The Toolbox for Conditional Zebrafish Cancer Models

Marie Mayrhofer and Marina Mione

Abstract Here we describe the conditional zebrafish cancer toolbox, which allows for fine control of the expression of oncogenes or downregulation of tumor suppressors at the spatial and temporal level. Methods such as the Gal4/UAS or the Cre/lox systems paved the way to the development of elegant tumor models, which are now being used to study cancer cell biology, clonal evolution, identification of cancer stem cells and anti-cancer drug screening. Combination of these tools, as well as novel developments such as the promising genome editing system through CRISPR/Cas9 and clever application of light reactive proteins will enable the development of even more sophisticated zebrafish cancer models. Here, we introduce this growing toolbox of conditional transgenic approaches, discuss its current application in zebrafish cancer models and provide an outlook on future perspectives.

Keywords Cancer models • Cre/lox • Gal4/UAS • Heat shock • Inducible systems • ER/tamoxifen • Tet-On • Transgenic methods • Zebrafish

Abbreviations

4-OHT 4-Hydroxy-Tamoxifen
AD Activation Domain
CRISPR/Cas9 Clustered Regularly Interspaced, Short Palindromic Repeats/CRISPR Associated System


M. Mayrhofer • M. Mione (*)
Institute of Toxicology and Genetics, Karlsruhe Institute of Technology, Eggenstein-Leopoldshafen, Germany
e-mail: marie.spitzner@kit.edu; maria.mione@kit.edu

© Springer International Publishing Switzerland 2016
D.M. Langenau (ed.), Cancer and Zebrafish, Advances in Experimental Medicine and Biology 910, DOI 10.1007/978-3-319-30654-4_2
Introduction

As a vertebrate, the zebrafish spontaneously develops tumors and together with its fast, transparent and ex utero embryonic development attracted the attention of researchers as a disease model organism [1]. The first cancer studies used chemicals and induced mainly liver and intestinal tumors in zebrafish [2–4]. Shortly after, the first stable transgenic zebrafish line was generated in 1988 [5]. However, due to difficulties with stable germline transmission and silencing of the integrated transgenes it took several years to establish reliable transgenic protocols [6].
Thereby, application of endonucleases such as the meganuclease I-SceI and transposases, including Sleeping Beauty or the tol2 system, improved the transgenic efficiency and stability. Especially, the medaka derived tol2 system, developed by Kawakami, greatly improved transgenesis as it integrates mostly single copies through a cut-and-paste mechanism without rearrangement or modification at the target site [7]. Its unbiased integration into nearly any genomic site and reliable expression of integrated transgenes made transgenesis in zebrafish easy, fast, efficient and stable [7].

A straightforward method to model a specific cancer type in zebrafish is to generate transgenes with oncogenes placed directly under the control of a tissue specific promoter. However, strong oncogenes will induce deleterious effects that may lead to lethality before the fish reach sexual maturity thus impeding the generation of stable lines. To overcome this drawback and to improve spatial and temporal control over oncogene expression, a toolbox of conditional transgenic approaches has been established in the zebrafish that provides the potential to (1) perform isolated gene analysis in specific tissues or cell type subsets, (2) determine the time point of gene expression onset and (3) offset and (4) control the strength of oncogene expression. The bipartite Gal4/UAS system facilitated the development of a wide range of aggressive cancer models [8–13] as it allows for independent inheritance of tissue specificity through Gal4 in one parental line and silent oncogenes in the second parental line. Other conditional approaches include the Cre/lox system, that generates stable recombination events but that in zebrafish still suffers of inefficient activity leading to a mosaic pattern [14, 15], and temporal control systems such as the heat shock promoter, the tetracycline dependent Tet-On and Tet-Off system or chemically inducible hormone receptors that can be activated or inactivated upon external stimuli.

**Spatial Control**

Genetic methods for restricting the expression of oncogenes to a specific tissue or cell type are typically based on the use of tissue specific promoters. Several systems have been developed that increase flexibility, including the possibility of turning oncogene expression on and off. Importantly, conditional transgenic approaches allow maintenance of lines, which are embryonic-lethal because of embryonic activity of the promoter. These include the Gal4/UAS system and its derivatives, the site specific recombinases (CRE/lox system & Co) and more recently, genome editing through CRISPR/Cas9.

**The Gal4/UAS System**

The Gal4/UAS system is one of the first conditional transgenic methods applied to zebrafish [16]. This binary system, derived from the yeast *Saccharomyces cerevisiae*, consists of the transcriptional activator Gal4 that controls gene expression through its DNA-binding motif, UAS (Upstream Activating Sequence) (Fig. 1a) [17].
Fig. 1 The Gal4/UAS system and its variants enable cell type specific oncogene expression. (A) Schematic of the Gal4/UAS system applied to a 5 dpf larva melanocyte (box): tissue specific gene expression is enabled through promoter dependent Gal4 expression (orange oval) which, upon binding to UAS in the same cell, activates expression of a UAS driven transgene such as green fluorescent protein (GFP) leading to fluorescent protein production (green stars). (B–D) Expression of Gal4 under tissue specific promoters allows for UAS controlled GFP expression in different tissues as indicated by white arrows: melanocytes (kita:Gal4, B), endothelial cells (fli1:Gal4, C) and brain (zic:Gal4, D) while UAS controlled expression of the oncogene HRAS^{G12V} in these lines induces formation of melanoma (B'), leukemia (C') and glioma (D'). (E) Inhibition of Gal4 activity can be mediated by Gal80 which binds to the activation domain (AD) of Gal4. (F) In the fusion protein Gal4-VP16, the AD of Gal4 is exchanged with a VP16 variant (pink triangle) disabling Gal80 mediated Gal4 inhibition while increasing the transcriptional activity of Gal4. (G) Increased transgene expression is achieved by application of up to 5 UAS repeats, while a much higher copy number reduces transcription efficiency. Systems similar to Gal4/UAS include the Q-system (H) and the TrpR/tUAS system (I) composed of the transactivator (pink oval) QF/TrpR and the upstream activation sequence (yellow rectangle) QUAS/tUAS, respectively. The QF can be inhibited by QS (blue half circle). (J) As these systems do not cross-react they can be combined in one model to study cell non-autonomous effects for example by expressing Gal4 and QF under different cell type specific promoters to elucidate the effect of disabled Notch-signaling (through expression of dominant-negative suppressor of hairless, grey rectangle) in B-cells, and assess the effect on oncogenic behavior of T-cells. Abbreviations: AD activation domain, DBD DNA-binding domain, DNSuH dominant negative suppressor of hairless, GFP green fluorescent protein, UAS upstream activating sequence
To achieve tissue specific gene expression, the gene of interest is placed downstream of the UAS motif and the transcriptional activator Gal4 is expressed under the control of a tissue specific promoter. As the UAS motif is exclusively active in presence of Gal4, the gene of interest will be expressed exclusively in the tissue where the chosen promoter is active. Separate integration of the two components in an activator line (Gal4) and an effector line (UAS-gene x) allows for silent inheritance of the gene of interest and its conditional activation only in double transgenic offsprings from crosses between activator and effector lines. A third component of this system that has been adopted for further control is the Gal4 inhibitor Gal80. Gal80 binds and inhibits the activating domain of Gal40 and thus inhibits expression of the UAS controlled gene (Fig. 1e) [18, 19].

