reporter protein will be able to maintain fluorescent labelling even if those cells stop expressing this transcription factor. This will ensure that the entire *foxj1a*$^+$ cell progeny will remain permanently labelled. The use of the modified version of Cre, CreER^{T2} [106], inserts a step of temporal control in the expression of a reporter protein, so that labelling of *foxj1a*$^+$ cells is only achieved when fish are treated with 4-hydroxytamoxifen (4-OHT), thus promoting CreER^{T2}-mediated recombination in the target cells at the desired time point.

In order to establish a transgenic line that would allow for efficient labelling of *foxj1a*$^+$ cells, a first attempt was made at recreating a line previously used in the lab: *Tg(ubi:Switch)x(foxj1a:CreER^{T2}; cmlc2:EGFP*). Due to the presence of the (*ubi:Switch*) cassette these fish show a ubiquitous expression of EGFP that, upon action of a Cre recombinase, will be exchanged for mCherry expression, thus labelling *foxj1a*$^+$ Cre$^+$ cells in red. Double transgenics are easily identified by their green heart due to the expression of EGFP under the control of the cardiac muscle promoter *cmlc2*.

Since both adult *Tg(ubi:Switch*) fish and the plasmid containing the (*foxj1a:CreER^{T2}; cmlc2:EGFP*) cassette flanked by Tol2 recognition sites were already available, the plasmid was microinjected along with Tol2 transposase mRNA into one-cell stage *Tg(ubi:Switch*) embryos according to standard techniques. At 24 hpf, injected embryos were screened for the presence of the (*foxj1a:CreER^{T2}; cmlc2:EGFP*) construct and positive embryos were selected based on EGFP expression in the heart (Figure 1.1 A, B). In a single injection event using 30ng.µL^{-1} DNA and 5ng.µL^{-1} mRNA, 121 (24%) positive embryos were selected and grown; of these, 57.9% survived to adulthood.

To look for germline transmission of the injected construct, all adult fish that had been injected at the zygote stage (considered as potential founders) were crossed to wildtypes and their progeny analysed for the presence of (*foxj1a:CreER^{T2}; cmlc2:EGFP*) cassette by EGFP fluorescence in the heart. Adult fish were considered founders when they showed the capacity to pass the injected transgenes to their offspring, despite any other factor like being hemizygous for the (*ubi:Switch*) cassette (which would

mean that only half the (*foxj1a:CreER^{T2}*)$^{+}$ progeny would also have the reporter transgene). In practice, this means that any potential founders that spawned at least one embryo with a green heart were considered effective founders. Of the 70 fish that survived to adulthood, only 11 (15.7%) were found to pass the desired transgenes to their offspring (Fig. 1.1 C), with the vast majority considered non-founders. Furthermore, more than half of the founders showed very low levels of germline transmission (as low as 2.0%), passing the injected construct to less than 10% of their progeny. The highest rate of germline transmission was found in one male fish (33.3%) and one female (18.0%), which were used to obtain embryos for the CreER^{T2} activation assay (see Chapter 2.2, Results section). The negative results obtained in this assay suggest that this double transgenic line is not suitable for the intended lineage tracing experiment; for that reason, founders with the highest germline transmission rate were maintained but no further steps were taken to establish a stable transgenic line.

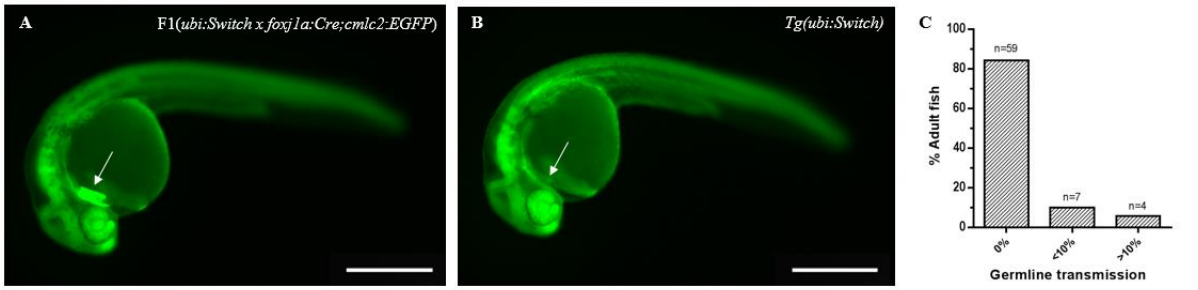


Figure 1.1 Screening of *Tg(ubi:Switch)* embryos injected with (*foxj1a:CreER^{T2}; cmlc2:EGFP*) construct and selection of founder fish. EGFP expression is clearly visible in the heart (arrow) of embryos possessing the desired construct containing *CreER^{T2}* (A), distinguishing them from (*ubi:Switch*)$^{+}$ embryos that did not inherit the injected construct (B). Scale bar: 500μm. **C:** Most larvae grown to adulthood did not pass the construct to the germline; the few founder fish present low germline transmission rates.

1.2 *Tg(foxj1a:mCherry-T2a-CreER^{T2})*

Given that the double transgenic line *Tg(ubi:Switch)x(foxj1a:CreER^{T2}; cmlc2:EGFP*) had been previously used with little success, a new strategy was devised for the time-specific labelling of *foxj1a*$^{+}$ cells. Adding to the features of the previous transgenic line, two major improvements were included. A heat shock inducible promoter was used to drive the expression of the reporter cassette (*Tg(hs:Red2Green)*), which could increase the level of protein production and lead to better visualization of labelled cells [132] while at the same time adding an additional step of temporal control over the cell labelling process. The detection of CreER^{T2} expression in the right place and time was also facilitated by coupling *mCherry* to *CreER^{T2}* coding sequences in frame, separated by the sequence for the viral self-cleaving peptide T2a [133], thus leading to co-expression of both proteins without potential drawbacks of fusion proteins (like partial loss of function) [134]. Transcription of both proteins from the same open reading frame (ORF) also prevents the possibility of transcriptional silencing being exerted over one of two closely placed expression cassettes, as in the case of the (*foxj1a:CreER^{T2}; cmlc2:EGFP*) construct.

To generate this line it was necessary to obtain first a suitable Tol2-based plasmid with the (*mCherry-T2a-CreER^{T2}*) ORF under the control of the *foxj1a* promoter. This was accomplished by switching the *her4.1* promoter in the pTol(*her4.1:mCherry-T2a-CreER^{T2}*) plasmid for the *foxj1a* promoter using molecular cloning methods. The cloning procedure is depicted in Figure 1.2A. Positive ligation clones

were partially sequenced (Supp. Fig. S.1 A), confirming *foxj1a* promoter integrity and the success of ligation necessary for effective transgene expression.

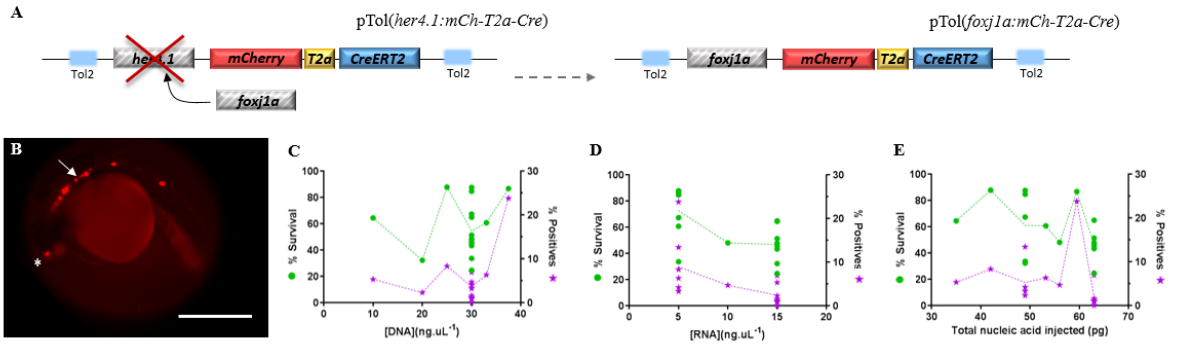


Figure 1.2 Molecular cloning scheme of pTol(*foxj1a:mCherry-T2a-CreER^{T2}*) plasmid, screening of injected *Tg(hs:Red2Green)* embryos and microinjection statistics. **A:** The *her4.1* promoter was switched for the *foxj1a* promoter, generating the final plasmid. Grey boxes represent promoter regions. **B:** At 24 hpf, embryos containing the injected construct show mCherry expression in pronephros (not shown here), floor plate (arrow) and olfactory bulb (*) cells Scale bar: 500µm. **C, D, E:** Survival rate (normalized to non-injected survival – see M&M section for details) and percentage of positive embryos is heterogeneous and does not follow a clear trend when associated with either DNA concentration or total nucleic acid injected. Each dot/star represents one of the n=16 independent batches of injected embryos. Dashed lines connect average Y-axis values for each X-axis category. DNA and mRNA concentrations detailed in M&M section.

Ligation clone #3 was selected for microinjection into one-cell stage *Tg(hs:Red2Green)* embryos along with Tol2 transposase mRNA according to standard techniques. Given that the reporter cassette is only expressed upon heat shock treatment, embryos are easily screened for the presence of (*foxj1a:mCherry-T2a-CreER^{T2}*) cassette by mCherry expression on *foxj1a*$^{+}$ cells. At 24 hpf, positive injected embryos were identified by red fluorescence in the floor plate of the neural tube, pronephros and olfactory bulb cells [104] (Fig. 1.2 B).

Common practices in zebrafish husbandry suggest that a minimum of 25-30 5 dpf larvae must be moved at once to grow on regular tanks due to higher likelihood of larvae death occurring in the first two weeks. Additionally, since Tol2-based transgenesis is known to originate variable rates of transgene integration and germline transmission [116], it is advisable that no less than 100-120 larvae are grown to adulthood. For these reasons, and because first attempts at injecting with lower DNA and mRNA concentrations yielded very few positive embryos, an optimization approach was implemented. A total of 8 combinations of DNA and mRNA concentrations were injected in order to find the injection mix that resulted in the higher yield of positive embryos without compromising embryo health. Figures 1.2 C-E relate survival rate of the injected embryos at 24 hpf (when screening for positives was conducted) and percentage of positives found in each of the 16 independent injection events to concentration and total amount of DNA and mRNA used. At first glance, increasing mRNA concentration seems to exert a negative influence on both variables, despite the large heterogeneity of values for the same mRNA concentration. A greater influence of mRNA concentration may explain the absence of any clear trend relating embryo survival and transgene integration with either DNA concentration and total nucleic acid. The lowest values for each variable are obtained with the highest amount of total nucleic acid, confirming that high concentrations of DNA and mRNA become toxic to the embryos.

Nevertheless, the rate of transgene integration was always under 15%, except for one injection event. Given that each day an average of 658 ± 237 embryos were injected, and absolute number of positives ranged from 2 (0.9%) to 63 (13.4%), only three injection events yielded enough positive embryos to be

grown: 86 with 5ng.μL^{-1} mRNA and 33 with 15ng.μL^{-1} mRNA (both with 30ng.μL^{-1} DNA). Embryos injected with 37.5ng.μL^{-1} DNA could not be grown because most of them presented severe developmental anomalies (like neural tube defects) probably due to toxic effects of DNA concentration, causing premature death before 5 dpf.

A total of 68 larvae survived to adulthood and were crossed to wildtype fish for observation of progeny and assessment of transgene germline transmission. Eight (12%) potential founders generated a very small fraction (4.6±1.7%) of embryos with red fluorescent labelling of a few cells that seemed to localize to the floor plate and pronephros at 24 hpf (similar to Fig. 1.6 C). However, this signal had disappeared completely by 28 hpf in some embryos and by 48 hpf in all the remaining ones. Following reported findings that a (*mCherry-T2a-[protein of interest])* cassette may result in protein of interest expression without mCherry fluorescence [135], it was hypothesized that a similar phenomenon could be happening with these embryos. Since functionality of (*mCherry-T2a-CreER^{T2}*) cassette for lineage tracing is bound to CreER^{T2} expression and function, and not to mCherry signal, *mCherry*$^{-}$ embryos that prove positive for CreER^{T2} are still useful. This pointed to the need of finding a different method to investigate the presence of the *CreER^{T2}* transgene in these embryos.

The first approach taken was to extract genomic DNA from F1 embryos of two incrosses of potential founders (Couple #1 and Couple #2) and perform PCR with *CreER^{T2}* specific primers that generated a 943bp amplicon located in the second half of the coding sequence. Results reveal that DNA extracted from embryos of both crosses generated a fragment matching in size to the predicted amplicon and to the positive control (a plasmid containing the *CreER^{T2}* coding sequence) (Fig. 1.3 A), suggesting these embryos carry the *CreER^{T2}* transgene. Several other fragments of lower molecular weight, consistent between lanes, are also visible, indicating that there was unforeseen amplification from other zebrafish sequences. This is not surprising given that a BLAST search of both primers used returned partial sequence similarity with some regions of the zebrafish genome (data not shown). However, the presence of visible faint bands for the negative control suggests template contamination may have occurred, prompting these results to be confirmed.

While looking for transgene presence in the genome can give a measure of transgenesis efficiency, the ultimate goal is to have the CreER^{T2} protein expressed and functional. One way to assess gene expression is to look for the presence of gene-specific transcripts; therefore, an *in situ* hybridization assay was performed to assess *CreER^{T2}* transcription at 24 hpf and 48 hpf using embryos from the incross of Couple #1. Staining for *CreER^{T2}* probe was found to span the entire length of floor plate and pronephros of 3 (12.5%) 24 hpf embryos (Fig. 1.3 B), matching the expression pattern of *foxj1a* gene, in which all AB embryos were found to express *foxj1a* (Fig. 1.3 E). However, the majority of 24 hpf embryos (Fig. 1.3 C), as well as all 48 hpf embryos (Fig. 1.3 D), did not show a specific *foxj1a* expression pattern, similarly to AB embryos tested for *CreER^{T2}* probe (Fig. 1.3 F, G), suggesting that *CreER^{T2}* was not being expressed. Embryos from incross of *Tg(fli:Cre; ubi:Switch*) fish were also assayed for *CreER^{T2}* transcription and intended to act as positive control for the probe. However, no specific staining was obtained either in 24 hpf or 48 hpf embryos (Fig. 1.3 H and I); these showed only unspecific staining in the general head area, very different from the expected expression pattern in the blood vessels typical of the *fli1* promoter. This result may have two readouts: either there is a problem with CreER^{T2} expression in the *Tg(fli:Cre; ubi:Switch*) line (and this would be consistent with the absence of recombination in a CreER^{T2} activation assay (data not shown)) or the *CreER^{T2}* probe was not working. It is then necessary to test the probe with another *CreER^{T2}*-expressing line and to further validate the results from embryos of potential founder fish.