The Gal4/UAS system has been used in a variety of organisms including plants [20], drosophila [21], xenopus [22] and mice [23]. Scheer and Campos-Ortega were the first to apply it to zebrafish [16]. They developed activator lines that express the full-length Gal4 gene under control of one of the two ubiquitous promoters SV40/thymidine kinase (svtk) or carp β-actin and effector lines that express an active form of the Notch1 receptor under UAS control. Double transgenic offspring from crosses between the stable transgenic activator and effector lines expressed the myc-notch1a:intra in the same regions where Gal4 was expressed and thus confirmed that Gal4 can regulate the expression of a UAS controlled gene. Still the system suffered from low gene induction, low transgenic efficiency and variegated/mosaic expression. Several alterations of the constructs and discoveries of new techniques have improved the system and made it widely used in zebrafish. Specifically, the discovery of the transposase Tol2 from medaka revolutionized the genomic integration of transgenes and reduced the mosaic and variegated expression pattern of the Gal4 system and facilitated its rapid development.

Further optimization of the system aimed to increase Gal4 activity through fusion of the Gal4 DNA-binding domain (DBD) with a strong transcriptional activation domain (AD). Intensive efforts have been undertaken to optimize the transcriptional AD from the VP16 protein of the herpes simplex virus. In a first attempt the Gal4-DBD was fused to the full-length VP16 [24, 25] which improved the strength of the UAS induction and maintained the promoter activity on Gal4 expression but also revealed massive toxic effects reflected in developmental defects and high embryonic lethality [26, 27]. These effects were thought to appear due to a phenomenon called “squelching” [28, 29]. VP16, as one of the most potent activation domains known, interacts with a number of transcriptional components and thus through a titration effect can lead to non-specific inhibition of transcription [28]. Therefore, to reduce toxicity several modifications were developed to reduce toxicity while keeping Gal4 activity as strong as when it is combined with full-length VP16. Such attempts include the fusion with the truncated form VP16$_{413-470}$ that lacks the last 20 amino acids important for its transcriptional activity. Albeit it is less active than full length VP16, VP16$_{413-470}$ still induces developmental defects. Asakawa et al. used an even smaller fragment of VP16 containing the 12 amino acids 436–447 and an N-terminal prolin [30, 31]. Two copies fused to Gal4 (Gal4FF) allowed sufficient and safe Gal4 expression while a further increase in copy num-
bers would increase Gal4 efficiency but also leads to side effects [32]. Using GFP (green fluorescent protein) and the newly developed Gal4FF, Kawakami et al. [33] generated a large library of diverse gene trap and enhancer trap lines which are published and available to the zebrafish community under the name zTrap (see Box 1 for web sites). Distel et al. applied a third version of the VP16 AD, called TA4 [24]. It contains three copies of VP16_{436-447} but two of them with weakening mutations in the position 442, a position crucial for its function [34, 35]. A Kozac sequence before the Gal4 and a β-globin intron positioned 3′ of TA4 improves the translation of the mRNA and gives the construct the new name KalTA4GI. This construct generates robust non-deleterious and non-variegated transactivation in a variety of lines [34].

Unfortunately, the use of the Gal4 inhibitor Gal80 to modulate the activity of Gal4 is limited by the modifications imposed to Gal4 to increase its efficiency. As Gal80 targets the C-terminal AD of Gal4 [19], which is missing in the Gal4-VP16 fusion variants, it is unable to bind and inhibit Gal4 activity in the vast majority of existing activator lines (Fig. 1f) [36, 37]. In alternative to VP16, the carboxyl terminal AD of the protein NFkB p65 was shown to significantly improve Gal4 activity with only minor toxic effects, but this combination has not been used by the zebrafish community [32].

To further optimize the binary Gal4/UAS system, the UAS promoter has been tested for its efficiency according to the number of UAS repeats. Thereby, a nearly linear dose effect was observed when increasing the UAS number from 1 to 5 (19.5 times stronger) repeats while further increase to 10 or even 14 repeats again reduces effects to a level comparable to 4 repeats (Fig. 1g) [34]. One drawback of the system is the gene silencing occurring over time through methylation probably due to the high GC content of the UAS sequence and the use of high copy numbers. Indeed, comparison of 4xUAS with 14xUAS showed that low copy number UAS constructs are less vulnerable to methylation induced silencing [38].

Since its development the Gal4/UAS system has been used for a variety of diverse zebrafish studies ranging from studies on early developmental processes to investigation of organization and function of individual brain structures [39–42].

The high flexibility of the Gal4/UAS system has promoted a number of cancer studies in zebrafish. In our lab, different cancer models have been established that utilize the same UAS controlled oncogene HRAS^G12V fused to GFP, but depend on different Gal4 lines. Gal4 expression under control of the kit/ or fli1 promoter for example induces development of melanoma [9] or leukemia [43], respectively. Remarkably, while the melanoma model develops aggressive tumors already in 20% of 1 month old double transgenic fish, most of them survive to fecundity, which allows to keep the line in a double transgenic state [9]. The developing melanomas resemble the human phenotype not only histologically through infiltrative behavior and polyploid nuclei but also immunologically through expression of human melanoma markers such as Tyrosinase, Melan-a, s100 and HMB45. The double transgenic larvae are also characterized by a hyperpigmentation phenotype at 3 dpf, which make them ideal for chemical screens. The leukemia model instead is highly lethal in double transgenic larvae and requires tol2 mediated
mosaic integration of the oncogene construct in somatic cells to allow development of leukemia in juveniles [43]. In larvae the disease is characterized by prominent expansion of the caudal hematopoietic tissue and disruption of the vascular system and in adults appears through an increased number of immature cells and arrest of myeloid maturation in kidney marrow. Applying the same GFP fused $HRAS^{G12V}$ oncogene, two further studies established novel cancer models. Burger and colleagues describe the first chordoma model through Gal4/UAS controlled $HRAS^{G12V}$ expression under the notochord prominent promoter tiggywinkle hedgehog ($twhh$) which shows histological features comparable to the human disease [13]. We also describe the first tissue specific glioma model in zebrafish based on $HRAS^{G12V}$ expression under control of the $zic4$ enhancer [11]. Expressing the oncogene exclusively in the dorsal central nervous system of double transgenic larvae induces an increased number of dedifferentiated neural cells and increased brain size due to hyper proliferation, which leads to early lethality. Juvenile tumor carriers, generated through somatic oncogene expression, reveal massive telencephalic tumors resembling human malignant glioma. In another glioma study, Ju et al describe the induction of malignant brain tumors in zebrafish by expression of mCherry fused $KRAS^{G12V}$ under two different neural promoters, $krt5$ and $gfap$ [44]. While both models develop brain tumors in about 50% of fish at 1 year of age, tumor type and histology are different with the $krt5$ promoter inducing mostly MPNST (malignant peripheral nerve sheath tumor)—tumors and $gfap$ promoter mainly tumors of the brain parenchyma. These studies using the same oncogenic line reveal the diverse impacts the same oncogene can have on different cell types (Fig. 1b–d). The availability of many diverse Gal4 lines and tested oncogenic constructs therefore allows to screen the existing libraries for new driver lines suitable for cancer modeling.