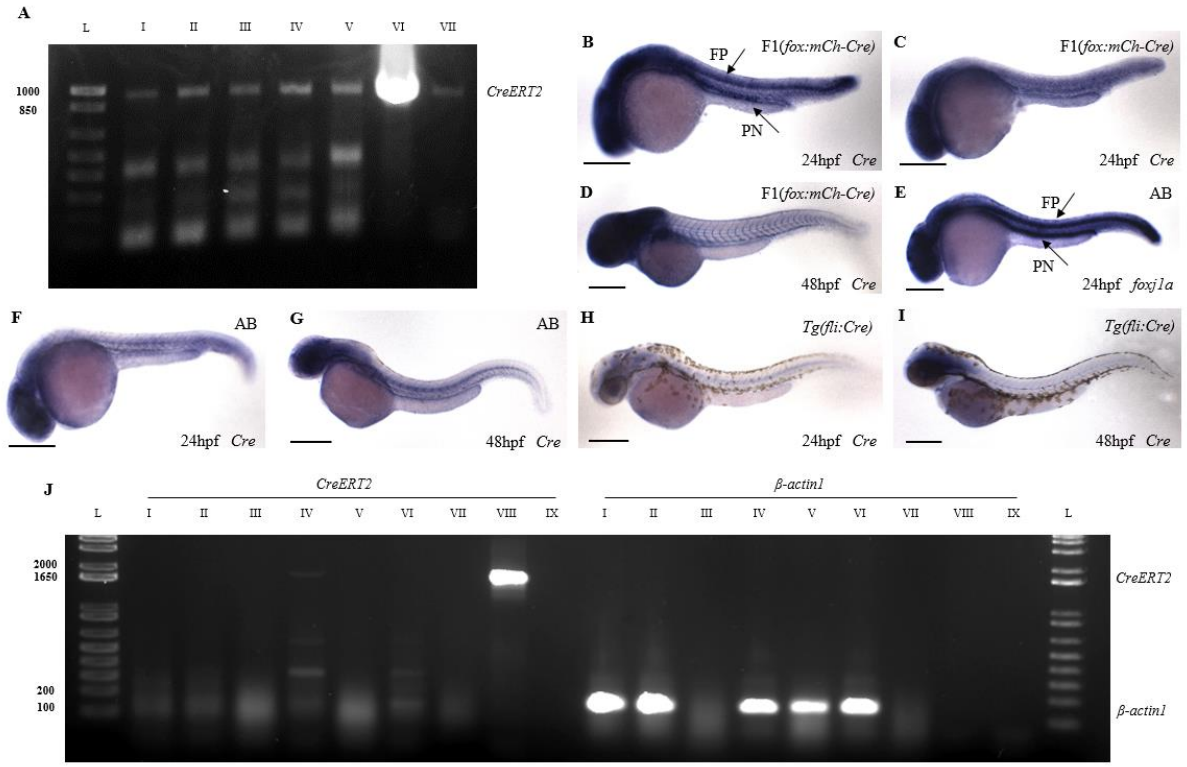


Figure 1.3 Alternative molecular methods for detection of founder fish carrying the (*foxj1a:mCherry-T2a-CreER^{T2}*) construct. **A:** Genomic PCR detected the presence of a DNA fragment matching predicted *CreER^{T2}* amplicon size (943bp) in progeny from two couples of potential founders (Couple 1: I-IV; couple 2: V). Lane VI: pTol(*her4.1:mCherry-T2a-CreER^{T2}*) as template for positive control. Lane VII: no DNA template for negative control. DNA Ladder (L) as molecular weight reference. **B-I:** WISH results identify some 24 hpf embryos from potential founders as positive (B) for *CreER^{T2}* expression, with the majority being negative (C). At 48 hpf, all embryos from the same batch stained negative (D), as well as AB (F, 24 hpf; G, 48 hpf) and *Tg(fli:Cre)* (H, 24 hpf; I, 48 hpf) embryos. 24 hpf AB embryos stained positive for *foxj1a* probe (E), validating the protocol. FP: floor plate; PN: pronephros. n=24(B,C), 28(D), 13(E), 14(F), 22(G), 28(H), 22(I) embryos. Scale bar: 300µm. **J:** RT-PCR reveals *CreER^{T2}* transcript absent from progeny of potential founders (II) and AB fish (I) but present 1h30min (IV) after heat shock in *Tg(hs:mCh-Cre)* embryos. Embryos collected 2h (V) and 3h (VI) after heat shock do not show trace of *CreER^{T2}* mRNA. Lanes III and VII: same as lanes II and IV (respectively) but without addition of reverse transcriptase enzyme. Lanes VIII: pTol(*her4.1:mCherry-T2a-CreER^{T2}*). Lanes IX: no template. *β-actin1* amplification validates RT-PCR protocol. DNA Ladder (L) as molecular weight reference.

In order to determine whether the *CreER^{T2}* transcript is present in the progeny of potential founders for (*foxj1a:mCherry-T2a-CreER^{T2}*) cassette, total RNA was extracted from 24 hpf embryos from incross of Couple #1 (see above) and RT-PCR was performed using *CreER^{T2}* specific primers (Fig. 1.3 J). This time there was no detection of the desired transcript in the test embryos, contrary to *Tg(hs:mCh-Cre*) embryos, acting as positive control for the method. In these, the 1758bp fragment is clearly visible for embryos collected 1h30min after heat shock (necessary to activate transcription from the *hsp70l* promoter). Surprisingly, embryos collected 2h and 3h after heat shock do not reveal the presence of *CreER^{T2}* mRNA, suggesting it may have already been degraded. It is worth to mention as a reference that mCherry fluorescence becomes visible from 2h30min after heat shock (data not shown). Given that *foxj1a^{+}* cells constitute a very small fraction of the embryo, it is possible that *mCherry-T2a-CreER^{T2}* mRNA existed in concentrations too small for PCR detection, even considering that the RT reaction was performed with pooled RNA from 60 embryos. This would not have been a problem for *Tg(hs:mCh-Cre*) embryos since the *hsp70l* promoter is expressed throughout the embryo upon heat shock. Conversely, this could also mean these embryos were negative for the transgenic construct, calling into question the results from previous screening approaches.

Overall, these three screening methods seem to indicate that genomic integration of the (*foxj1a:mCherry-T2a-CreER*^{T2}) cassette occurred in some of the injected embryos and was passed to their germline, generating some F1 embryos with the desired transgene. However, it was not possible to establish a stable transgenic line due to the low transmission efficiency and the absence of straightforward screening methods.

1.3 *Tg(foxj1a:DsRed2(floxed)-EGFP)*

In order to maximize the probability of having at least one functional transgenic line to perform the lineage tracing experiment, a third strategy was planned to generate a line reciprocal to the one presented in Chapter 1.2 (Results section). This new line would then use the same coding sequences for the *Cre* driver construct and for the reporter construct, but each would be driven by the promoter sequence assigned to the other construct in 1.2. In practice, the goal was to generate the double transgenic line *Tg(hs:mCh-Cre)x(foxj1a:DsRed2(floxed)-EGFP*), with the CreER^{T2} enzyme being expressed in the whole embryo upon heat shock and *foxj1a*$^{+}$ cells changing from a DsRed2^{+} to a EGFP^{+} phenotype after recombination.

To generate this transgenic line it was necessary to assemble a Tol2-based plasmid with the (*foxj1a:DsRed2(floxed)-EGFP*) cassette. This was accomplished by first replacing the *ntr-EGFP* coding sequence in plasmid pTol(*ef1α:DsRed2(floxed)-ntr-EGFP*) for the *EGFP* coding sequence obtained from PCR amplification. The resulting pTol(*ef1α:DsRed2(floxed)-EGFP*) had the *ef1α* promoter switched for the *foxj1a* promoter obtained from PCR amplification, originating the final pTol(*foxj1a:DsRed2(floxed)-EGFP*). Figure 1.4 A depicts main events in the cloning process. Positive ligation clones were partially sequenced (Supp. Fig. S.1 B and C), confirming both ligation events were successful.

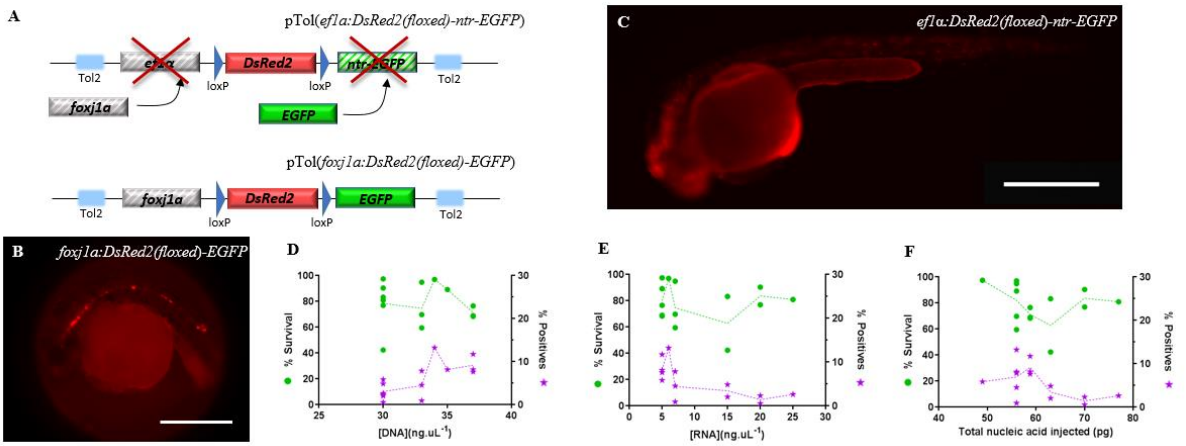


Figure 1.4 Molecular cloning scheme of pTol(*foxj1a:DsRed2(floxed)-EGFP*) plasmid, screening of injected *Tg(hs:mCh-Cre*) embryos and microinjection statistics. **A:** The *ef1α* promoter was switched for the *foxj1a* promoter and the *ntr-EGFP* ORF was switched for the *EGFP* ORF, generating the final plasmid. Grey boxes represent promoter regions. **B:** At 24 hpf, embryos containing the injected construct show DsRed2 expression in pronephros, olfactory bulb (not shown here) and floor plate cells. Scale bar: 500µm. **C:** Embryos injected with (*ef1α:DsRed2(floxed)-ntr-EGFP*) construct show overall DsRed2 expression at 24 hpf. Scale bar: 500µm. **D, E, F:** Survival rate (normalized to non-injected survival – see M&M section for details) and percentage of positive embryos is heterogeneous and does not follow a clear trend when associated with either DNA or RNA concentration or total nucleic acid injected. Each dot/star represents one of the n=14 independent batches of injected embryos. Dashed lines connect average Y-axis values for each X-axis category. DNA and mRNA concentrations detailed in M&M section.

Ligation clone #7 was selected for microinjection into one-cell stage *Tg(hs:mCh-Cre*) embryos along with Tol2 transposase mRNA according to standard techniques. Embryos were screened for the presence of the (*foxj1a:DsRed2(floxed)-EGFP*) cassette at 24 hpf by looking for DsRed2 fluorescence in *foxj1a*$^{+}$ cells, which could be found in the neural tube floor plate (Fig. 1.4 B), pronephros and olfactory bulbs by that developmental stage.

Injection of this construct presented similar problems to the ones found with pTol(*foxj1a:mCherry-T2a-CreER*^{T2}): the number of positive embryos was seldom enough to be worth growing and in many cases embryos showed only a very small number of DsRed^{+} cells. An identical optimization approach was undertaken, using 8 combinations of DNA and mRNA concentrations in order to achieve a better yield of positive embryos without the toxic effects of high nucleic acid concentration. Figures 1.4 D-F relate survival rate of the injected embryos at 24 hpf (when screening for positives was conducted) and percentage of positives found in each of the 14 independent injection events to DNA and mRNA concentration used, as well as to the total amount of nucleic acid for each injection event. It is clear at first that there is no monotonic trend determining embryo survival or transgene integration rate based on nucleic acid concentration. Nevertheless, increasing DNA concentrations and decreasing mRNA concentrations seem to associate with a higher yield of positives, while an increase of both concentrations appears to have a negative impact in transgene integration. Survival rates are generally high and heterogeneous but still appear to follow a similar trend to the rate of positives. The general low rate of positives (always under 15% in an average of 632 ± 206 injected embryos per event) made it difficult to gather in one day at least 30 embryos to grow. The peak on both variables observed in Figure 1.4 D (34ng.μL^{-1}) corresponds to one of the 4 injection events yielding enough positives; others were obtained with 33ng-μL^{-1} and 37ng.μL^{-1}, to a total of 116 DsRed2^{+} embryos.

It was postulated during the optimization process that lack of DsRed2 signal could be due to the fact that the protein emission spectrum is only partially captured by the stereo microscope emission filters. The original plasmid pTol(*ef1α:DsRed2(floxed)-ntr-EGFP*) was injected in embryos from the same line, using 30ng.μL^{-1} DNA and 10ng.μL^{-1} mRNA, to assess if DsRed2 signal could be easily seen with the same filters. At 24 hpf, red fluorescence could be seen throughout the whole embryo (Fig. 1.4 C) in 65.7% of injected embryos, not only validating this emission filter as ideal for DsRed2 visualization but also confirming functionality of the Tol2 transposition sites in the vector backbone.

In order to determine whether any of the 30 (25.9%) fish that survived to adulthood passed the transgenic cassette to the germline and could then be used as a line founder, each of them was crossed to wildtype fish and their progeny was observed at 24 hpf. All potential founders were considered negative for germline transmission given that DsRed2 fluorescence could not be observed in any of the progeny embryos. To ensure that even small rates of germline integration would be detected, only clutches of at least 100 embryos were analysed. It was considered that any founder fish with a germline transmission rate lower than 1% would not generate in one spawning enough embryos for growing or for lineage tracing experiments.

Contrary to what was done with (*foxj1a:mCherry-T2a-CreER*^{T2}) potential founders, no further screening strategies were attempted for this line given that DsRed2 fluorescence is essential for the lineage tracing experiments, being the reporter protein that signals absence of CreER^{T2}-driven recombination.

1.4 Lines for genetic ablation of *foxj1a*$^{+}$ cells

In biology, one classical experimental design used to discover whether one element of a given system has an important role, and what that role may be, is to deprive the system of that one element and

measure how the system is affected by that loss. This means that to study the functional role of ependymal cells in zebrafish spinal cord (SC) regeneration, it is essential to remove these cells from the central canal of a regenerating SC. A genetic ablation strategy was employed because it allows specific targeting for ependymal cells (with the use of the *foxj1a* promoter) and temporal control over the cell ablation process, which is triggered by delivering a specific drug to the organism. The choice of the HSV Thymidine Kinase enzyme was due to the fact that this enzyme is only active in proliferating cells, thus ablating only those that respond to injury while maintaining the integrity of the rest of the central canal [105].

1.4.1 *Tg(foxj1a:TK-GFP)*

The generation of a transgenic line expressing the *thymidine kinase (TK)* gene under the control of the *foxj1a* promoter depended on the insertion of *TK* coding sequence into a Tol2-based plasmid (Fig. 1.5 A). Using molecular cloning methods and plasmid pG1(*ins:TK-GFP*), the *ins* promoter was first switched for the *foxj1a* promoter that had been PCR amplified, generating pG1(*foxj1a:TK-GFP*). PCR amplification of the expression cassette (*foxj1a:TK-GFP*) was required to insert it into a Tol2-based plasmid; several optimization attempts were made using a forward primer for the *foxj1a* promoter and a reverse primer for *EGFP*, but unspecific amplification could still not be prevented. Despite this setback, the fragment of expected size was inserted into the backbone of pTol(*ef1α:DsRed2(floxed)-ntr-EGFP*) and partially sequenced to confirm ligation success.

Sequencing results revealed that the *GFP* coding sequence present in this construct did not match the *EGFP* coding sequence present in other plasmids used before, which could explain all the previous problems in amplifying this construct. In the Methods section of the publication where the original pG1(*ins:TK-GFP*) was used [113], it is mentioned the use of *mmGFP* coding sequence to create the fusion protein TK-GFP. Since a search in the NCBI nucleotide database did not return any match to this particular gene designation, the coding sequence of *mmGFP* had to be recreated from *mGFP5* sequence using the method described in [127]. The resulting sequence proved an exact match to the partial sequence of pG1(*ins:TK-GFP*) that had since been obtained by sequencing, confirming *mmGFP* as the gene present in this construct. Alignment of *EGFP* and *mmGFP* coding sequences reveals major sequence mismatches (data not shown), mostly in the third position of codons, which results in small differences in aminoacid composition between the proteins (Supp. Fig. S.2) but explains why a *EGFP* primer did not effectively amplify the *mmGFP* sequence. For this reason, a reverse primer was designed specifically for the 3' end of *mmGFP* and used to repeat PCR amplification of (*foxj1a:TK-GFP*) cassette, leading to a second ligation event to the backbone of pTol(*ef1α:DsRed2(floxed)-ntr-EGFP*).

Partial sequencing of ligation clone #6 confirmed a successful ligation and the absence of PCR-introduced mutations, enabling the use of this plasmid for microinjection. Wildtype embryos at one-cell stage were used for injection of pTol(*foxj1a:TK-GFP*) along with Tol2 transposase mRNA and embryos were screened at 24 hpf for transgene integration. Positive embryos showed mmGFP expression in a few scattered cells matching the known localization of *foxj1a*$^{+}$ cells, like pronephros and floor plate cells (Fig. 1.5 B). Despite optimization attempts with 5 combinations of DNA and mRNA concentrations, the average amount of mmGFP^{+} cells per embryo remained unchanged. Relative yield of positive embryos was also low in all 13 injection events (Fig. 1.5 C-E), never reaching more than 10% in an average of 700 ± 243 injected embryos per day.

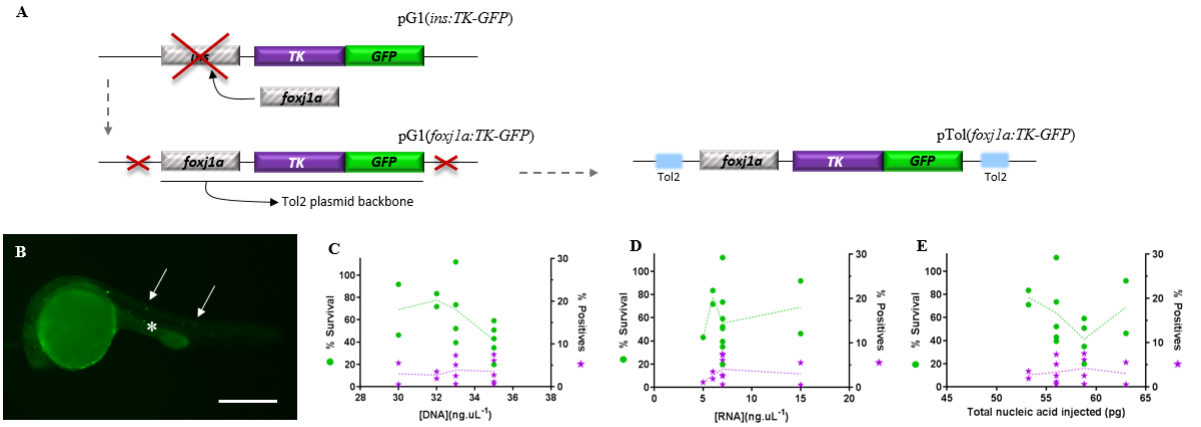


Figure 1.5 Molecular cloning scheme of pTol(*foxj1a:TK-GFP*) plasmid, screening of injected embryos and microinjection statistics. **A:** The *ins* promoter was switched for the *foxj1a* promoter and the whole expression cassette was cloned into a Tol2-based plasmid backbone, generating the final plasmid. Grey boxes represent promoter regions. **B:** At 24 hpf, embryos containing the injected construct show mmGFP expression in olfactory bulb (not shown here), pronephros (*) and floor plate (arrow) cells. Scale bar: 500µm. **C, D, E:** Survival rate (normalized to non-injected survival – see Materials and Methods section for details) and percentage of positive embryos is heterogeneous and does not follow a clear trend when associated with either DNA or RNA concentration or total nucleic acid injected. Each dot/star represents one of the n=13 independent batches of injected embryos. Dashed lines connect average Y-axis values for each X-axis category. DNA and mRNA concentrations detailed in M&M section.

As mentioned in Chapter 1.3 (Results section), injection of pTol(*ef1α:DsRed2(floxed)-ntr-EGFP*) resulted in a high yield of positive embryos with strong DsRed2 signal. Given that this plasmid's backbone was used to insert the construct (*foxj1a:TK-GFP*), it seemed safe to assume that defective Tol2 sites in pTol(*foxj1a:TK-GFP*) were not responsible for the observed low levels of mmGFP expression. Attention was then turned to other sequence features that could interfere with gene expression levels [136], namely the Kozak sequence, associated with the targeting of the small ribosomal unit to the correct translational start site in a mRNA molecule, and the polyadenylation (poly(A)) signal, responsible for 3' cleavage of mRNA and addition of poly(A) tail preventing premature mRNA degradation [137]. Sequencing data from ligation clone #6 revealed the absence of any previously described Kozak sequence (Figure S.3 A), including original Kozak and the natural *D. rerio* variant. Moreover, of the two key nucleotides conserved across vertebrate groups, only the G in position +4 is found at the start of *TK* coding sequence (with the other being a purine in position -3 [138]). When comparing the sequence downstream of the *mmGFP* stop codon with available sequences of other plasmids, it was also observed that the injected plasmid did not seem to contain a poly(A) signal. A BLAST search of the same sequence also failed to identify any polyadenylation signal in that region.

Codon usage (CU) preference is also known to be associated with transcription rate through the relative abundance of charged tRNA isoacceptors for each codon [136, 139]. It was predicted that genes with codon usage preferences similar to the general zebrafish CU would be more easily transcribed than genes with greater dissimilarities. To measure the level of CU similarity between *TK* and *D. rerio* protein coding genes, two variables were analysed: the correlation between the relative codon usage (a measure of how often a given codon was used among all codon for that aminoacid) of all *TK* and *D. rerio* codons (Supp. Fig. S.3 B-G) and the codon usage preference score (CUPS), which shows the fraction of aminoacids sharing a similar CU between *TK* and *D. rerio*. The same comparisons were made between *D. rerio* CU and *mmGFP*, *EGFP*, *DsRed2* and *mCherry*; zebrafish *β-actin1* (*actb1*) gene was also added for reference as a housekeeping gene. Surprisingly, *TK* and *mmGFP* genes proved to be the genes with higher similarities to the overall zebrafish codon usage (Supp. Fig. S.3 B, C), both with a higher score in CUPS than even endogenous *β-actin1* (Supp. Fig. S.3 G). Conversely, genes coding for proteins that

had already shown a high expression level in zebrafish, like *EGFP* (Supp. Fig. S.3 D), score very low on codon usage preference and show minimal correlation to zebrafish CU. This results may mean that, for zebrafish transgenesis, codon usage is not a critical factor in determining expression levels of heterologous genes. It was later found that hydrolysis of pTol(*ef1α:DsRed2(floxed)-ntr-EGFP*) for the cloning procedure had unintentionally removed the poly(A) signal of said plasmid, which likely explained the low number of $EGFP^+$ cells.

In any case, even though codon usage preferences do not seem to favour a slower transcription rate for either *TK* or *mmGFP*, it was considered that a construct without a Kozak sequence and a poly(A) signal would not be able to sustain an adequate level of transgene expression for the intended cell ablation experiment. The use of (*foxj1a:TK-GFP*) construct was therefore rejected in favour of a new strategy for expression of the *TK* gene in $foxj1a^+$ cells.

1.4.2 *Tg(foxj1a:TK-T2a-EGFP)*

A new strategy to generate a transgenic line expressing the TK protein in $foxj1a^+$ cells depended on the creation of a different construct containing the best possible features for transgene integration and expression. Construct (*foxj1a:TK-T2a-EGFP*) was then designed bearing in mind the need to include the natural variant of *D. rerio* Kozak sequence flanking the *TK* start codon, as well as using a plasmid backbone with reliable Tol2-recognition sequences and a good poly(A) signal (pTol(*foxj1a:EGFP*) plasmid). To remove potential problems in protein folding and function arising from a fusion of TK with a fluorescent reporter protein (as was the case in the previous construct), coding sequences for both proteins were separated by the sequence for the self-cleaving peptide T2a, which also ensures they are produced in equimolar amounts. EGFP was chosen as the fluorescent reporter given that this particular sequence had already proved to result in high protein expression levels under the control of the *foxj1a* promoter.

To assemble the three coding sequences forming the (*TK-T2a-EGFP*) ORF, a series of fusion PCR [114, 115] reactions was set up as depicted in Supp. Figure S.4. Fusion PCR is a technique that allows joining of DNA sequences without the use of restriction enzymes, by adding to the amplified sequences a stretch of nucleotides that is complementary between those sequences, allowing them to be joined in a second PCR reaction. Supp. Figure S.4 depicts the method applied to this particular construct, in which a first set of reactions was used to amplify individual coding sequences and the following two reactions were used to join them in the intended order. The complete sequence was then cloned into pTol(*foxj1a:EGFP*) after removal of its *EGFP* coding sequence, thus creating pTol(*foxj1a:TK-T2a-EGFP*) (Fig. 1.6 A). Positive ligation clones were partially sequenced, confirming ligation success (Supp. Fig. S.1 D).

Ligation clone #11 was selected for injection into one-cell stage wildtype embryos along with Tol2 transposase mRNA and at 24 hpf embryos expressing EGFP in pronephros, floor plate and olfactory bulb cells were identified by fluorescence stereo microscopy (Fig. 1.6 C-D). Of the 5 independent injection events, only one exceeded the established minimum of positive embryos for growing, resulting in 52 (12.4%) $EGFP^+$ embryos injected with $30ng.\mu L^{-1}$ DNA and $25ng.\mu L^{-1}$ mRNA. Other injection events, using higher and lower nucleic acid concentrations, yielded from 0.0% to 10.8% of positives in an average of 514±367 injected embryos per day.

At 5 dpf, the same 52 larvae were observed again to confirm transgene expression before being transferred to the main aquaria system for growing. However, this second observation revealed that in four days these larvae had lost most of their initial EGFP expression (Fig. 1.6 B, G), with only 6 larvae retaining some green-labelled cells. In fact, detection of $EGFP^+$ cells was only possible using higher

magnifications, and these seemed to reside now in the posterior-most part of the brain, with no cells being observed either in the spinal cord or olfactory bulbs. Interestingly, larvae that had been injected with a much lower mRNA concentration show similar levels of EGFP expression at 5 dpf (Fig. 1.6 F), despite having had much fewer positive cells at 24 hpf (Fig. 1.6 C) than embryos injected with more nucleic acid (Fig. 1.6 D). The ratio of positive embryos is also unaltered from 24 hpf to 5 dpf (Fig. 1.6 B). These results suggested that, contrary to what had been thought, transposase mRNA might not be promoting genome integration. Moreover, this could mean that reporter expression levels at 24 hpf may not be considered a reliable measure of transgene integration, thus not an adequate criterion for selection of positives. To understand whether this could be due to an undetected problem with the plasmid backbone (namely with Tol2 recognition sequences) or with transposase translation or function, plasmid pTol(*foxj1a:EGFP*) was injected in one-cell stage embryos from the same wildtype line, using the lower DNA and mRNA concentrations from before (25ng.µL^{-1} and 5ng.µL^{-1}, respectively), and embryos were observed at 24 hpf and as 5 dpf larvae.

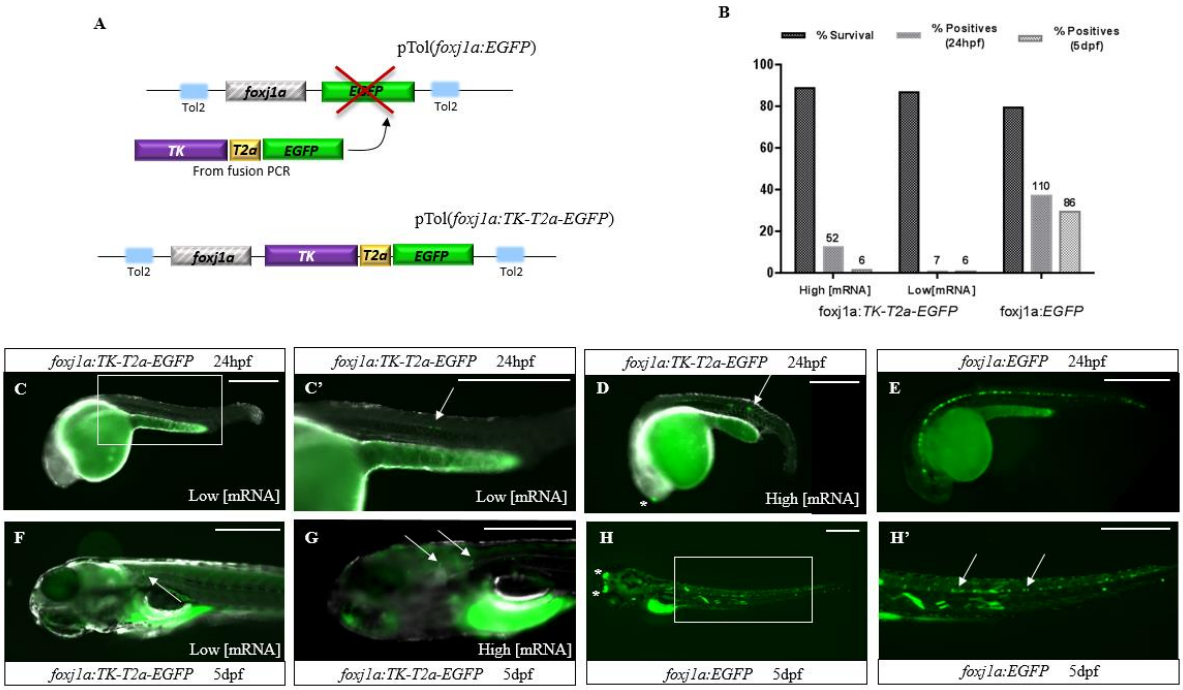


Figure 1.6 Molecular cloning scheme of pTol(*foxj1a:TK-T2a-EGFP*) plasmid, microinjection statistics and screening of embryos injected with (*foxj1a:TK-T2a-EGFP*) or (*foxj1a:EGFP*) constructs. **A:** The (*TK-T2a-EGFP*) coding sequence was inserted into a Tol2-based plasmid that retained the *foxj1a* promoter, generating the final plasmid. Grey boxes represent promoter regions. **B:** Injection of higher mRNA concentration yields more positives at 24 hpf for the (*foxj1a:TK-T2a-EGFP*) construct but at 5 dpf this number is equal to the obtained with low concentrations. (*foxj1a:EGFP*) construct yields the highest positive rate even at low mRNA concentrations. Absolute number of positives displayed above each column. **C-H':** Screening of injected embryos. Embryos injected with low mRNA concentration (C, C') show fewer EGFP^{+} cells (arrow) than when injected with higher concentrations (D). At 5 dpf, larvae from both situations (F, low; G, high) retain similar low numbers of EGFP^{+} cells (arrow). Embryos injected with low mRNA concentrations and (*foxj1a:EGFP*) construct display stronger widespread labelling of *foxj1a*$^{+}$ cells from 24 hpf (E) until 5 dpf (H, H'), including in floor plate (arrow) and olfactory bulb (*) cells. C' and H' show boxed areas of C and H respectively. C, C', D, F and G show merged images from green channel and darkfield. Scale bar: 500µm.