Additionally, the same GAL4 line can be used to study the effects of different oncogenes on cancer type/progression. In a pancreas cancer model the oncogenic Ras variant $KRAS^{G12V}$ was placed under UAS control and induced by Gal4 under control of the $ptf1a$ promoter. While this promoter activates gene expression in the cerebellum, hindbrain and pancreas, double transgenic $KRAS^{G12V}$ expressing fish developed only pancreatic cancer [10]. Overexpression of another oncogene (dominant active Akt) under the $ptf1a$ promoter instead leads to glioma development in nearly $1/3$ of double transgenic 6 months old fish [12]. The $ptf1a:Gal4$ line was further used to evaluate different $KRAS$ mutants for their oncogenic potential through injection of the constructs into one-cell stage $ptf1a:GAL4$ carriers [45]. From the 12 mutant $KRAS$ forms used, all the 7 variants found in human pancreatic cancer also induced pancreatic cancer in zebrafish, while only 1 of the 5 mutant $KRAS$ forms found in human non-pancreatic cancers was able to induce pancreatic cancer in zebrafish [45]. While confirming the similarity between zebrafish and human oncogenic processes, this study also revealed that the Gal4/UAS system is suitable to screen the oncogenic potential of different mutant proteins in a tissue specific manner. The possibility to apply different transgenes to the Gal4 line directly through injection into one-cell stage Gal4 carriers obviates the generation of large numbers of stable transgenic lines, although the expression of the transgenes will be mosaic and therefore differences are difficult to interpret.
The Gal4/UAS system can also be utilized for gene interaction studies. In a crossing intensive approach the pancreas cancer model has been tested in mutant background lines for the ribosomal proteins rpl36\(^{-/-}\), rpl23a\(^{-/-}\) or rpl36\(^{-/-}\)/rpl23a\(^{-/-}\) which showed that rpl36 but not rpl23a functions as haploinsufficient tumor suppressor [46]. By far easier is the co-expression of different transgenes. In an activator line driving Gal4 expression under the cytokeratin promoter krt4, Ju et al. investigated the oncogenic effect of UAS controlled mutant sonic hedgehog (SHH) and mutant AKT1 (myrhAkt1) [8]. While neither of them alone induce tumors, injection of double transgenic carriers krt4:Gal4, UAS:SHH with the construct UAS:myrhAKT1 at one-cell stage led to development of diverse tumor types in the trunk, the eye and the head region that were identified as spindle cell sarcoma, rhabdomyoma, ocular melanoma, myoma, astrocytoma and glioblastoma. This latter study illustrates the strength of the Gal4/UAS system for easy and fast combined directed gene expression studies without the need of long and costly transgenic approaches, besides the promiscuity of the krt4 promoter.

Site directed gene silencing instead still remains an obstacle in zebrafish. Help may come from a technique tested by Dong et al. [47]. Through combination of inhibitory short hairpin RNA (shRNA) with the backbone of the micro RNA miR-30 they generated a stable transgenic line (miR-shRNA), which is able to inhibit translation through mRNA degradation. Similar approaches have been described using artificial microRNAs [48], but have not yet been exploited in zebrafish. Controlled by a UAS promoter, these effects on mRNAs can be directed to specific cells or tissues [47]. While these methods have not been tested in zebrafish cancer studies yet, they have the potential to revolutionize model development as it not only accomplishes gene silencing, but also allows for spatial and temporal control of gene silencing. A schematic representation of the GAL4/UAS system (and its derivatives) applied to zebrafish cancer models is depicted in Fig. 1.

**Gal4/UAS Alternative Systems**

Two alternative systems to Gal4/UAS have been tested in zebrafish that promise to undergo less silencing and will increase transgenic potential as they are not interacting with the Gal4/UAS system (Fig. 1h, j). The TrpR/tUAS system [49] and the Q-system [50] are derived from *Escherichia coli* and the fungus *Neurospora crassa*, respectively. Due to a lower CG content in their UAS sequences both systems are less vulnerable to methylation induced silencing than the Gal4/UAS system. In a pioneering zebrafish study the wildtype tryptophan repressor (TrpR) and two mutants with 5x (T81M) and 11x (T81A) reduced transcriptional activity have been tested [49]. While no silencing occurred over four generations, TrpR appeared toxic to zebrafish larvae when expressed under strong promoters, thus requiring further optimization steps. In addition to the transcription factor (QF) and its binding site (QUAS) the Q-system includes a repressor (QS) which itself can be inhibited in some organisms by quinic acid [51, 52]. However, while QF and QUAS are functional in zebrafish, efficient inhibition of the system could only be achieved through
injection of QS mRNA in embryos [50]. Stable effective QS expression instead could not be achieved, probably due to titration effects. Thus, the functionality of quinic acid has not been tested in zebrafish. As both systems have been generated through the Gateway cloning system they can easily be adapted to different needs and thus provide a novel platform for transgenic zebrafish models.

Site Specific Recombinases

Site specific recombinases (SSR) allow for site specific genome manipulation. Currently, three main systems are used for site specific recombination: the Cre/loxP, Flp/frt and phiC31 system, plus the newly discovered Dre/rox-system. While the Cre, Flp and Dre are similar and recombine between a pair of identical sites, the phiC31 system recombines between two different sites. However, all these systems allow for three different reactions: excision, insertion and inversion.

The Cre/lox System

The Cre/lox system relies on the 343 amino acid long Cre recombinase of the bacteriophage P1 that catalyzes site specific recombination between two 34 bp long loxP (locus of cross-over in P1) or simply lox sites. The lox sites consist of an asymmetric 8 bp spacer that determines site orientation and is flanked by two 13 bp inverted repeats, to which Cre binds [53]. By manipulating the orientation of two lox sites, Cre functionality can be controlled in the way that head-to-tail orientation induces almost irreversible insert excision, while head-to-head orientation leads to insert inversion. As after inversion the two lox sites still reside in cis, this reaction is reversible and thus inefficient (Fig. 2a, b, d).

Different variants of the wildtype loxP sites exist, which increase the efficiency and flexibility of the system and allow for more sophisticated transgenic approaches such as orientation coordinated “recombinase-mediated cassette exchange” (RMCE) or stable insert inversion through the “flip-excision” (FLEX) approach or the inverted-repeat method that depends on two partly mutated lox sites (reviewed in [54, 55]). This set of techniques made the Cre/lox system the first widely used transgenic approach in fly [56], mice [57] and mammalian cell cultures [58].