When observed at 24 hpf, embryos injected with (*foxj1a:EGFP*) construct present strong EGFP signal in cells from the floor plate (Fig. 1.6 E) and olfactory bulbs, with pronephros labelling being evident in some cases. Contrary to construct (*foxj1a:TK-T2a-EGFP*) injected previously, EGFP labelling is found

in the majority of floor plate *foxj1a*$^{+}$ cells, highlighting the whole structure from the posterior-most part of the brain to the end of the tail. This plasmid was also able to yield a higher number of positives, with 36.9% of the 478 injected embryos being EGFP^{+} (Fig. 1.6 B); it is worth noting that this was the highest expression efficiency ever obtained among all the injected constructs. Four days later, the larvae were observed for the second time and it was visible that EGFP expression was still present in the structures identified at 24 hpf in most larvae (Fig. 1.6 B, H, H'). Despite the appearance of some *foxj1a*$^{-}$ structures also expressing EGFP (which is a common artefact found when injecting some plasmids), floor plate and olfactory bulb cells can be easily identified by a strong green fluorescence, thus confirming the injected construct is still expressed five days after injection. These observations indicate that the (*foxj1a:EGFP*) cassette has probably been inserted in these embryos' genomes upon injection, depending by necessity on the Tol2 transposase mRNA that was co-injected. Given that the same mRNA concentration was unable to sustain a similar level of transgene integration with the (*foxj1a:TK-T2a-EGFP*) cassette, it can be assumed that transgenesis efficiency is being limited by the injected DNA and not by the quality of the mRNA or its protein product. The reasons leading to this discrepancy between two constructs that share the same plasmid backbone are not clear. Nevertheless, it seems clear that the level of transgene integration obtained with this plasmid may not result in a sufficient number of positive embryos for the cell ablation experiment or for the generation of a transgenic line.

2. CreER^{T2} activation with 4-OHT

The role of ependymal cells during spinal cord regeneration can be studied by the genetic labelling of *foxj1a*$^{+}$ cells and subsequent tracing of their progeny, making use of one of the CreER^{T2}/LoxP transgenic zebrafish lines [140, 141] mentioned before. Each of these lines would be composed of two main expression cassettes: a reporter construct, consisting of two ORFs for different fluorescent proteins in which the first would be flanked by LoxP sites, and a driver construct, coding for a version of Cre recombinase (CreER^{T2}) that is able to promote recombination only upon activation with 4-OHT [106]. Thus, to ensure the success of lineage tracing experiments, and since previous attempts with this strategy were not fortunate, a protocol for efficient CreER^{T2} activation had to be formulated. Optimization was conducted in embryos or larvae of different transgenic lines given that drug access to tissues would be more difficult in adult fish and also because embryos are easier to obtain in larger numbers.

2.1 Recombination in the whole embryo: *Tg(hs:mCh-Cre)xTg(hs:Red2Green)*

Activation of CreER^{T2} is known to occur in a mosaic fashion, with the recombination rate of CreER^{T2+} cells never reaching 100% [30]. Thus, to improve the likelihood of observing even small recombination rates, CreER^{T2} activation was attempted in embryos that express both the reporter and the driver constructs in the majority of their cells. Embryos obtained from mating of *Tg(hs:mCh-Cre)* with *Tg(hs:Red2Green)* fish were subject to heat shock to promote expression of CreER^{T2} at 7 hpf. At 24 hpf (16h after the heat shock) mCherry expression was confirmed by fluorescence microscopy and embryos were treated with 4µM, 8µM or 10µM 4-OHT and left at 33ºC in the dark for 6h. In a second group of embryos treated with 8µM 4-OHT there was addition of pronase to promote embryo release from the chorion and thus investigate if the chorion posed a substantial barrier to drug diffusion. Two negative controls were also performed: a control for the function of CreER^{T2}, with embryos from the same batch that did not express any red fluorescent protein after heat shock (*mCh-Cre*$^{-}$ *Red2Green*$^{-}$) being treated in the same conditions as test embryos, and a control for the function of 4-OHT, with the same embryos

used for testing but using EtOH instead of the drug. To induce expression of the reporter proteins, a second heat shock was performed at the end of the 6h of 4-OHT treatment.

At 48 hpf embryos were observed again, and this time it was possible to detect several EGFP^{+} cells in most treated embryos with all 4-OHT concentrations, indicating that CreER^{T2}-mediated recombination had occurred. In order to improve visualization and confirm that the observed green fluorescence was in fact generated by EGFP proteins, embryos were immediately fixed and processed for immunochemical detection of EGFP.

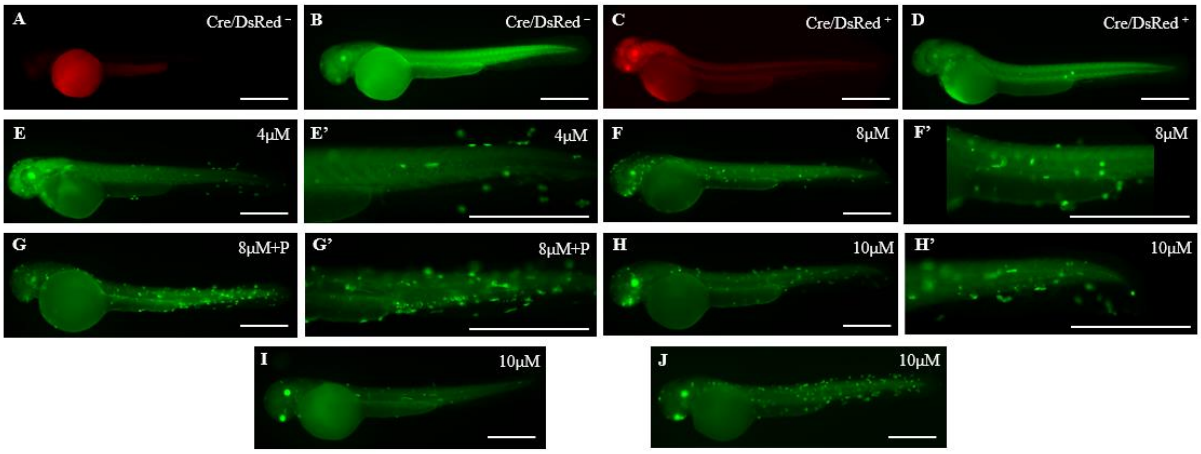


Figure 2.1 4-OHT treatment was tested in *Tg(hs:mCh-Cre)xTg(hs:Red2Green)* embryos after heat-shock induced CreER^{T2} expression. Negative controls for the transgenes (CreER^{T2-} DsRed2^{-} with 4μM 4-OHT)(A, B) and for the drug (CreER^{T2+} DsRed2^{+} without 4-OHT)(C, D) show absence of Cre-mediated recombination. Embryos treated with 4μM 4-OHT (E, E') show fewer recombined EGFP^{+} cells than embryos treated with 8μM (F, F'), or with 8μM+pronase (G, G'), with the latter showing the highest recombination efficiency. Treatment with 10μM was heterogeneous, resulting in low (I), medium (H, H') and high (J) recombination efficiencies. E', F', G' and H' show zoomed areas of embryos treated with the same conditions as E, F, G and H respectively. All embryos show red fluorescence similar to C. All pictures depict 48 hpf embryos processed for immunochemical detection of EGFP. Scale bar: 500μm.

Figure 2.1 depicts representative images of embryos from each treatment, along with both controls, confirming that EGFP^{+} cells are present in embryos from all treatment groups. Recombined cells were found in different tissues throughout the body of the embryos, both deep and superficial. It was also observed that the activation effect of 4-OHT increased as drug concentration doubled from 4μM (Fig. 2.1 E, E') to 8μM (Fig. 2.1 F, F'), as was expected, and was further enhanced by addition of pronase (Fig. 2.1 G, G'). In fact, even though embryos treated with 10μM 4-OHT showed variable amounts of EGFP^{+} cells (Fig. 2.1 H-J), with some similar to 4μM treated embryos (Fig. 2.1 I), they never reached the maximum level of recombined cells visible in embryos treated with 8μM+pronase. Nevertheless, a concentration of 10μM 4-OHT resulted in a higher total number of embryos showing recombined cells than treatment with 8μM+pronase. Both negative controls showed that in absence of CreER^{T2}, reporter protein or 4-OHT there was no phenotype switch from DsRed2^{+} to EGFP^{+} (Fig. 2.1 A-D) (green spots in Figure 2.1 D are artefacts introduced by the processing of embryos for antibody staining).

This assay confirmed for the first time in this lab the functionality of a CreER^{T2} driver line to perform recombination of LoxP sites and generate a phenotype switch. Using this protocol as a starting point, it may then be possible to achieve recombination of *foxj1a*$^{+}$ cells in other transgenic lines.

2.2 Recombination in *foxj1a*$^{+}$ cells: *Tg(ubi:Switch)x(foxj1a:CreER^{T2}; cmlc2:EGFP)*

After finding a protocol that effectively led to CreER^{T2} activation, it was necessary to test it in one of the lineage tracing transgenic lines, and adapt it if required. Adult fish carrying both constructs (*ubi:Switch*) and (*foxj1a:CreER^{T2}; cmlc2:EGFP*) were already available (see Chapter 1.1, Results section) so embryos were obtained from incross of these fish. Double positives were selected at 24 hpf, treated with 10µM 4-OHT and pronase to improve drug accessibility and kept at 33ºC in the dark. Following reports that this drug takes approximately 2 hours to promote CreER^{T2} activation and recombination but fluorescent protein expression and maturation may need 90 minutes to 4 hours [142], observation was first conducted 6h after 4-OHT addition to the medium and again at 48 hpf. However, detection of DsRed2^{+} cells was never possible (data not shown).

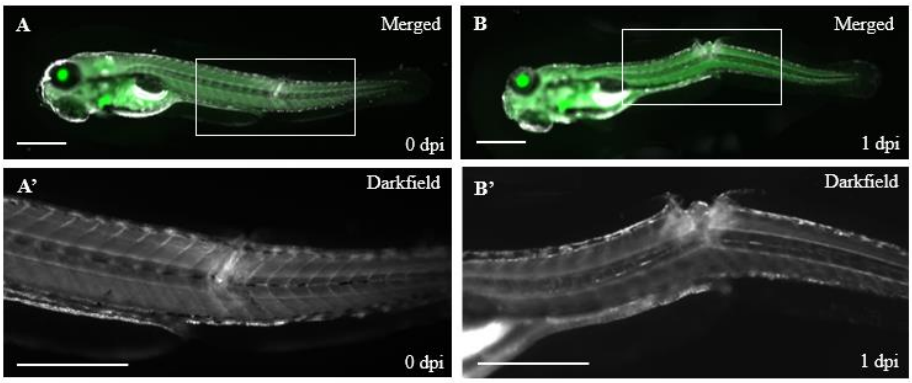


Figure 2.2 4-OHT treatment was tested in *Tg(ubi:Switch)x(fox:CreER^{T2}; cmlc2:EGFP*) larvae after SCI. **A, A':** SCI was performed in 5 dpf larvae by transecting the spinal cord (SC) without affecting the whole width of the notochord. **B, B':** At 1 dpl cell proliferation at the lesion site had created a blastema that would lead to wound closure and tissue regeneration. A' and B' show boxed areas of A and B respectively. A and B show merged images from green channel and darkfield. Scale bar: 500µm.

Because the CreER^{T2} enzyme was only expressed in *foxj1a*$^{+}$ cells, it was possible that recombination was occurring at a rate too low to be detected. It has been described that one of the hallmarks of spinal cord ependymal cell reaction to injury is a strong proliferative response [91]. Therefore, one way to invoke a proliferative response in *foxj1a*$^{+}$ cells, thus increasing the likelihood of detecting recombined cells, would be to inflict a spinal cord injury (SCI) before treatment with 4-OHT. To accomplish this, double transgenic embryos from an incross of the same founder fish were selected and as 5 dpf larvae they were subjected to SCI according to standard techniques (Fig. 2.2 A, A'). Immediately after the surgery, larvae were transferred to medium with 10µM 4-OHT and kept at 33ºC in the dark, this time for 4h30min due their higher fragility. At 6 dpf larvae were observed and, while no recombination was detected, it seemed that the regeneration process had not been affected by treatment conditions (Fig. 2.2 B, B'). Regenerating larvae were incubated in 10µM 4-OHT at 33ºC for a second time, but observation at 7 dpf did not reveal the presence of any DsRed2^{+} cells, indicating that the protocol used is not efficient in promoting CreER^{T2} activation in this line. Given that *foxj1a*$^{+}$ cells reside deep within the larvae it may be a problem of drug accessibility to the tissue, which could be solved by using a permeabilization agent other than DMSO or in higher concentrations; raising 4-OHT concentration may not be advisable since 10µM has been described as the maximum non-toxic concentration [107, 142]. On the other hand, the negative results observed in this experiment could be due to an intrinsic problem of this transgenic line, which had already been used by others with little success, thus prompting the need to try variations of this protocol in other lines that use a different *CreER^{T2}* construct.

Discussion

Spinal cord injury is a debilitating condition to which there is yet no cure or effective treatment. In contrast to mammals, the zebrafish displays remarkable regeneration capacities that are being studied in order to find key features that can be activated or inactivated in humans to promote functional recovery. This is only possible due to the existent structural similarities between *D. rerio* and mammalian models, such as the presence of ependymal cells lining the central canal of the SC that derive from similar precursors in both models and exhibit hallmark features of NSC (*in vivo* in zebrafish, *in vitro* in mammals). Understanding how EC are able to generate new neurons and promote axonal regeneration after a SCI in fish can give invaluable insight into how endogenous human EC can potentially be manipulated to obtain the same functional output.

Zebrafish ependymal cells are multiciliated and therefore can be identified by the expression of Foxj1a transcription factor. The major aim of this work was to generate transgenic fish lines that used this marker to label spinal cord EC and track and identify their progeny after a SCI and, in a parallel manner, to ablate proliferating EC responding to injury and determine their role in regeneration. A widely used method for transgene integration, the Tol2 system [117], was chosen to generate transgenics. This method relies on the co-injection in one-cell stage embryos of a donor plasmid, carrying the desired construct between Tol2 recognition sequences, and mRNA for the Tol2 transposase; the enzyme is translated inside the early embryo and catalyses the transposition of the construct from the donor plasmid to the embryo genome. Germline transmission depends on this transposition event occurring early in development so that the primordial germ cells (PGC) already carry this insertion and can pass it to their progeny, the germ cells [143]. In this work, five separate constructs were injected along with transposase mRNA in order to produce transgenic fish, of which: the (*foxj1a:TK-T2a-EGFP*) did not yield enough positive embryos for growing, the (*foxj1a:TK-GFP*) resulted in almost imperceptible levels of transient expression, the (*foxj1a:DsRed2(floxed)-EGFP*) did not produce any founders and the (*foxj1a:mCherry-T2a-CreER^{T2}*) resulted in inconclusive founder screening. Only with the construct (*foxj1a:CreER^{T2}; cmlc2:EGFP*) was it possible to obtain an adequate rate of positive injected embryos that were grown and passed the transgene to their offspring. Nevertheless, the success rate of this transgenesis was below that expected for the Tol2 system, which has been reported to result in 50-70% of founders and a germline transmission rate between 3 and 100% among the injected embryos using much lower DNA and mRNA concentrations [116] (against 15.7% and 2-33% observed in this work, respectively).

For the four other constructs several optimization attempts were made to improve both the level of transient transgene expression and the number of positive embryos, adjusting DNA and mRNA concentrations in the range of 10-37.5ng.μL^{-1} and 5-27.5ng.μL^{-1} respectively. It was found that each construct had a different optimal concentration for each variable (DNA or mRNA concentration or total nucleic acid) and that survival rate and percentage of positives could not always be predicted from the same variable (Figures 1.2, 1.4, 1.5). In fact, both percentages were highly fluctuating even for the same value of DNA or mRNA concentration or total nucleic acid, suggesting the role of an unknown stochastic factor (such as egg quality) that could not be eliminated by normalizing survival rates. Differences between constructs can also be attributed to the fact that injected embryos possessed distinct genetic backgrounds which could confer separate receptivity and sensitivity to transgenesis. Nevertheless, the number and quality of positive embryos remained unsatisfactorily low and could not be justified in some constructs. Sequencing results revealed no mutations either in promoter regions or in the coding sequences of (*foxj1a:mCherry-T2a-CreER^{T2}*) and (*foxj1a:DsRed2(floxed)-EGFP*), also confirming the success of ligation procedures (Supp. Fig. S.1). Injection of (*ef1α:DsRed2(floxed)-ntr-*

EGFP) as a positive control (Fig. 1.4 C) also dismissed concerns that construct size might be impairing transposition (this construct is longer than any of the others) or that (*foxj1a:DsRed2(floxed)-EGFP*) was not being integrated into the genome due to undetected mutations in Tol2 recognition sequences (these constructs share the same backbone). Additionally, Tol2 mediated transgenesis ensures that only a single copy of a construct is inserted in a given locus [144], preventing transgene silencing that is known to occur when transgenes are inserted in concatamers [61, 145] or when there are several copies of the promoter [146].

Low GFP expression obtained with injection of (*foxj1a:TK-GFP*) construct was also investigated by looking at sequence features known to impact translation rate (Supp. Fig. S.3). Codon usage (CU) is known to not be random but, on the opposite, to present specific frequencies depending on the taxon, including at the species level [139, 147]. Therefore, it was hypothesized that coding sequences whose codon usage presented more dissimilar to the *Danio rerio* codon usage (obtained from analysis of all known protein coding genes [130]) would be translated at a slower rate in zebrafish, with a potential negative impact on protein expression levels. Examination of the relationship between *mmGFP* and *D. rerio* CU did not expose the differences expected for an impairment in protein translation when compared with an endogenous housekeeping zebrafish protein coding gene (*β-actin1*), which presented more differences to *D. rerio* CU than *mmGFP* did (CUPS score of 16/18 for *mmGFP* and 9/18 for *β-actin1*). The same analysis was extended to the rest of the sequences coding for fluorescent proteins used in this work, with another unexpected result arising from the *EGFP* sequence. While *EGFP* was by far the gene that showed the best *foxj1a* expression pattern when injected ((*foxj1a:EGFP*) construct, Fig 1.6 E), CU comparisons resulted in a CUPS score of only 3/18 (very few similarities in CU). On one hand, this result may mean that, despite not being "finely tuned" to the zebrafish codon usage, *EGFP* coding sequence can still be translated at an adequate rate and, conversely, impairment of mmGFP expression is due to other factors (see further below). On the other hand, these results may reflect an incorrect approach to CU comparisons. One of the main theories explaining codon usage evolution posits that it is biased towards the use of synonymous codons for which there is a higher abundance of isoaccepting tRNAs, while at the same time diminishing the frequency of codons for which there are less tRNAs (translational selection) [147-149]. Logically, it follows that if such selection exists then highly expressed and more essential genes will display a CU more biased towards tRNA abundance than less expressed genes [136, 147, 150]. In this perspective, only the codon usage of these highly expressed genes will be optimized to produce high protein expression levels, thus being the preferred comparison to the coding sequences under study. In fact, more than one platform has been created to optimize coding sequences relative to the CU of highly expressed genes of a given organism [136, 150]. The approach used in this work would have been more appropriate to compare two individual coding sequences, as was done by Plotkin *et al* [139] to show that tissue-specific genes differ in CU from genes of other tissues. The Kozak sequence is also known to impact gene expression due to its role in signalling the initiation codon [137, 151]. The original sequence compiled by Marilyn Kozak in 1987 was obtained mainly from analysis of mammalian genes; only recently have individual organisms' genomes been studied for the occurrence of natural Kozak sequence variants [152]. In particular, *D. rerio* natural variant was found to be more abundant and almost twice as efficient as the original Kozak sequence for promoting translation initiation of exogenous genes in zebrafish. Analysis of the 5' UTR of the (*TK-GFP*) coding sequence revealed the absence of a sequence with Kozak-like features (Supp.Fig. S.3 A), namely a correct alignment to either the original Kozak or the *D. rerio* variant in the positions closer to the starting ATG or the presence of a purine in position -3 (considered a critical component of any Kozak sequence [138]). *In vivo* studies have shown that a translational start site consisting of a non-Kozak sequence lead to a measurable impairment in translation efficiency and protein expression levels [152]; it is then expected that a construct without a Kozak sequence, such as the (*foxj1a:TK-GFP*) construct,

would show reduced expression levels. Finally, it was found by further sequencing that plasmid pTol(*foxj1a:TK-GFP*) lacked a transcription termination and polyadenylation (poly(A)) signal, inadvertently removed during cloning procedure. The role of poly(A) signals in transcription termination and mRNA processing has been long known [153], and functions such as mRNA stabilization and promotion of translation initiation are essential to ensure an adequate level of protein expression [154, 155]. In that view, a construct intended for heterologous expression that does not possess a poly(A) signal will probably result in very low expression levels, such as the ones observed after injection of pTol(*foxj1a:TK-GFP*). Compared to the other reasons presented before, it is likely that lack of a poly(A) signal is the major obstacle preventing correct expression of TK-GFP fusion protein. In contrast, constructs (*foxj1a:mCherry-T2a-CreER^{T2}*), (*foxj1a:DsRed2(floxed)-EGFP*) and (*foxj1a:EGFP*) were confirmed for the presence of both Kozak sequence and poly(A) signal for each ORF.

The construct (*foxj1a:TK-T2a-EGFP*) is a particular case in which transient expression after injection (at 24 hpf) proved successful but was lost at 5 dpf (retained only in 11.5% of initially detected larvae and in a very restricted domain) (Fig. 1.6). Previous studies have shown that one day after injection up to 50% of the inserted DNA has still resisted degradation, persisting extrachromosomally and potentially contributing to the observed expression [145, 156]. However, it is arguable whether a relevant fraction of the initial injected DNA that was not integrated still persists after five days and is capable of sustaining considerable expression levels; for this reason, it was considered that 5 dpf expression derives from transgene copies integrated into the host genome. From this perspective, it appeared that the efficiency of transgene integration (seen by expression at 5 dpf) was independent of the transposase mRNA concentration used, despite transient expression levels being higher when more mRNA was injected. This pointed to a failure in the transposition mechanism, either on the part of the construct or of the enzyme. Injection of pTol(*foxj1a:EGFP*) as a positive control suggests the transposase is functional, given that reporter expression remains high at 5 dpf. However, sequence features of the (*foxj1a:TK-T2a-EGFP*) construct were planned during its construction so that it should be capable of sustaining adequate reporter expression. This plasmid contains Tol2 recognition sequences, promoter sequence, *EGFP* coding sequence and poly(A) signal derived from plasmid pTol(*foxj1a:EGFP*) (used as a positive control throughout this work) as well as the zebrafish Kozak sequence variant; the T2a peptide sequence is the same as in *Tg(hs:mCh-Cre*) embryos, in which both mCherry and CreER^{T2} are expressed and functional (see further below and Chapter 2.1 of Results section). Additionally, the whole ORF and ligation sites of pTol(*foxj1a:TK-T2a-EGFP*) were sequenced before injection and no mutations were detected, supporting 24 hpf observations in which there was EGFP expression in *foxj1a*$^{+}$ cells.

Ultimately, it was possible to obtain enough positive embryos from injection of (*foxj1a:DsRed2(floxed)-EGFP*) and (*foxj1a:mCherry-T2a-CreER^{T2}*) constructs, which were grown to adulthood. While no founders were obtained with the first construct, results from the second one were not conclusive. The first screening approach (looking for mCherry fluorescence in F1 embryos) indicated that there had been no germline transmission in any of the potential founders. However, it was possible that the construct was present and being expressed in F1 embryos without detectable mCherry fluorescence, as observed by others [135]. Alternative screening approaches were then applied, revealing conflicting results (Fig. 1.3). PCR detection of the CreER^{T2} coding sequence performed on DNA extracted from F1 embryos revealed the presence of the transgene in progeny of two couples of potential founders, suggesting that at least two of these fish passed the transgene to their germline. However, there was also unforeseen amplification thought to derive from other unknown genomic regions (since the primers used have partial sequence similarity to the zebrafish genome), bringing into question the identity of the *CreER^{T2}*-sized amplicon observed. *CreER^{T2}* transcription was analysed separately by WISH and RT-PCR: while the first revealed the presence of *CreER^{T2}* expression in the great majority of *foxj1a*$^{+}$ cells in a fraction of F1 embryos (consistent with an expected low germline transmission rate, as observed with the

(*foxj1a:CreER^{T2}; cmlc2:EGFP*) construct), it was not possible to detect the presence of *CreER^{T2}* transcripts using RT-PCR. However, one of the positive controls may indicate this negative result as a consequence of the lower detection threshold (or sensitivity) of the method. When RNA was extracted from *Tg(hs:mCh-Cre*) embryos that had been subjected to heat shock-induced expression of *CreER^{T2}*, the band resulting from RT-PCR was very faint compared to the other positive control (a plasmid containing the *CreER^{T2}* coding sequence). If this reflects accurately the level of *CreER^{T2}* expression in these embryos, which occurs in the great majority of cells after a heat shock, then the amount of *CreER^{T2}* mRNA produced by positive F1 test embryos (in *foxj1a^{+}* cells only) may be so low as to be undetectable by this RT-PCR protocol. Overall, these screening methods suggest that some of the potential founder fish may have passed the transgene to their germline, and these F1 embryos could be used to establish a stable line. Still, a fast identification of founders requires methods that can be easily applied with no doubt as to their results, contrary to the ones described above. Furthermore, since these methods involve killing the embryos, lineage tracing experiments would have to be conducted without screening them, leading to the use of a much higher number of embryos to ensure a minimum amount of positives in each batch. This problem would be minimized with the establishment of the transgenic line, when all fish would be hemizygous and 75% of the progeny from an incross would carry the transgene. However, the time required to establish a transgenic line with these screening methods would be too long to allow generation of results to be included in this work.

Overall, it is clear that the simple, straightforward Tol2 method for generation of transgenics in zebrafish did not yield the expected results, and the transgenic lines necessary for lineage tracing and cell ablation experiments could not be obtained. One of the main reasons for this, the low germline transmission rate, could in the future be overcome by the addition of the *nanos1* 3' UTR to the transposase mRNA, which has been observed to promote mRNA accumulation in PGC [157]. As a result, recombination in PCG is increased and germline transmission rates rise. Still, one of the drawbacks of the Tol2 system is that transgene insertion can happen in any region of the host genome, being affected by the chromatin dynamics of that site [156]. In practice, this means that different F0 organisms injected with the same construct often display significant levels of variability in transgene expression pattern and level and even transgene silencing [62, 158]. Recently, a novel method for zebrafish transgenesis was developed in which a phage integrase can be targeted for a desired genome sequence [62, 158]. This method has been used for *Drosophila* transgenesis with great success for more than a decade but depends on the previous establishment of several "recipient" lines, containing the phage integration site in specific genomic regions. The future development of recipient zebrafish lines will facilitate the generation of transgenics with this new site-specific integration method, overcoming the undesirable position effects resulting from Tol2 transposition.

For the lineage tracing experiment, all three different lines designed to permanently label *foxj1a^{+}* cells and their progeny relied on the CreER^{T2}/LoxP system. Adding to the known properties of conditional expression of conventional Cre/LoxP technology, the fusion of a modified version of the oestrogen receptor (ER^{T2}) to Cre makes the system inducible, since it is necessary to add tamoxifen (or its metabolite 4-OHT) for CreER^{T2} to be translocated to the nucleus and promote recombination [106, 159]. This system has been used in zebrafish to promote inducible and conditional gene expression [64, 160, 161], but so far it had not been possible to activate CreER^{T2}-mediated recombination with 4-OHT in this lab. The first attempt to obtain a functional protocol was performed in double transgenic embryos that expressed both the Cre-driver construct and the reporter construct in the majority of cells after heat shock induction (*Tg(hs:mCh-Cre)xTg(hs:Red2Green*)). The protocol consisted in 4-OHT application 16h after induction of CreER^{T2} expression, followed by a second heat shock to promote a second wave

of expression from the reporter gene (*DsRed2* before recombination and *EGFP* afterwards). Recombination efficiency increased as 4-OHT concentration was raised and it was noted that removal of the chorion with pronase further enhanced this effect by improving drug accessibility to tissues (Fig. 2.1). Nevertheless, the best results obtained (with 8μM 4-OHT plus pronase) indicate that recombination efficiency was probably lower than previously reported even with inferior 4-OHT concentrations [162], given that recombined cells can almost be individualized at any plane of focus (showing that they are few compared with all embryo cells). Despite the success of this protocol in this specific line, attempts to promote recombination in *foxj1a*$^{+}$ cells of embryos carrying the (*ubi:Switch*) and (*foxj1a:CreER*^{T2}*;cmlc2:EGFP*) constructs (F1 of identified founders) were fruitless, even after *foxj1a*$^{+}$ cell proliferation was induced by SCI (Fig.2.2). It is possible that a low recombination rate resulted in a DsRed2 signal below the threshold of visual detection, even though a small number of EGFP^{+} cells had been detected before, after injection of some constructs (Fig.1.6). However, it is also possible that recombination in these cells has failed, either due to problems in enzyme activity or in *CreER*^{T2} expression. Chromatin dynamics are known to be closely associated with transcription status (activation *vs* repression), with chromatin modulation factors (such as DNA methyltransferases and histone methyltransferases and acetyltransferases) working synergistically with promoter sequence binding proteins to activate or prevent transcription [163-165]. The construct (*foxj1a:CreER*^{T2}*; cmlc2:EGFP*) includes coding sequences whose expression is driven by promoters that are not active in the same cell type at any time (*cmlc2* is specific of cardiac muscle cells [166], which are not multiciliated, therefore not expressing *foxj1a*). It is then conceivable that *foxj1a*$^{+}$ cells may be repressing *cmlc2* expression by imposing a state of compressed chromatin to any loci possessing the *cmlc2* promoter, in this case the endogenous *cmlc2* locus and the construct insertion site. Such a repressive state would make the *foxj1a* promoter present in the inserted construct also inaccessible to its cognate transcription factor. Assessment of *CreER*^{T2} transcription status in these embryos (by RT-PCR for instance) is then necessary to determine whether this line may be used for further optimization of 4-OHT mediated CreER^{T2} activation and lineage tracing experiments.

Given the impossibility to generate functional transgenic lines in this work, further studies will be necessary to obtain biological tools with which to understand the role of *foxj1a*$^{+}$ cells in the zebrafish spinal cord regeneration.

References

1. Organization WH, Society ISC: **International perspectives on spinal cord injury**: World Health Organization; 2013.
2. Vander A, Sherman J, Luciano D: **Human physiology 8th edition**. In.: McGraw Hill; 2001.
3. Kahle W, Frotscher M: **Color atlas of human anatomy. Vol. 3**. In.: Thieme Medical Publishers, Incorporated; 2002.
4. Silva NA, Sousa N, Reis RL, Salgado AJ: **From basics to clinical: a comprehensive review on spinal cord injury**. *Progress in neurobiology* 2014, **114**:25-57.
5. Norenberg MD, Smith J, Marcillo A: **The pathology of human spinal cord injury: defining the problems**. *Journal of neurotrauma* 2004, **21**(4):429-440.

6. Dumont RJ, Okonkwo DO, Verma S, Hurlbert RJ, Boulos PT, Ellegala DB, Dumont AS: **Acute spinal cord injury, part I: pathophysiologic mechanisms**. *Clinical neuropharmacology* 2001, **24**(5):254-264.

7. Hulsebosch CE: **Recent advances in pathophysiology and treatment of spinal cord injury**. *Advances in physiology education* 2002, **26**(1-4):238-255.

8. Varma AK, Das A, Wallace G, Barry J, Vertegel AA, Ray SK, Banik NL: **Spinal cord injury: a review of current therapy, future treatments, and basic science frontiers**. *Neurochemical research* 2013, **38**(5):895-905.