In zebrafish the Cre/lox system was first tested for its excision ability. In 2005 Pan et al. applied Cre mRNA to one-cell stage eggs of a transgenic zebrafish line expressing lox flanked GFP, which leads to strong mosaic GFP reduction through excision of the GFP sequence from the zebrafish genome [59]. Few years later Liu et al. proved that the Cre/lox system can also be applied for gene inversion in zebrafish, although with very low efficiency [60]. Using the LE (left excision)/RE (right excision) mutant sites lox71 and lox66, respectively, they showed that a lox66 flanked RFP (red fluorescent protein) coding sequence can be integrated into a stable single-site lox71 line through co-injection of Cre mRNA and the vector pZk-lox66RFP. Shortly after, it was shown that—just as in mouse—Cre can be stably...
Fig. 2 Site specific recombinases (SSRs) enable cell type specific stable transgene integration. (A) The group of available SSR systems comprises the proteins Cre (blue shape), Flp (pink shape) and Dre (green shape) that target pairs of the respective sites lox, frt and dre (small shapes with corresponding colors), and the protein phiC31 (violet shape) that recombines between attP and attB sites transforming them into nonreactive attL and attR sites. (B–E) To recombine, all SSRs require a pair of target loci in close proximity and catalyze different reactions depending on loci orientation. (B, C) Excision and insertion of fragments between target sites in head-to-tail orientation can be catalyzed reversibly by Cre/Flp/Dre (B) or irreversibly by phiC31 (C). (D, E) Inversion of fragments between target sites in head-to-head orientation can be catalyzed reversibly by Cre/Flp/Dre (D) or irreversibly by phiC31 (E). (F, H) Smart combination of different mutants of loci of the same SSR system allow for enhanced reaction control. (F) A promising application example for zebrafish cancer studies is the *ubi:zebrabow* line that encodes the three fluorescent proteins RFP (red), CFP (blue) and YFP (yellow) flanked by pairs of the mutant lox2272 sites (white triangles) and the wildtype loxP sites (blue triangles) in head-to-tail orientation. Cre can either be inactive, labeling cells in red, or can be active and induce recombination between pairs of homologous lox sites, labeling cells and their respective clones either in blue (RFP excision, blue rectangles with blue halo) or yellow (RFP and CFP excision, yellow rectangles with yellow halo). (G) Tail region of a 2 dpf *ubi:zebrabow* larva injected with the ubiquitous Cre expressing plasmid CMV:Cre. *Insets* show single channel expression (RFP, CFP and YFP) of the boxed melanocyte. (H) An example for the application of the zebrabow tool to study tumor origin by injecting UAS controlled Cre into crosses between the *ubi:zebrabow* and a *ptfa:Gal4, UAS:loxSTOPloxKRAS<sup>G12V</sup>* line (Gal4 expression in the pancreas and UAS controlled oncogenic KRAS which must be activated by Cre dependent excision of the STOP codon) that would allow to distinguish different tumor clones in a somatic pancreas cancer model. Calibration bar: 25 μm. *Abbreviations*: CFP cyan fluorescent protein, dpf days post fertilization, RFP red fluorescent protein, SSR site specific recombinase, UAS upstream activating sequence, YFP yellow fluorescent protein, *ubi* ubiquitin promoter
expressed under specific promoters in zebrafish without significant toxic effects [61, 62] which allows for promoter dependent tissue and time specific onset of Cre expression in this organism.

Despite these promising results until today only few groups developed cancer models using the Cre/lox system. The first group to apply it for cancer modeling was Langenau et al. in order to develop their early lethal, c-Myc dependent T-ALL (T-cell acute lymphatic leukemia) model into a stable transgenic line [15]. Briefly, this model depends on the construct rag2-EGFP-mMyc that drives expression of mouse c-Myc under the rag2 promoter and leads to GFP labeled, m-Myc expressing T-lymphocytes. These fish develop clonal lymphoblastic leukemia that expresses the human T-ALL oncogenes tal1/scl and Imo2 and can be transplanted into irradiated recipients. Due to its aggressiveness the fish die before reaching sexual maturity, and studies on the model require injections of the transgene, which leads to variability in integration. To overcome this hurdle, Langenau et al. generated a stable line expressing a modified construct containing the loxP flanked red fluorescent protein dsRED between the rag2 promoter and GFP. Without Cre this line expresses rag2-controlled dsRed, while Cre mRNA injection mediates dsRed excision and allows GFP and m-Myc expression as well as development of leukemia even though with much reduced frequency compared to the Cre/lox independent version of this model. In another study Seok et al. used the Cre/lox system to investigate the ability of oncogenic KRAS<sup>G12V</sup> under control of the nestin promoter to cause brain tumors, but instead found that activation of KRAS<sup>G12V</sup> expression through injection of Cre mRNA only induces neural cell differentiation, apoptosis or brain edema which leads to early lethality [63].

Besides conditional tumor induction, the Cre/lox system also proved its value for detailed lineage tracing [64, 65], which is especially interesting for studies on tumor clones. The so called zebrabow line is derived from the mouse Brainbow model [66, 67] and is based on a construct that contains a promoter, the three red, blue and yellow fluorescent proteins RFP, CFP and YFP, respectively and pairs of lox2272 and loxP sites spanning the first or the first two fluorescent proteins, respectively [68]. In this way Cre/lox activity leads to different excision events and allows only one fluorescent protein to be expressed per recombination event, hence, enabling differentiation between cells belonging to different clones. Further variability is added through copy numbers as each copy can add a different color to the cells increasing color variability in the tissue. Therefore, hetero- or homozygosity of the fish severely influences cell labeling (Fig. 2f, g). Pan et al. generated two zebrabow lines. The first line known as ubi:zebrabow uses the ubi promoter, established by Mosimann et al. [69], to label cells during all stages of development. The second line known as UAS:zebrabow uses the Gal4 specific promoter sequence UAS, hence allowing tissue specific cell labeling. This latter line may be a valuable tool in zebrafish cancer studies to examine tumor heterogeneity, clone development and progression and decision fate in metastatic models [70]. Examples of the flexibility of the CRE/lox system are shown in Fig. 2.
The Flp/frt System

The Flp/frt system derived from *Saccharomyces cerevisiae* is highly comparable to the Cre/lox system in terms of recognition sites and functionality (Fig. 2a, b, d). The recombinase Flp (flipase) recognizes 34 bp sequences, called frt (flp recognition target). Even though they have a similar design to the lox sites, these frt sites are inert to Cre, which allows simultaneous application of both systems in a line, thus enhancing transgenic flexibility. As the best activity of Flp is achieved at around 30 °C, an enhanced version (FLPe) was developed which retains its activity at higher temperatures, useful for applications in mammals [71]. The Flp/frt system has been intensively used over a long time in drosophila, where it is an established tool (reviewed in [72, 73]) and in the mouse both alone [74, 75] and in combination with the Cre/lox system [76, 77]. Only in 2011 it was tested in zebrafish through an excision experiment to proof its functionality in this model organism [78]. Since then Cre/Flp combination tools have been developed allowing highly sophisticated gene expression control. The FlipTrap was developed by Trinh et al. as a novel tool for fluorophore labeling of endogenous proteins for analysis without the bias of overexpression approaches [79]. For their study, they generated a library of about 170 FlipTrap lines (see Box 1 for web site). The plasmid to label endogenous proteins contains the yellow fluorophore *Citrine* flanked by splice acceptor and donor sequence instead of initiation and stop codon and an adjacent inverted *mCherry* sequence. Both sequences are embedded in two heterologous pairs of lox sites that allow simultaneous Cre induced excision of citrine and inversion and hence activation of *mCherry*, which results in truncated *mCherry* labeled endogenous proteins. Frt sites flanking the construct allow for exchange of the construct at the very same locus. The FT1 vector developed by the Chen lab allows conditional sequential activation and inhibition of gene expression [80, 81]. The Tol2 based vector combines both Flp and Cre sensitivity by strategically flanking a gene and a five repeat stop sequence with pairs of heterotypical lox and frt sites so that Cre and Flp activity in either order can invert the gene and so activate or inactivate its expression. Dependent on the expression pattern of Cre and Flp this allows for (1) tissue specific or general knock-out experiments followed by tissue specific rescue or, if the construct is initially integrated in active orientation, it allows for (2) tissue specific or general rescue followed by tissue specific knock-out. Even though none of these systems has been used so far in cancer studies, they illustrate the gain in flexibility obtained through combination of SSR systems and may serve as inspiration for the development of more sophisticated zebrafish cancer models.

The Dre/rox System

A third recombination system is the Dre/rox system, which has recently been discovered in the phage D6, related to bacteriophage P1 (Fig. 2a, b, d). It recognizes sequences called rox, which resemble lox sites in size, sequence and structure but are not compatible with Cre. Equally, also Dre does not function on lox sites [82].
Until now this system has been tested in embryonic stem cells and mouse [83, 84], and also in zebrafish, where it proved its flexibility and compatibility with the Cre system [85]. Thus, it provides an additional tool for SSR dependent conditional transgenesis allowing for more sophisticated lineage labeling and gene expression control, for example through intersection of partially overlapping promoter activities.

The phiC31 System

The phiC31 integrase system offers highly specific spatial control over any transgene integration. It is derived from the PhiC31 bacteriophage, which uses it to integrate its genome into the Streptomyces chromosome. The recombination is highly specifically performed between the attP site of the phage and the attB site of the Streptomyces and requires no cofactors [86]. Successful recombination results in hybrid attL and attR sites. As these are incompatible with phiC31 the reaction is irreversible and provides a unique characteristic over the reversible reactions catalyzed by Cre and Flipase (Fig. 2a, c, e). This makes the phiC31 integrase preferable for transgene insertion strategies. Only recently, a recombination directionality factor has been found which allows reversion of the reaction when combined to the phiC31 integrase [87]. Of note, despite the nomenclature is the same, the phiC31 system sites are different from the attP and attB sites of the lambda phage-derived Gateway system and thus, are not interacting with those.

Since its discovery, the phiC31 system has been applied to various eukaryotic systems including human cells [88], mouse [89], drosophila [90] and frog embryos [91]. In these, it was established as an advantageous system for site directed insertional transgenesis, while in mice and humans pseudo-attP sites in the genome complicate the application of phiC31 in these models [92]. Indeed, gene function analysis in transgenics suffers from unpredictable integration sites of diverse transgenes making quantitative and qualitative analyses difficult and labor intensive. With the phiC31 system, comparable transgenic lines can easily be generated. Well-characterized transgenic lines with a single attP site integration can be used to integrate different transgenes into a predefined region to compare them under similar genomic conditions. This facilitates for example comparison between native and mutant proteins. Further, as it naturally recombines sequences of 100 kb, the phiC31 integrase sets basically no limit to insert size allowing a wide range of applications.

Despite its success in other organisms, its functionality in zebrafish has first been shown in 2010 [93, 94]. As in other model organisms, the phiC31 integrase also proved its advantage for comparative transgene analyses. After Lu et al. [93], also Lister showed that phiC31 is generally active in zebrafish and able to excise inserts flanked by attP and attB sites [93, 94]. Hu et al. proposed a method for RMCE in zebrafish through phiC31 [95]. Using a tol2 construct they expressed attP flanked GFP under a tissue specific promoter and could exchange it with an attB flanked mCherry. A drawback however is the relatively low efficiency of phiC31 integrase in genomic approaches. Another approach described by Mosimann et al. aims to establish single-insertion zebrafish lines that harbor only one attP site [96].
They developed three functional lines with single \textit{attP} integration sites of which they characterized two with homogenous gene expression. Using the \textit{phiC31} mRNA inserts as large as 8 kb flanked by \textit{attB} sites could be inherited. They showed that for at least four generations the signal is stable appearing with the same intensity among all larvae of a clutch. Similarly, Roberts et al. confirmed the site specificity of \textit{phiC31} mediated transgene integration as well as a low rate of unspecific integration. This leads to the conclusion that the fidelity of the recombination is high [97].

Further development of this system has been applied in a medaka study by Kirchmaier et al. who constructed an elegant example for the combination of the \textit{phiC31} system with the Cre/\textit{lox} system to integrate different genes and promoters at specific sites and described a fluorescence based control mechanism for efficient recombination [98]. This system allows for directed gene integration into a pre-defined locus through the \textit{phiC31} integrase, while fluorescence confirms successful integration. In a second step, the \textit{lox} sites allow for locus cleanup removing fluorescent proteins and promoters and leaving only the insert behind [98].

To summarize, during the last 5 years, the \textit{phiC31} system has been developed as a powerful tool for highly controlled spatial transgenesis, which may improve comparison between gene effects independently of integration site but also allows the study of enhancer and promoter characteristics. Its ability to integrate constructs of yet unlimited size adds additional value to this system. However, in order to gain full use of the \textit{phiC31} integrase system, more \textit{attP} site single-insertion zebrafish need to be developed and characterized. Here, the novel discovery of the CRISPR/Cas9 system could be used to develop designer lines. Also concentration dependent toxicity [95] could be reduced and activity improved through modifications of the \textit{phiC31} integrase as it has been already done for the mouse.

\textbf{Conditional Cancer Gene Manipulation Through Genome Editing}

Genome editing includes a range of methods for targeted genome manipulation. Combination of an engineered nuclease containing a sequence specific DBD with a generic DNA cleavage module enables highly locus specific knock-out and knock-in of individual genes. This provides a great advantage over timely limited knock-down methods such as morpholinos or mRNA injections and unspecific mutagenesis through chemicals or radiation. The three main systems applied are known under the acronyms ZFN, TALEN and CRISPR/Cas9.

Zink-finger nucleases (ZFNs) were the first genome editing tool applied to zebrafish in 2008 [99]. Only 3 years later the discovery of the transcription activator-like effector nucleases (TALENs) provided a new and more efficient tool for targeted zebrafish mutagenesis [100, 101]. Both ZFNs and TALENs rely on protein-DNA interaction with relatively restrictive combinations between grouped amino acids and nucleotides. For detailed information we refer the readers to the many excellent reviews published on these methods [102–105]. Both ZFNs and TALENs effectively allow genome manipulation through mutation, but also deletions and insertions and are even applied for therapeutic approaches to combat HIV/
AIDS currently in clinical trials [106]. However, reports of deletions or insertions of cancer-related genes using ZFNs or TALENs in zebrafish are lacking, thus testifying the complexity in tool design, and restricting their use to experienced labs.

The clustered, regularly interspaced, short palindromic repeats (CRISPR)/CRISPR associated (Cas) system (CRISPR/Cas) provides a great alternative to ZFNs and TALENs as it relies on Watson-Crick base pairing instead of the potentially less specific protein-DNA interaction [106]. The CRISPR/Cas system can be found in many bacteria and archaea as an adaptive immune system against viral infection where it integrates viral DNA into the CRISPR loci as so called protospacer. The protospacers are recognized through their PAM sites (protospacer adjacent motif), which flank them at their 5′ or 3′ end as GGN or NGG, where N stands for any nucleotide. At the second infection, protospacers are transcribed into CRISPR RNA (crRNA), which anneals to transactivating crRNA (tracrRNA) that in combination with the endonuclease Cas specifically digests foreign DNA sequences. The CRISPR/Cas system II is the system of choice as it is the only of the three known CRISPR/Cas systems that requires only a single Cas endonuclease (Cas9) [107].