9. McDonald JW, Sadowsky C: **Spinal-cord injury**. *Lancet* 2002, **359**(9304):417-425.

10. Fehlings MG, Vaccaro A, Wilson JR, Singh A, Cadotte DW, Harrop JS, Aarabi B, Shaffrey C, Dvorak M, Fisher C *et al*: **Early versus delayed decompression for traumatic cervical spinal cord injury: results of the Surgical Timing in Acute Spinal Cord Injury Study (STASCIS)**. *PloS one* 2012, **7**(2):e32037.

11. Kwon BK, Tetzlaff W, Grauer JN, Beiner J, Vaccaro AR: **Pathophysiology and pharmacologic treatment of acute spinal cord injury**. *The spine journal : official journal of the North American Spine Society* 2004, **4**(4):451-464.

12. Thuret S, Moon LD, Gage FH: **Therapeutic interventions after spinal cord injury**. *Nature reviews Neuroscience* 2006, **7**(8):628-643.

13. Silver J, Miller JH: **Regeneration beyond the glial scar**. *Nature reviews Neuroscience* 2004, **5**(2):146-156.

14. Antonic A, Sena ES, Lees JS, Wills TE, Skeers P, Batchelor PE, Macleod MR, Howells DW: **Stem cell transplantation in traumatic spinal cord injury: a systematic review and meta-analysis of animal studies**. *PLoS biology* 2013, **11**(12):e1001738.

15. Mar FM, da Silva TF, Morgado MM, Rodrigues LG, Rodrigues D, Pereira MI, Marques A, Sousa VF, Coentro J, Sa-Miranda C *et al*: **Myelin Lipids Inhibit Axon Regeneration Following Spinal Cord Injury: a Novel Perspective for Therapy**. *Molecular neurobiology* 2016, **53**(2):1052-1064.

16. Mar FM, Bonni A, Sousa MM: **Cell intrinsic control of axon regeneration**. *EMBO reports* 2014, **15**(3):254-263.

17. Cajal SR: **Estudios sobre la degeneración y regeneración de los centros nerviosos**. *Madrid: Moya* 1914.

18. Martino G, Pluchino S, Bonfanti L, Schwartz M: **Brain regeneration in physiology and pathology: the immune signature driving therapeutic plasticity of neural stem cells**. *Physiol Rev* 2011, **91**(4):1281-1304.

19. Emsley JG, Mitchell BD, Kempermann G, Macklis JD: **Adult neurogenesis and repair of the adult CNS with neural progenitors, precursors, and stem cells**. *Progress in neurobiology* 2005, **75**(5):321-341.

20. Ming GL, Song H: **Adult neurogenesis in the mammalian central nervous system**. *Annual review of neuroscience* 2005, **28**:223-250.

21. Bernstock J, Verheyen J, Huang B, Hallenbeck J, Pluchino S: **Typical and atypical stem cell niches of the adult nervous system in health and inflammatory brain and spinal cord diseases**. *Adult Stem Cell Niches* 2014, **8**.

22. Weiss S, Dunne C, Hewson J, Wohl C, Wheatley M, Peterson AC, Reynolds BA: **Multipotent CNS stem cells are present in the adult mammalian spinal cord and ventricular neuroaxis**. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 1996, **16**(23):7599-7609.

23. Shihabuddin LS, Ray J, Gage FH: **FGF-2 is sufficient to isolate progenitors found in the adult mammalian spinal cord**. *Experimental neurology* 1997, **148**(2):577-586.

24. Johansson CB, Momma S, Clarke DL, Risling M, Lendahl U, Frisen J: **Identification of a neural stem cell in the adult mammalian central nervous system**. *Cell* 1999, **96**(1):25-34.

25. Barnabe-Heider F, Goritz C, Sabelstrom H, Takebayashi H, Pfrieger FW, Meletis K, Frisen J: **Origin of new glial cells in intact and injured adult spinal cord**. *Cell stem cell* 2010, **7**(4):470-482.

26. Hamilton LK, Truong MK, Bednarczyk MR, Aumont A, Fernandes KJ: **Cellular organization of the central canal ependymal zone, a niche of latent neural stem cells in the adult mammalian spinal cord**. *Neuroscience* 2009, **164**(3):1044-1056.

27. Ellis P, Fagan BM, Magness ST, Hutton S, Taranova O, Hayashi S, McMahon A, Rao M, Pevny L: **SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult**. *Developmental neuroscience* 2004, **26**(2-4):148-165.

28. Scott CE, Wynn SL, Sesay A, Cruz C, Cheung M, Gomez Gaviro MV, Booth S, Gao B, Cheah KS, Lovell-Badge R *et al*: **SOX9 induces and maintains neural stem cells**. *Nature neuroscience* 2010, **13**(10):1181-1189.

29. Kaneko Y, Sakakibara S, Imai T, Suzuki A, Nakamura Y, Sawamoto K, Ogawa Y, Toyama Y, Miyata T, Okano H: **Musashi1: an evolutionally conserved marker for CNS progenitor cells including neural stem cells**. *Developmental neuroscience* 2000, **22**(1-2):139.

30. Meletis K, Barnabe-Heider F, Carlen M, Evergren E, Tomilin N, Shupliakov O, Frisen J: **Spinal cord injury reveals multilineage differentiation of ependymal cells**. *PLoS biology* 2008, **6**(7):e182.

31. Lendahl U, Zimmerman LB, McKay RD: **CNS stem cells express a new class of intermediate filament protein**. *Cell* 1990, **60**(4):585-595.

32. Eng LF: **Glial fibrillary acidic protein (GFAP): the major protein of glial intermediate filaments in differentiated astrocytes**. *Journal of neuroimmunology* 1985, **8**(4-6):203-214.

33. Mao Y, Mathews K, Gorrie CA: **Temporal Response of Endogenous Neural Progenitor Cells Following Injury to the Adult Rat Spinal Cord**. *Frontiers in cellular neuroscience* 2016, **10**:58.

34. Sabelstrom H: **An ependymal cell quest: identification and functional role of spinal cord neural stem cells**: Inst för cell-och molekylärbiologi/Dept of Cell and Molecular Biology; 2014.

35. Dimou L, Götz M: **Glial cells as progenitors and stem cells: new roles in the healthy and diseased brain**. *Physiological reviews* 2014, **94**(3):709-737.

36. Dromard C, Guillon H, Rigau V, Ripoll C, Sabourin JC, Perrin FE, Scamps F, Bozza S, Sabatier P, Lonjon N *et al*: **Adult human spinal cord harbors neural precursor cells that generate neurons and glial cells in vitro**. *J Neurosci Res* 2008, **86**(9):1916-1926.

37. Mothe AJ, Tator CH: **Proliferation, migration, and differentiation of endogenous ependymal region stem/progenitor cells following minimal spinal cord injury in the adult rat**. *Neuroscience* 2005, **131**(1):177-187.

38. Attar A, Kaptanoglu E, Aydin Z, Ayten M, Sargon MF: **Electron microscopic study of the progeny of ependymal stem cells in the normal and injured spinal cord**. *Surgical neurology* 2005, **64 Suppl 2**:S28-32.

39. Cizkova D, Nagyova M, Slovinska L, Novotna I, Radonak J, Cizek M, Mechirova E, Tomori Z, Hlucilova J, Motlik J *et al*: **Response of ependymal progenitors to spinal cord injury or enhanced physical activity in adult rat**. *Cellular and molecular neurobiology* 2009, **29**(6-7):999-1013.

40. Moreno-Manzano V, Rodriguez-Jimenez FJ, Garcia-Rosello M, Lainez S, Erceg S, Calvo MT, Ronaghi M, Lloret M, Planells-Cases R, Sanchez-Puelles JM *et al*: **Activated spinal cord ependymal stem cells rescue neurological function**. *Stem cells* 2009, **27**(3):733-743.

41. Sabelstrom H, Stenudd M, Reu P, Dias DO, Elfineh M, Zdunek S, Damberg P, Goritz C, Frisen J: **Resident neural stem cells restrict tissue damage and neuronal loss after spinal cord injury in mice**. *Science* 2013, **342**(6158):637-640.

42. Shihabuddin LS, Horner PJ, Ray J, Gage FH: **Adult spinal cord stem cells generate neurons after transplantation in the adult dentate gyrus**. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 2000, **20**(23):8727-8735.

43. Martens DJ, Seaberg RM, Van Der Kooy D: **In vivo infusions of exogenous growth factors into the fourth ventricle of the adult mouse brain increase the proliferation of neural progenitors around the fourth ventricle and the central canal of the spinal cord**. *European Journal of Neuroscience* 2002, **16**(6):1045-1057.

44. Chen J, Knowles HJ, Hebert JL, Hackett BP: **Mutation of the mouse hepatocyte nuclear factor/forkhead homologue 4 gene results in an absence of cilia and random left-right asymmetry**. *The Journal of clinical investigation* 1998, **102**(6):1077-1082.

45. Blatt EN, Yan XH, Wuerffel MK, Hamilos DL, Brody SL: **Forkhead transcription factor HFH-4 expression is temporally related to ciliogenesis**. *American journal of respiratory cell and molecular biology* 1999, **21**(2):168-176.

46. Yu X, Ng CP, Habacher H, Roy S: **Foxj1 transcription factors are master regulators of the motile ciliogenic program**. *Nature genetics* 2008, **40**(12):1445-1453.

47. Vij S, Rink JC, Ho HK, Babu D, Eitel M, Narasimhan V, Tiku V, Westbrook J, Schierwater B, Roy S: **Evolutionarily ancient association of the FoxJ1 transcription factor with the motile ciliogenic program**. *PLoS genetics* 2012, **8**(11):e1003019.

48. Jacquet BV, Salinas-Mondragon R, Liang H, Therit B, Buie JD, Dykstra M, Campbell K, Ostrowski LE, Brody SL, Ghashghaei HT: **FoxJ1-dependent gene expression is required for differentiation of radial glia into ependymal cells and a subset of astrocytes in the postnatal brain**. *Development* 2009, **136**(23):4021-4031.

49. Schilling TF: **The morphology of larval and adult zebrafish**. *Zebrafish* 2002(261):59-94.

50. Engeszer RE, Patterson LB, Rao AA, Parichy DM: **Zebrafish in the wild: a review of natural history and new notes from the field**. *Zebrafish* 2007, **4**(1):21-40.

51. Briggs JP: **The zebrafish: a new model organism for integrative physiology**. *American journal of physiology Regulatory, integrative and comparative physiology* 2002, **282**(1):R3-9.

52. Brand MG, Nüsslein-Volhard, C: **Keeping and raising zebrafish**. In: *Zebrafish: A Practical Approach.* 1st edition edn: Oxford University Press; 2002.

53. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF: **Stages of embryonic development of the zebrafish**. *Developmental dynamics : an official publication of the American Association of Anatomists* 1995, **203**(3):253-310.

54. Lieschke GJ, Currie PD: **Animal models of human disease: zebrafish swim into view**. *Nature reviews Genetics* 2007, **8**(5):353-367.

55. Streisinger G, Walker C, Dower N, Knauber D, Singer F: **Production of clones of homozygous diploid zebra fish (Brachydanio rerio)**. *Nature* 1981, **291**(5813):293-296.

56. Hasley AO: **Exploring questions in early vertebrate development using the zebrafish (Danio rerio): From cleavage patterns to transcriptomes**. The University Of Wisconsin-Madison; 2016.

57. Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, Amora R, Hocking TD, Zhang L, Rebar EJ: **Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases**. *Nature biotechnology* 2008, **26**(6):702-708.

58. Sander JD, Cade L, Khayter C, Reyon D, Peterson RT, Joung JK, Yeh JRJ: **Targeted gene disruption in somatic zebrafish cells using engineered TALENs**. *Nature biotechnology* 2011, **29**(8):697.

59. Hwang WY, Fu Y, Reyon D, Maeder ML, Kaini P, Sander JD, Joung JK, Peterson RT, Yeh JRJ: **Heritable and precise zebrafish genome editing using a CRISPR-Cas system**. *PloS one* 2013, **8**(7):e68708.

60. Ekker SC: **Morphants: a new systematic vertebrate functional genomics approach**. *Yeast* 2000, **17**(4):302-306.

61. Kawakami K: **Transposon tools and methods in zebrafish**. *Developmental dynamics : an official publication of the American Association of Anatomists* 2005, **234**(2):244-254.

62. Roberts JA, Miguel-Escalada I, Slovik KJ, Walsh KT, Hadzhiev Y, Sanges R, Stupka E, Marsh EK, Balciuniene J, Balciunas D *et al*: **Targeted transgene integration overcomes variability of position effects in zebrafish**. *Development* 2014, **141**(3):715-724.

63. Soroldoni D, Hogan BM, Oates AC: **Simple and efficient transgenesis with meganuclease constructs in zebrafish**. *Methods in molecular biology* 2009, **546**:117-130.

64. Pan X, Wan H, Chia W, Tong Y, Gong Z: **Demonstration of site-directed recombination in transgenic zebrafish using the Cre/loxP system**. *Transgenic research* 2005, **14**(2):217-223.

65. Scheer N, Campos-Ortega JA: **Use of the Gal4-UAS technique for targeted gene expression in the zebrafish**. *Mechanisms of development* 1999, **80**(2):153-158.

66. Shoji W, Sato-Maeda M: **Application of heat shock promoter in transgenic zebrafish**. *Development, growth & differentiation* 2008, **50**(6):401-406.

67. Jopling C, Sleep E, Raya M, Marti M, Raya A, Izpisua Belmonte JC: **Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation**. *Nature* 2010, **464**(7288):606-609.

68. Emelyanov A, Parinov S: **Mifepristone-inducible LexPR system to drive and control gene expression in transgenic zebrafish**. *Developmental biology* 2008, **320**(1):113-121.

69. Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Matthews L *et al*: **The zebrafish reference genome sequence and its relationship to the human genome**. *Nature* 2013, **496**(7446):498-503.

70. Santoriello C, Zon LI: **Hooked! Modeling human disease in zebrafish**. *The Journal of clinical investigation* 2012, **122**(7):2337-2343.

71. Seth A, Stemple DL, Barroso I: **The emerging use of zebrafish to model metabolic disease**. *Disease Models and Mechanisms* 2013, **6**(5):1080-1088.

72. Sehnert AJ, Stainier DY: **A window to the heart: can zebrafish mutants help us understand heart disease in humans?** *Trends in genetics* 2002, **18**(10):491-494.

73. Jing L, Zon LI: **Zebrafish as a model for normal and malignant hematopoiesis**. *Disease Models and Mechanisms* 2011, **4**(4):433-438.

74. Berger J, Currie PD: **Zebrafish models flex their muscles to shed light on muscular dystrophies**. *Disease Models and Mechanisms* 2012, **5**(6):726-732.

75. Xi Y, Noble S, Ekker M: **Modeling neurodegeneration in zebrafish**. *Current neurology and neuroscience reports* 2011, **11**(3):274-282.

76. White R, Rose K, Zon L: **Zebrafish cancer: the state of the art and the path forward**. *Nature reviews Cancer* 2013, **13**(9):624-636.

77. MacRae CA, Peterson RT: **Zebrafish as tools for drug discovery**. *Nature Reviews Drug Discovery* 2015, **14**(10):721-731.

78. Brockes JP, Kumar A: **Comparative aspects of animal regeneration**. *Annual review of cell and developmental biology* 2008, **24**:525-549.

79. Trembley A: **Mémoires pour servir à l'histoire d'un genre de polypes d'eau douce, à bras en forme de cornes. Par A. Trembley**: chez Jean and Herman Verbeek; 1744.

80. Iten LE, Bryant SV: **Forelimb regeneration from different levels of amputation in the newt, Notophthalmus viridescens: length, rate, and stages**. *Wilhelm Roux'Archiv für Entwicklungsmechanik der Organismen* 1973, **173**(4):263-282.