Since its first application in mammalian cells in 2013 [108, 109], it has already been applied to a variety of species including plants [110], mice [111, 112], rats [113], rabbits [114], goats [115], pigs [116] and zebrafish [117, 118], to name some. The main advantage of the CRISPR/Cas9 system lies in its simple design. It relies on two components (1) a ~20 bp long guide RNA (gRNA) sequence, encoding the sequence of the target and (2) the Cas9 which can be applied either as protein or as mRNA. The Cas9 endonuclease contains two enzymatic activities that cleave the complementary and non-complementary strands, allowing Cas9 to induce double strand breaks in any target sequence [107] (Fig. 3a). Thus, for transgenesis only the gRNA needs to be designed. In zebrafish the CRISPR/Cas9 system is applied by injection of a gRNA and the Cas9 mRNA from Streptococcus pyogenes into one-cell stage embryos and reveals high efficiency for knock-out of single genes (70–90% [117–119]), as well as for simultaneous knock-out of multiple genes which allows analysis of both single and combinatorial knock-out effects already in the F0 generation. Through injection of two different gRNAs and Cas9 mRNA even huge chromosomal deletions are possible [120]. To reduce mutation events the knock-out efficiency can also be titrated through varying RNA concentrations [118]. As the mutations are inherited to the next generation the CRISPR/Cas9 system allows to generate stable knock-out lines with a fast and easy protocol [121]. While off-target effects have been found to appear in only few cases, still deeper investigation is required [119]. However, the system displays variable efficiencies for different genomic targets with some genes yielding effects in <1% of injected embryos [117, 118, 122]. To optimize the CRISPR/Cas9 system, a zebrafish codon-optimized version of Cas9 (zCas9) was constructed and flanked with nuclear localization signals that showed a superior performance compared to wildtype and human optimized Cas9 [118, 123]. Application of different Cas9 variants from other prokaryotes or generated through mutation may lead to further improvements in efficiency and specificity. The gRNA sequence is especially important for specific and directed Cas9 activity and can be optimized by various means [122]. For design and specificity
Fig. 3 CRISPR/Cas9 facilitates genome manipulation on demand. (A) Schematic of CRISPR/Cas9 technology applied to zebrafish. For Cas9 dependent genome editing the gRNA and Cas9 mRNA are injected into one-cell stage embryos, where the Cas9 mRNA is translated into Cas9 protein and together with the gRNA transported into the cell nucleus. In the nucleus the gRNA binds the corresponding DNA sequence which allows Cas9 to digest the DNA at the adjacent PAM sequences and induce a double strand break through its two enzymatic activities (scissors). The cell can repair this damage through different mechanisms that allow various applications (B–D).

(B) Through Non-Homologous End Joining (NHEJ) the double strand break is repaired by random integration of nucleotides (boxes with A, T, G or C) which often induces a frame shift and generates abnormal and mostly non-functional proteins. (C) Coinjection of gRNA and Cas9 mRNA with small single stranded DNA (ssDNA) allows strand repair through homology-dependent recombination which leads to the integration of a desired small sequence such as a stop codon (green box) to generate a truncated protein, tags of human influenza hemagglutinin (HA) (blue box) to allow easy protein detection, or target sites for SSRs such as lox, frt, rox or attP (yellow boxes) to allow to use this line for site specific integration of constructs into a defined locus. (D) Co-injection of gRNA and Cas9 mRNA with a plasmid encoding a bait sequence (yellow oval) and a gene (green rectangle) allows strand repair through homology-independent recombination and insertion. gRNA dependent linearization of the plasmid at the bait sequence allows its integration into the genomic DNA at the site of gRNA induced double strand break resulting in expression of a transgene under an endogenous promoter. Abbreviations: A adenine, C cytosine, G guanine, gRNA guide RNA, HA human influenza hemagglutinin, NHEJ non-homologous end joining, PAM protospacer adjacent motif, ssDNA single strand DNA, SSR site specific recombinase, T thymidine.
testing several freely available software tools have been developed (see Box 1 for web sites) to enable even inexperienced labs to implement the CRISPR/Cas9 system in their daily routine.

Besides successful knock-out, also knock-in strategies have been developed with the CRISPR/Cas9 system (Fig. 3b–d). Homology directed recombination through injection of single strand DNA together with gRNA and Cas9 mRNA allows insertion of short fragments of ~50 bp such as loxP or HA tags into the zebrafish genome but frequently leads to inverted or partial insertion [119, 123]. Homology-independent knock-in methods instead, make use of an insert flanked by one or two gRNA target sites (baits) as double strand DNA or as plasmid [124–126]. As plasmid DNA is less toxic than linearized DNA, it can be used at higher concentrations for more efficient integration [125]. In several studies, integration efficiencies of 20–30% have been reached in the F0 generation with inheritance to the next generation by 10% of carriers [124–126]. A step-by-step protocol for homologous-independent knock-in is available from Auer and colleagues [127].

While none of the genome editing approaches has been used to generate zebrafish cancer models, yet, they have so far found their application in gene function analyses [128, 129] and disease modeling [130–132]. Hence, we propose that genome editing systems provide a valuable tool for detailed analysis of tumor suppressor genes. With the CRISPR/Cas9 system, the zebrafish research community now gained a flexible, cheap and easy tool to generate mosaic knock-out or lines on demand which shall improve the analysis of otherwise difficult to investigate tumor suppressor genes.

**Temporal Control**

Timing of oncogene expression or inactivation of tumor suppressors may be critical in cancer modeling, as it may trap specific cell types which are more capable of cancer initiation. In addition, duration of oncogenic stimuli or withdrawal from oncogene expression may reveal the strength of oncogene addiction at different stages of tumor progression. Therefore, systems for temporal control of expression of oncogenes or tumor suppressors have been developed in zebrafish. Some of these systems will also allow spatial control, or may be used in combination with spatial control methods. The most popular systems for temporal control of gene expression are listed below.

**Heat Shock Promoters**

Control of gene expression through temperature regulation has been achieved in zebrafish thanks to promoters of heat shock proteins [133]. Understandable, this approach is difficult in mammals due to their tight regulation of body temperature [134]. In poikilotherm organisms instead, such as the zebrafish, temperature controllable systems have become widely implemented. The zebrafish heat shock promoter 70l is a 1.5 kb sequence that contains tandem arrays of 5 bp DNA
consensus-sequences named heat shock elements (HSE). The transcription factor to activate the HSE is called heat shock factor (HSF). At physiological temperatures HSF is bound in a protein complex containing the protein hsp90. Increase in temperature forces the proteins in the complex to change their conformation and release HSF to bind and activate the HSE [135, 136]. While its application in zebrafish is easy and straightforward, it also holds a significant drawback. Besides its activation through a lot of stressors such as toxins or starvation, it is also active during early development [137]. However, this did not impede its wide use in zebrafish models. The heat shock promoter has been used instead of ubiquitous promoters to drive general gene expression as well as in combination with laser irradiation for tight local expression control to study gene function, search for transcription enhancers and for lineage labeling (reviewed in [138]).

Also cancer models have been generated in zebrafish based on the heat shock promoter. Due to its easy timel y controllable induction, the heat shock promoter allows for development of stable lines that encode otherwise lethal oncogenes. The induction of HRASG12V expression during larval stages induces hyperproliferation and cellular senescence [139], similar to the expression of oncogenic HRAS in human fibroblasts [140]. Additionally, the heat shock promoter has been used to construct inducible leukemia models. Yeh and colleagues expressed the human oncogenic fusion protein AML1-ETO under control of the heat shock promoter 70l [141]. Activation of this oncogene in larvae resulted in a lack of circulating blood cells despite an intact cardiovascular system, accumulation of immature hematopoietic blast cells, and gene expression comparable to the myeloid leukemia in human with upregulation of lmo1, hoxa9 and hoxa10 and downregulation of cmyb, gata and scl. Remarkably, injection of scl mRNA and treatment with the drug Trichostatin A could partially rescue the AML1-ETO induced phenotype which makes the line suitable for drug screens. Shen et al. developed a line that expresses the oncogene Mycn and the fluorescence protein GFP under control of a sequence of bicistronic HSE [142]. Already one single heat shock at larval stage induces the development of leukemia like lesions that are characterized by leukemia typical changes in blood cell populations to more immature types and upregulation of scl and lmo2, genes that have also been found upregulated in human leukemia. Additionally, heat shocked fish have a significantly reduced life span (9.5 months compared to 2.5 years in adults) and even at 8 months show a similar histopathology as the 2 month old juveniles.

A model for Ewing sarcoma, a malignant tumor characterized as small-round-blue cell tumor, depends on heat shock promoter controlled expression of the human EWS-FLI1, a fusion oncoprotein of the 5′ transactivation domain of EWS and the 3′ Ets domain of FLI-1 [143]. Induction of this construct in a p53−/− background through a heat shock at 50% epiboly induces development of solid tumors, leukemia like tumors and malignant peripheral nerve sheath tumors with the typical histology of small, round, blue cells similar to human Ewing’s sarcoma and mouse models, which could also be transplanted.

Besides heat shock promoters also heat sensitive splicing has been described. An example is the mitf<sup>vc7</sup> mutant which allows inactivation of mitf at temperatures above 28 °C. This variant has been used in a melanoma model to elucidate the function of mitf in melanoma development [144].
Hormones

Hormone dependent activation of transgenes depends on so called nuclear receptors. A large group of steroid and thyroid receptors, which are located in the cytoplasm, upon ligand binding translocate to the nucleus where they bind specific genomic regions through their DBD and function as transcription factors. The homodimeric functional steroid receptors estrogen receptor (ER), progesterone receptor (PR) and glucocorticoid receptor (GR) are the most frequently used ones for conditional transgenic approaches. To utilize their hormone specificity, the ligand-binding domain is fused to a transcription factor of interest thus rendering the activity of the transcription factor hormone dependent [145]. To activate the ligand-binding domains of ER, PR and GR the most frequently applied hormone analogs are tamoxifen, mifepristone (also known as RU486) and dexamethasone, respectively. While mifepristone and dexamethasone are directly effective, tamoxifen is a prodrug, which is metabolized to 4-hydroxy-tamoxifen (4-OHT) through the liver enzyme CYP 450 2D6.

In the zebrafish, 4-OHT can be applied directly to the water to avoid irregularities due to enzymatic interactions and reduce variations in dose response curves. In different studies it was shown to reliably control gene expression in a dose dependent manner and without measurable leakiness [146, 147]. In a further evolution of the rag2 controlled leukemia model developed in the lab of T. Look [15] the cMyc oncogene is fused to the ligand binding domain of ER [147]. This allows for conditional activation of the oncogenic effect of cMyc through application of 4-OHT which leads to development of fully established T-ALL already after 5 weeks of treatment, while withdrawal of 4-OHT blocks cMyc activity and leads to apoptosis of cancer cells. The untreated siblings instead do not develop blood abnormalities. Also in an in vitro study of a zebrafish derived liver cell line transfected with ER bound Raf1 only cells treated with 4-OHT showed hyperactivation of the MEK-ERK cascade and increased proliferation in cell culture as well as after transplantation while untreated cells showed no alterations [146]. The other two systems, namely PR/mifepristone and GR/dexamethasone, are also functional in zebrafish [85, 148]. As all three systems are not cross reacting [85], they allow for simultaneous or sequential activation of different target genes and thus add enormous flexibility in transgenic zebrafish cancer models.

Tet-On/Tet-Off

The Tet-system is derived from Escherichia coli, where it controls expression of genes mediating tetracycline resistance. Its activity is controlled through the tet-Repressor (tetR), which is reactive to tetracycline or its more stable derivative doxycycline (Dox). Through fusion of tetR to the activation domain of VP16 the Dox controlled transactivator (tTA) is generated. In absence of Dox tTA binds specifically to tetracycline operator (tetO) sequences and activates transcription from minimal promoters adjacent to tetO. As tTA is inhibited through Dox binding, this system is called Tet-Off [149]. The Tet-On system instead relies on a mutant
transactivator (rtTA) with four substituted amino acids and is physiologically inactive. It can bind and activate the tetO element only upon Dox binding [150]. This system is mostly preferred for gene silencing and controlled activation of gene expression [151] as it circumvents problems that may arise with the Tet-Off system such as slow withdrawal of Dox from the organism and thus delays in induction of gene expression. The Tet-On system has been applied to many different organisms including plants [152], drosophila [153], mice [154, 155] and zebrafish. However, in the first studies the system appeared leaky, which hampered its further use [156]. Through viral mutagenesis, variants have been designed with increased stability and reduced leakiness [157]. One of them is the Tet-On Advanced (rtTA2s-M2), which was achieved through random mutation of the tetracycline repressor gene and screening for alleles with reverse binding properties. Additionally, the VP16 activation domain is replaced with a minimal activation domain (12 amino acids long) and a coding sequence optimized for expression in human cells [158]. These changes make rtTA2s-M2 10 times more sensitive to Dox than rtTA, while it has severely reduced background activity and is more stable in eukaryotic cells. However, as in zebrafish some background activity is measurable, a fusion system has been developed. Binding of rtTA to the ligand binding domain of GR or the ecdysone receptor (EcR’) from the silkworm (Bombix mori) allows for additional control through dexamethasone or tebufenozide, respectively [148]. Both hormone systems reduce leakiness to undetectable levels. While dexamethasone at higher concentrations may deplete T-cells [159], side effects of tebufenozide are not known. Additionally, the EcR’ bound rtTA system exhibits higher sensitivity which makes it preferable.

In zebrafish cancer studies, the basic rtTA system has been sufficient. The initially Gal4/UAS based glioma model from Ju and colleagues [44] was also adapted for conditional oncogene expression through transversion to the Tet-On system. Expression of rtTA under either of the promoters krt5 and gfap and control of the mCherry fused KRASG12V oncogene through a tetO allowed for analysis of early effects and thus for chemical treatment approaches [44]. Additionally, different Tet-On dependent liver cancer models have been generated with oncogenes expressed under the control of the liver specific promoter fabp10. The most aggressive model depends on conditional expression of Xmrk, the Xiphophorus derived mutant form of the epithelial growth factor receptor (EGFR) [160] and leads to hepatocellular carcinomas (HCC) already after 3 weeks of treatment accompanied by diminished growth and increased lethality, while withdrawal of Dox leads to a complete rescue after only 4 weeks. A less severe model relies on overexpression of Myc [161], a gene, often amplified and associated with bad prognosis in human HCC. While it can lead to induction of liver tumors with gene expression profiles comparable with human HCC and mouse models, it shows less malignancy than models depending on expression of Xmrk or KRASG12V. However, this can be an advantage, as Myc expression can generate a sensitized background that allows to discovery and evaluate new, yet undetermined oncogenes in HCC.
In a different approach the Tet-On system was used to investigate the complicated relationship between oncogenic \( kras \) and \( rhoA \) [162]. Through Dox dependent co-expression of both \( kras^{G12V} \) and wt, dominant active (DA) or dominant negative (DN) \( rhoA \), Chew et al. [162] found that \( rhoA \) has an inhibitory effect on the oncogenic effects of \( kras^{G12V} \). This underlines the flexibility of the Tet-On system for gene interaction studies.