81. Morgan TH: **Regeneration in teleosts**. *Development Genes and Evolution* 1900, **10**(1):120-134.

82. Gemberling M, Bailey TJ, Hyde DR, Poss KD: **The zebrafish as a model for complex tissue regeneration**. *Trends in genetics : TIG* 2013, **29**(11):611-620.

83. Poss KD: **Getting to the heart of regeneration in zebrafish**. *Seminars in cell & developmental biology* 2007, **18**(1):36-45.

84. Zhou W, Boucher RC, Bollig F, Englert C, Hildebrandt F: **Characterization of mesonephric development and regeneration using transgenic zebrafish**. *American journal of physiology Renal physiology* 2010, **299**(5):F1040-1047.

85. Moss JB, Koustubhan P, Greenman M, Parsons MJ, Walter I, Moss LG: **Regeneration of the pancreas in adult zebrafish**. *Diabetes* 2009, **58**(8):1844-1851.

86. He J, Lu H, Zou Q, Luo L: **Regeneration of liver after extreme hepatocyte loss occurs mainly via biliary transdifferentiation in zebrafish**. *Gastroenterology* 2014, **146**(3):789-800. e788.

87. Harris JA, Cheng AG, Cunningham LL, MacDonald G, Raible DW, Rubel EW: **Neomycin-induced hair cell death and rapid regeneration in the lateral line of zebrafish (Danio rerio)**. *Journal of the Association for Research in Otolaryngology* 2003, **4**(2):219-234.

88. Becker CG, Becker T: **Adult zebrafish as a model for successful central nervous system regeneration**. *Restorative neurology and neuroscience* 2008, **26**(2-3):71-80.

89. Becker T, Wullimann MF, Becker CG, Bernhardt RR, Schachner M: **Axonal regrowth after spinal cord transection in adult zebrafish**. *The Journal of comparative neurology* 1997, **377**(4):577-595.

90. Hui SP, Dutta A, Ghosh S: **Cellular response after crush injury in adult zebrafish spinal cord**. *Developmental dynamics : an official publication of the American Association of Anatomists* 2010, **239**(11):2962-2979.

91. Reimer MM, Sorensen I, Kuscha V, Frank RE, Liu C, Becker CG, Becker T: **Motor neuron regeneration in adult zebrafish**. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 2008, **28**(34):8510-8516.

92. Ferretti P, Prasongchean W: **Adult neurogenesis and regeneration: focus on nonmammalian vertebrates**. In: *Neural Stem Cells in Development, Adulthood and Disease.* Springer; 2015: 1-21.

93. Hui SP, Nag TC, Ghosh S: **Characterization of Proliferating Neural Progenitors after Spinal Cord Injury in Adult Zebrafish**. *PloS one* 2015, **10**(12):e0143595.

94. Goldshmit Y, Sztal TE, Jusuf PR, Hall TE, Nguyen-Chi M, Currie PD: **Fgf-dependent glial cell bridges facilitate spinal cord regeneration in zebrafish**. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 2012, **32**(22):7477-7492.

95. Guo Y, Ma L, Cristofanilli M, Hart RP, Hao A, Schachner M: **Transcription factor Sox11b is involved in spinal cord regeneration in adult zebrafish**. *Neuroscience* 2011, **172**:329-341.

96. Kuscha V, Frazer SL, Dias TB, Hibi M, Becker T, Becker CG: **Lesion-induced generation of interneuron cell types in specific dorsoventral domains in the spinal cord of adult zebrafish**. *The Journal of comparative neurology* 2012, **520**(16):3604-3616.

97. Ogai K, Nakatani K, Hisano S, Sugitani K, Koriyama Y, Kato S: **Function of Sox2 in ependymal cells of lesioned spinal cords in adult zebrafish**. *Neuroscience research* 2014, **88**:84-87.

98. Barami K, Iversen K, Furneaux H, Goldman SA: **Hu protein as an early marker of neuronal phenotypic differentiation by subependymal zone cells of the adult songbird forebrain**. *Journal of neurobiology* 1995, **28**(1):82-101.

99. Reimer MM, Kuscha V, Wyatt C, Sorensen I, Frank RE, Knuwer M, Becker T, Becker CG: **Sonic Hedgehog Is a Polarized Signal for Motor Neuron Regeneration in Adult Zebrafish**. *Journal of Neuroscience* 2009, **29**(48):15073-15082.

100. Pinto L, Gotz M: **Radial glial cell heterogeneity--the source of diverse progeny in the CNS**. *Progress in neurobiology* 2007, **83**(1):2-23.

101. Reimer MM, Norris A, Ohnmacht J, Patani R, Zhong Z, Dias TB, Kuscha V, Scott AL, Chen YC, Rozov S *et al*: **Dopamine from the brain promotes spinal motor neuron generation during development and adult regeneration**. *Developmental cell* 2013, **25**(5):478-491.

102. Stubbs JL, Oishi I, Izpisua Belmonte JC, Kintner C: **The forkhead protein Foxj1 specifies node-like cilia in Xenopus and zebrafish embryos**. *Nature genetics* 2008, **40**(12):1454-1460.

103. Hellman NE, Liu Y, Merkel E, Austin C, Le Corre S, Beier DR, Sun Z, Sharma N, Yoder BK, Drummond IA: **The zebrafish foxj1a transcription factor regulates cilia function in response to injury and epithelial stretch**. *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**(43):18499-18504.

104. Aamar E, Dawid IB: **Isolation and expression analysis of foxj1 and foxj1.2 in zebrafish embryos**. *The International journal of developmental biology* 2008, **52**(7):985-991.

105. Heyman RA, Borrelli E, Lesley J, Anderson D, Richman DD, Baird SM, Hyman R, Evans RM: **Thymidine kinase obliteration: creation of transgenic mice with controlled immune deficiency**. *Proceedings of the National Academy of Sciences of the United States of America* 1989, **86**(8):2698-2702.

106. Feil R, Wagner J, Metzger D, Chambon P: **Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains**. *Biochemical and biophysical research communications* 1997, **237**(3):752-757.

107. Mosimann C, Zon LI: **Advanced zebrafish transgenesis with Tol2 and application for Cre/lox recombination experiments**. *Methods in cell biology* 2011, **104**:173-194.

108. Mullins MC, Nusslein-Volhard C: **Mutational approaches to studying embryonic pattern formation in the zebrafish**. *Current opinion in genetics & development* 1993, **3**(4):648-654.

109. Hans S, Freudenreich D, Geffarth M, Kaslin J, Machate A, Brand M: **Generation of a non-leaky heat shock-inducible Cre line for conditional Cre/lox strategies in zebrafish**. *Developmental dynamics : an official publication of the American Association of Anatomists* 2011, **240**(1):108-115.

110. Mosimann C, Kaufman CK, Li P, Pugach EK, Tamplin OJ, Zon LI: **Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish**. *Development* 2011, **138**(1):169-177.

111. Kroehne V, Freudenreich D, Hans S, Kaslin J, Brand M: **Regeneration of the adult zebrafish brain from neurogenic radial glia-type progenitors**. *Development* 2011, **138**(22):4831-4841.

112. Caron A, Xu X, Lin X: **Wnt/beta-catenin signaling directly regulates Foxj1 expression and ciliogenesis in zebrafish Kupffer's vesicle**. *Development* 2012, **139**(3):514-524.

113. Moro E, Gnugge L, Braghetta P, Bortolussi M, Argenton F: **Analysis of beta cell proliferation dynamics in zebrafish**. *Developmental biology* 2009, **332**(2):299-308.

114. Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR: **Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension**. *Gene* 1989, **77**(1):61-68.

115. Heckman KL, Pease LR: **Gene splicing and mutagenesis by PCR-driven overlap extension**. *Nature protocols* 2007, **2**(4):924-932.

116. Kawakami K: **Tol2: a versatile gene transfer vector in vertebrates**. *Genome biology* 2007, **8 Suppl 1**:S7.

117. Kawakami K, Takeda H, Kawakami N, Kobayashi M, Matsuda N, Mishina M: **A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish**. *Developmental cell* 2004, **7**(1):133-144.

118. Urasaki A, Morvan G, Kawakami K: **Functional dissection of the Tol2 transposable element identified the minimal cis-sequence and a highly repetitive sequence in the subterminal region essential for transposition**. *Genetics* 2006, **174**(2):639-649.

119. Suster ML, Sumiyama K, Kawakami K: **Transposon-mediated BAC transgenesis in zebrafish and mice**. *BMC genomics* 2009, **10**:477.

120. Xu Q: **Microinjection into Zebrafish Embryos**. In: *Molecular Methods in Developmental Biology: Xenopus and Zebrafish.* Edited by Guille M. Totowa, NJ: Humana Press; 1999: 125-132.

121. Gilmour DT, Jessen JR, Lin S: **Manipulating gene expression in the zebrafish**. *Zebrafish: A Practical Approach* 2002, **261**:121-143.

122. Meeker ND, Hutchinson SA, Ho L, Trede NS: **Method for isolation of PCR-ready genomic DNA from zebrafish tissues**. *BioTechniques* 2007, **43**(5):610, 612, 614.

123. Westerfield M: **A guide for the laboratory use of zebrafish (Danio rerio)**. In: *The zebrafish book.* University of Oregon Press Eugene; 2000: 4-5.2.

124. Neugebauer JM, Amack JD, Peterson AG, Bisgrove BW, Yost HJ: **FGF signalling during embryo development regulates cilia length in diverse epithelia**. *Nature* 2009, **458**(7238):651-654.

125. Matsuda T, Cepko CL: **Controlled expression of transgenes introduced by in vivo electroporation**. *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104**(3):1027-1032.

126. Briona LK, Dorsky RI: **Spinal cord transection in the larval zebrafish**. *Journal of visualized experiments : JoVE* 2014(87).

127. ZernickaGoetz M, Pines J, Ryan K, Siemering KR, Haseloff J, Evans MJ, Gurdon JB: **An indelible lineage marker for Xenopus using a mutated green fluorescent protein**. *Development* 1996, **122**(12):3719-3724.

128. Notredame C, Higgins DG, Heringa J: **T-Coffee: A novel method for fast and accurate multiple sequence alignment**. *Journal of molecular biology* 2000, **302**(1):205-217.

129. Stothard P: **The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences**. *BioTechniques* 2000, **28**(6):1102, 1104-1102, 1104.

130. Nakamura Y, Gojobori T, Ikemura T: **Codon usage tabulated from international DNA sequence databases: status for the year 2000**. *Nucleic acids research* 2000, **28**(1):292-292.

131. Cardillo G: **MyFisher: the definitive function for the Fisher's exact**. In.: http://www.mathworks.com/matlabcentral/fileexchange/26883; 2010.

132. Lele Z, Engel S, Krone PH: **hsp47 and hsp70 gene expression is differentially regulated in a stress- and tissue-specific manner in zebrafish embryos**. *Developmental genetics* 1997, **21**(2):123-133.

133. Provost E, Rhee J, Leach SD: **Viral 2A peptides allow expression of multiple proteins from a single ORF in transgenic zebrafish embryos**. *Genesis* 2007, **45**(10):625-629.

134. Snapp E: **Design and use of fluorescent fusion proteins in cell biology**. *Current Protocols in Cell Biology* 2005:21-4.

135. de Almeida AIL: **An approach to molecular genetics of thyroid cancer: from novel mutations to a zebrafish model**. 2016.

136. Horstick EJ, Jordan DC, Bergeron SA, Tabor KM, Serpe M, Feldman B, Burgess HA: **Increased functional protein expression using nucleotide sequence features enriched in highly expressed genes in zebrafish**. *Nucleic acids research* 2015, **43**(7):e48.

137. Lodish HF: **Molecular cell biology**, 4th edn. New York: W.H. Freeman; 2000.

138. Kozak M: **Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes**. *Cell* 1986, **44**(2):283-292.

139. Plotkin JB, Robins H, Levine AJ: **Tissue-specific codon usage and the expression of human genes**. *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101**(34):12588-12591.

140. Hoess RH, Abremski K: **Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system**. *Journal of molecular biology* 1985, **181**(3):351-362.

141. Thummel R, Burket CT, Brewer JL, Sarras MP, Jr., Li L, Perry M, McDermott JP, Sauer B, Hyde DR, Godwin AR: **Cre-mediated site-specific recombination in zebrafish embryos**. *Developmental dynamics : an official publication of the American Association of Anatomists* 2005, **233**(4):1366-1377.

142. Hans S, Kaslin J, Freudenreich D, Brand M: **Temporally-controlled site-specific recombination in zebrafish**. *PloS one* 2009, **4**(2):e4640.

143. Kaufman OH, Marlow FL: **Methods to study maternal regulation of germ cell specification in zebrafish**. *Methods in cell biology* 2016, **134**:1-32.

144. Kawakami K, Asakawa K, Muto A, Wada H: **Tol2-mediated transgenesis, gene trapping, enhancer trapping, and Gal4-UAS system**. *Methods in cell biology* 2016, **135**:19-37.

145. Stuart GW, McMurray JV, Westerfield M: **Replication, integration and stable germ-line transmission of foreign sequences injected into early zebrafish embryos**. *Development* 1988, **103**(2):403-412.

146. Akitake CM, Macurak M, Halpern ME, Goll MG: **Transgenerational analysis of transcriptional silencing in zebrafish**. *Developmental biology* 2011, **352**(2):191-201.

147. Duret L: **Evolution of synonymous codon usage in metazoans**. *Current opinion in genetics & development* 2002, **12**(6):640-649.

148. Ikemura T: **Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the E. coli translational system**. *Journal of molecular biology* 1981, **151**(3):389-409.

149. Sorensen MA, Kurland CG, Pedersen S: **Codon usage determines translation rate in Escherichia coli**. *Journal of molecular biology* 1989, **207**(2):365-377.

150. Puigbo P, Guzman E, Romeu A, Garcia-Vallve S: **OPTIMIZER: a web server for optimizing the codon usage of DNA sequences**. *Nucleic acids research* 2007, **35**(Web Server issue):W126-131.

151. Kozak M: **An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs**. *Nucleic acids research* 1987, **15**(20):8125-8148.

152. Grzegorski SJ, Chiari EF, Robbins A, Kish PE, Kahana A: **Natural variability of Kozak sequences correlates with function in a zebrafish model**. *PloS one* 2014, **9**(9):e108475.

153. Connelly S, Manley JL: **A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II**. *Genes & development* 1988, **2**(4):440-452.

154. Keller W, Minvielle-Sebastia L: **A comparison of mammalian and yeast pre-mRNA 3'-end processing**. *Current opinion in cell biology* 1997, **9**(3):329-336.

155. Libri D, Dower K, Boulay J, Thomsen R, Rosbash M, Jensen TH: **Interactions between mRNA export commitment, 3′-end quality control, and nuclear degradation**. *Molecular and cellular biology* 2002, **22**(23):8254-8266.

156. Williams DW, Muller F, Lavender FL, Orban L, Maclean N: **High transgene activity in the yolk syncytial layer affects quantitative transient expression assays in zebrafish *Danio rerio*) embryos**. *Transgenic research* 1996, **5**(6):433-442.

157. Hu G, Goll MG, Fisher S: **PhiC31 integrase mediates efficient cassette exchange in the zebrafish germline**. *Developmental dynamics : an official publication of the American Association of Anatomists* 2011, **240**(9):2101-2107.

158. Mosimann C, Puller AC, Lawson KL, Tschopp P, Amsterdam A, Zon LI: **Site-directed zebrafish transgenesis into single landing sites with the phiC31 integrase system**. *Developmental Dynamics* 2013, **242**(8):949-963.

159. Jungke P, Hammer J, Hans S, Brand M: **Isolation of Novel CreERT2-Driver Lines in Zebrafish Using an Unbiased Gene Trap Approach**. *PloS one* 2015, **10**(6):e0129072.

160. Pinzon-Olejua A, Welte C, Chekuru A, Bosak V, Brand M, Hans S, Stuermer CA: **Cre-inducible site-specific recombination in zebrafish oligodendrocytes**. *Developmental dynamics : an official publication of the American Association of Anatomists* 2016.

161. Felker A, Mosimann C: **Contemporary zebrafish transgenesis with Tol2 and application for Cre/lox recombination experiments**. *Methods in cell biology* 2016, **135**:219-244.

162. Kikuchi K, Holdway JE, Werdich AA, Anderson RM, Fang Y, Egnaczyk GF, Evans T, Macrae CA, Stainier DY, Poss KD: **Primary contribution to zebrafish heart regeneration by gata4(+) cardiomyocytes**. *Nature* 2010, **464**(7288):601-605.

163. Wolffe AP, Matzke MA: **Epigenetics: regulation through repression**. *Science* 1999, **286**(5439):481-486.

164. Lenhard B, Sandelin A, Carninci P: **Metazoan promoters: emerging characteristics and insights into transcriptional regulation**. *Nature reviews Genetics* 2012, **13**(4):233-245.

165. Lee TI, Young RA: **Transcriptional regulation and its misregulation in disease**. *Cell* 2013, **152**(6):1237-1251.

166. Huang CJ, Tu CT, Hsiao CD, Hsieh FJ, Tsai HJ: **Germ-line transmission of a myocardium-specific GFP transgene reveals critical regulatory elements in the cardiac myosin light chain 2 promoter of zebrafish**. *Developmental dynamics : an official publication of the American Association of Anatomists* 2003, **228**(1):30-40.

Appendix

I. Restriction enzymes

Table S.1 Designation and features of restriction enzymes used for molecular cloning steps

Enzyme	Restriction site (5' → 3')	Cleavage ends	Incubation temperature	Brand
ApaI	GGGCC^C	Cohesive	25ºC	Promega
AscI	GG^CGCGCC	Cohesive	37ºC	NEB
BamHI	G^GATCC	Cohesive	37ºC	NEB
ClaI	AT^CGAT	Cohesive	37ºC	NEB
FseI	GGCCGG^CC	Cohesive	37ºC	NEB
HindIII	A^ACGTT	Cohesive	37ºC	NEB
SmaI	CCC^GGG	Blunt	25ºC	NEB
XhoI	C^TCGAG	Cohesive	37ºC	Promega

II. Microinjection

Table S.2 Microinjection conditions for each line

Construct	Injected into line…	[DNA] (ng.µL^{-1})	[mRNA] (ng.µL^{-1})
foxj1a:CreER^{T2}; cmlc2:EGFP (generated by José Ramalho, FMUNL)	*Tg(ubi:Switch)*	30	5
foxj1a:mCherry-T2a-CreER^{T2}	*Tg(hs:Red2Green*)	10, 20, 25, 30, 33, 37.5	5, 10, 15
foxj1a:DsRed2(floxed)-EGFP	*Tg(hs:mCh-Cre*)	30, 33, 34, 35, 37	5, 6, 7, 15, 20
foxj1a:TK-GFP	AB	30, 32, 33, 35	5, 6, 7, 15
foxj1a:TK-T2a-EGFP	AB	25, 30, 35	5, 10, 25, 27.5
ef1α:DsRed2(floxed)-ntr-EGFP	AB	30	10
foxj1a:EGFP	AB	25	5

III. Oligonucleotide primer sequences

Table S.3 Designation and features of oligonucleotides used for PCR amplification and/or sequencing

A. Primers for DNA amplification

Reference number	Designation	Sequence (5'→ 3')	Tm (ºC)	Features
A1*	foxj1a-For-ApaI	CGCGGGGCCCTGTGTGTGTGTTTGAGAAGC	58	5': CGCG; ApaI restriction site
A2*	foxj1a-Rev-FseI	GCGCGGCCGGCCAGTGGTAGTTGATGGTCAGCAG	66	5': GCGC; FseI restriction site
A3*	EGFP-For-SmaI	CGCGCCCGGGATGGTGAGCAAGGGCGAGGAGC	72	5': CGCG; SmaI restriction site
A4*	EGFP-Rev-AscI	GCGCGGCGCGCCTTACTTGTACAGCTCGTCCATGC	68	5':GCGC; AscI restriction site
A5*	foxj1a-For-XhoI	CGCGCTCGAGTGTGTGTGTGTTTGAGAAGC	58	5': CGCG; XhoI restriction site
A6*	foxj1a-For-HindIII	CGCGAAGCTTTGTGTGTGTGTTTGAGAAGC	58	5': CGCG; HindIII restriction site
A7*	foxj1a-Rev-BamHI	GCGCGGATCCAGTGGTAGTTGATGGTCAGCAG	66	5': GCGC; BamHI restriction site
A8	mmGFP-Rev-AscI	GCGCGGCGCGCCTTATTTGTATAGTTCATCCA	50	5': GCGC; AscI restriction site
A9	TK-For-BamHI	CGCGGATCCAAACATGGCTTCGTACCCCGGC	60	5': CGC; BamHI restriction site
A10	TK-Rev-T2a	AGAAGACTTCCTCTGCCCTCTCCGTTAGCCTCCCCCATC	62	5': 20nt complementary to T2a peptide (fusion PCR)
A11	T2a-For-TK	AGATGGGGGAGGCTAACGGAGAGGGCAGAGGAAGTCTTC	60	5': 20nt complementary to *TK* (fusion PCR)

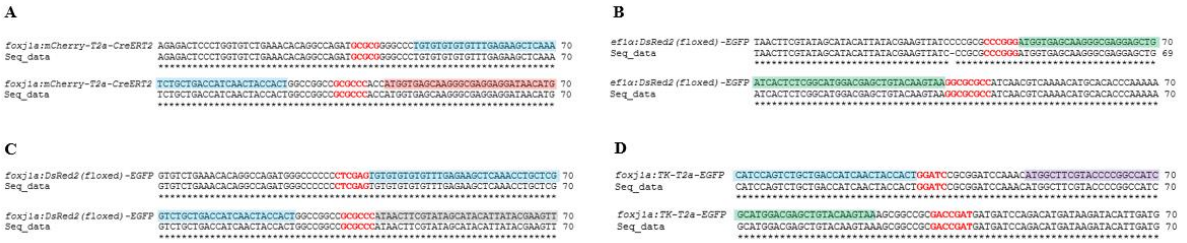
A12	T2a-Rev-EGFP	TCCTCGCCCTTGCTCACCATAGGGCCGGGATTCTCCTC	60	5': 20nt complementary to *EGFP* (fusion PCR)
A13	EGFP-For-T2a	TGGAGGAGAATCCCGGCCCTATGGTGAGCAAGGGCGAGG	62	5': 20nt complementary to T2a peptide (fusion PCR)
A14	EGFP-Rev-NotI	GTCGCGGCCGCTTTACTTG	62	5': GTC; NotI restriction site
A15	Cre-For-1	CGCGGGCTGGAGACATGAGAGCTG	64	5': CGCG
A16	Cre-For-2	CGCGTATATCTTCAGGCGCGCGGT	62	5': CGCG
A17	Cre-Rev-2	GCGCATCAAGCTGTGGCAGGGAAA	60	5': GCGC
A18	βactin-For	TGGCATTGCTGACCGTATGCAG	68	--
A19	βactin-Rev	ACTCCTGCTTGCTGATCCACATC	70	--

B. Primers for sequencing

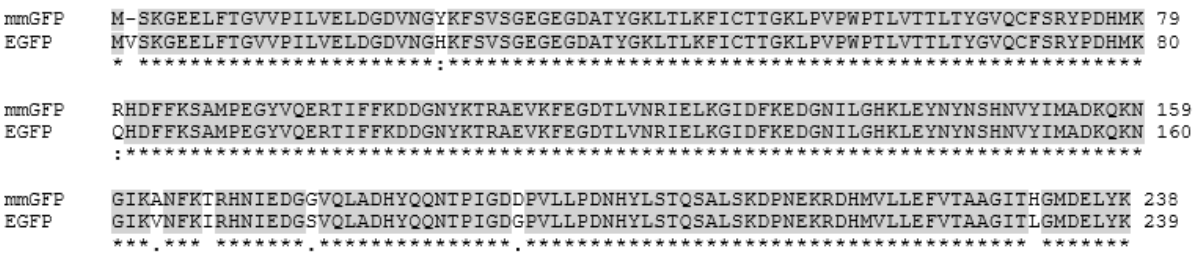
Reference number	Designation	Sequence (5'→ 3')	Tm (ºC)	Features
B1	TK-For-497	ATCGCCGCCCTCCTGTGCTA	66	--
B2	mmGFP-For	ATGAGTAAAGGAGAAGAAC	52	--

Note: all amplification primers marked with * were also used for sequencing purposes.

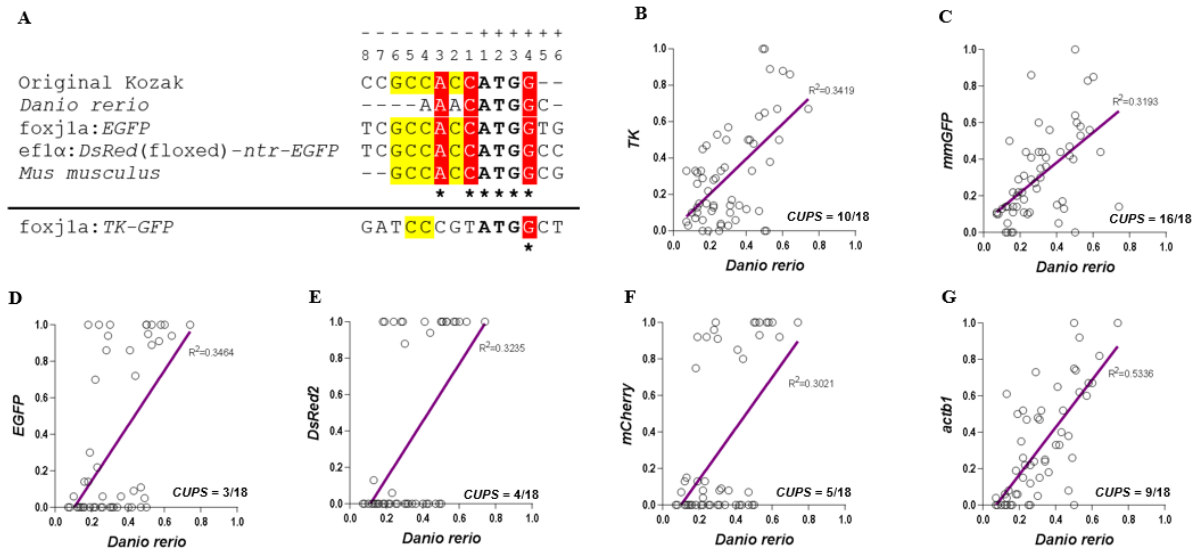
IV. Supplementary Figures



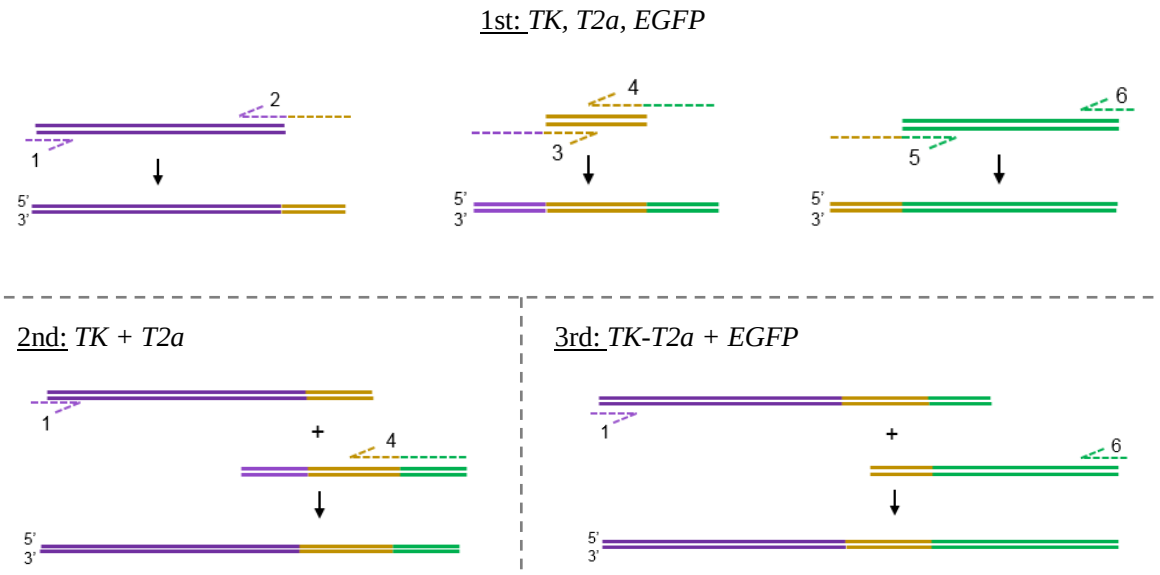
Supplementary Figure S.1 Confirmation of major molecular cloning steps by sequencing. **A:** Sequencing data from clone #3 confirming a successful ligation with a perfect alignment (*) to the predicted sequence of pTol(*foxj1a:mCherry-T2a-CreER^{T2}*) in both 5' (top) and 3' (bottom) insert ends. Blue shading indicates beginning (top) and end (bottom) of *foxj1a* promoter, red shading indicates beginning of *mCherry* coding sequence. **B:** Sequencing data from a positive clone confirming a successful ligation of *EGFP* coding sequence with a perfect alignment (*) of the ligation site to the predicted sequence of pTol(*ef1α:DsRed2(floxed)-EGFP*) in both 5' (top) and 3' (bottom) insert ends. Green shading indicates beginning (top) and end (bottom) of *EGFP* sequence. **C:** Sequencing data from clone #7 confirming a successful ligation with a perfect alignment (*) to the predicted sequence of pTol(*foxj1a:DsRed2(floxed)-EGFP*) in both 5' (top) and 3' (bottom) insert ends. Blue shading indicates beginning (top) and end (bottom) of *foxj1a* promoter, grey shading indicates beginning of loxP sequence. **D:** Sequencing data from clone #11 confirming a successful ligation with a perfect alignment (*) to the predicted sequence of pTol(*foxj1a:TK-T2a-EGFP*) in both 5' (top) and 3' (bottom) insert ends. Blue shading indicates beginning of *foxj1a* promoter, purple shading indicates beginning of *TK* coding sequence, green shading indicates end of *EGFP* coding sequence. In all alignments red sequences indicate ligation sites.



Supplementary Figure S.2 Protein alignment between mmGFP and EGFP aminoacid sequences reveal exact matches (*) for all but 8 residues. (:) marks aminoacids of same size and hydropathy, (.) marks aminoacids of similar size or evolutionary preserved hydropathy.



Supplementary Figure S.3 Sequence features influencing expression of transgenic proteins. **A:** Unlike other injected constructs, (*foxj1a:TK-GFP*) does not display a Kozak sequence similar to the original one or the zebrafish natural variant. *D. rerio*, *M. musculus* and original Kozak sequences obtained from [152]. Reference nucleotide positions relative to start codon are depicted above the alignment. **B-G:** Relative codon usage for selected genes used in this work versus *D. rerio* protein coding genes with linear regression displayed as a purple line. Each circle represents one codon; Met, Trp and stop codons not included. CUPS score reads as fraction of aminoacids with a similar codon usage between the set of genes in question.



Supplementary Figure S.4 Schematic representation of the fusion PCR reactions leading to the assembly of (*TK-T2a-EGFP*) coding sequence. 1-6 depict individual primers used (see M&M section for details).