**Optogenetics**

Optogenetic tools only emerged during the last 10 years. Nonetheless, optogenetics gained tremendous interest and has been established in a wide range of organisms. In a broad definition optogenetics is the science of genetic encoding of light-gated proteins for targeted control of cellular behavior, which has been shown to provide a superior tool for precise spatial and temporal control of stimulation or inhibition of cellular activities. However, optogenetics is often restricted to the application of light-sensitive microbial membrane proteins which can be classified into the two groups of “optogenetic activators” (light sensitive channels and pumps) and “optogenetic sensors” (fluorescent proteins that sense \( Ca^{2+} \) concentrations or membrane potentials). Mainly applied in neuroscience and cardiac studies, optogenetic activators and sensors have intensively furthered our understanding in these fields. Especially the zebrafish, as a transparently developing organism, provides easy application of optogenetic activators and sensors [163–165]. However, as this subset of optogenetic tools provides only limited benefit for cancer research it will not be further discussed in this chapter. The interested reader is referred to recent reviews [166–170].

Another branch of optogenetics makes use of a range of light-sensitive non-channel proteins to spatiotemporally control cell signaling pathways and gene transcription (reviewed in [171]). Among the most frequently applied protein systems are (1) the LOV (light, oxygen and voltage) domain, a small domain that reacts to blue light (450 nm) and requires the cofactor flavin mononucleotide (FMN), ubiquitously expressed in mammalian cells and zebrafish, (2) the PhyB (phytochrome B)—PIF (phytochrome interacting factor) system, a binary system that heterodimerizes through red light (650 nm) and dissociates through infrared light (750 nm), (3) the CRY2 (cryptochrome 2)-CIB1 (cryptochrome-interacting basic helix-loop-helix) system that heterodimerizes under blue light exposure and (4) the UVR8 (UV-resistance locus 8), a small molecule that forms homodimers which dissociate under UV-light. These tools allow a broad range of applications. The LOV domain can be used to mask protein domains and release them upon blue light illumination, allowing for example light-induced apoptosis in a caspase7-LOV human cell model [172] or light-induced protein degradation when fused to a transgenic protein capped with degrons, small sequences that induce protein degradation [173]. In another mammalian study UVR8 fusion proteins have been used to study secretion...
processes. As UVR8 forms dimers, it leads to protein aggregation in the endoplasmic reticulum, which prevents further secretion. Only upon UV-light exposure UVR8 dissociates allowing further trafficking of the protein [174]. Protein localization can also be affected through both, the PhyB-PIF and the CRY2-CIB1 system, which allows close and direct control over cell signaling pathways. By anchoring PhyB or CIB1 to the cell membrane and fusing PIF or Cry2 to a membrane active protein, activation or inhibition of for example the mitogen-activated protein kinase (MAPK) signaling pathway [175] or the phosphatidylinositol 3-kinase (PI3K) signaling pathway [176, 177] is possible. To control gene transcription, the AD and the DBD of the transcription activator Gal4 can be fused to either component of the PhyB-PIF [178] or CRY2-CIB1 [179] system. Light-induced fusion of the systems allows dimerization of the AD and DBD and activates Gal4 dependent gene expression. In drosophila, a model of CRY2-CIB1 controlled Cre activation allows for highly specific neural cell labeling [180].

Besides their application in a wide range of organisms, optogenetic non-channel proteins have also been used in zebrafish. Currently evaluated systems include the CYP2-CIB1 controlled Gal4 activation [179] and a single-protein transcription regulator that uses the fast cycling LOV containing protein EL222 which additionally contains a helix-turn-helix (HTH) DBD that is covered by the LOV domain in the dark to prevent DNA binding [181]. In both systems gene expression is activated through exposure of target cells to blue light. Thereby, the strength of gene expression can be regulated through intensity and length of laser exposure, which allows fine tuning of the target gene. Another interesting light sensitive protein tested in zebrafish is the KRed (killerRed), an optogenetically activatable inducer of reactive oxygen species [182]. Several lines of tissue specific KRed expression are provided on the line database ZETRAP2.0 [183] that may be helpful in elucidating the role of oxidative stress in cancer development or cure.

In summary, optogenetics provides a tool with high spatial, temporal and dose control with broad applications in zebrafish due to the transparent nature of its early developmental stages. Several studies in mammalian systems manipulating potential cancerous pathways (MAPK, PI3K) through optogenetic techniques promise to be valuable also for cancer models in the zebrafish. Additionally, bipartite systems such as the CYP2-CIB1 system, which has already been tested in zebrafish, offer the chance to target distinct subpopulations of cells through expression of both components under promoters with intersectional activity. This would for example allow to elucidate effects of different subpopulation of hematopoietic cells on leukemia development at different ages. However, as expression dependent systems such as Gal4 activation would rely on continues light exposure, more stable tools such as the Cre system should be implement for zebrafish studies to achieve stable oncogenic activity while implementing the red light dependent PhyB-PIF system would allow for more versatile combination studies such as studies on tumor clone competition or consecutive oncogene expression.
Combination Examples from Zebrafish

The diversity of available conditional transgenic systems already provides high flexibility in stage, tissue and cell type specific gene expression control. Smart combination of spatial regulation systems under control of different gene regulatory elements enables precise regulation of endogenous and exogenous gene expression in primarily spatial but also temporal manner. Temporal regulation systems instead allow for flexible gene expression independent of native gene expression timing. However, for many studies, such as evaluation of oncogene effects at different stages, high flexibility is required for both, spatial and temporal control of gene expression. In such cases, limitation to either of these control techniques would require intensive cloning and crossing workload. Thus, the combination of spatial control with temporal control systems is an important step in order to unleash the full potential of the large set of conditional transgenic approaches already established in zebrafish.
Temporal Control of SSRs

Since the first application of SSRs in zebrafish transgenesis, Cre & co have intensively been developed and combined in order to increase flexibility in gene expression manipulation. Thereby, the early combination with temporal control mechanisms boosted their improvement and increased their application field. One of the first attempts was to set the expression of Cre-GFP fusion proteins under control of a heat shock promoter [184]. This study not only confirmed the functionality of Cre in zebrafish, but additionally showed that it is not toxic when expressed systemically and can be induced through a rise in temperature. However, as the heat shock promoter is not exclusively heat sensitive but also active during early zebrafish development, leakiness was observed. This did not prevent its use in inducible cancer models.

Taking advantage of the heat shock controllable Cre activity Feng et al. further improved their zebrafish model of acute T-cell leukemia [14]. This model relies on a line encoding the construct rag2-LoxP-dsRED2-LoxP-EGFP-mMyc and was previously activated through injection of Cre mRNA. Crossing the line with a hsp:Cre line [14] allows for controlled induction of oncogene expression at different larval stages which leads to development of T-ALL in 81% of heat shocked larvae in contrast to the injection approach (13%) thereby showing that Cre expression from chromosomal integration is far more effective than from mRNA injection. Despite the positive outcomes, the leakiness of the heat shock promoter also leads to few problems. Thus, it also causes leukemia in 13% of non-heat shocked fish and the attempt to produce a line encoding both, the Cre and the lox constructs, reveals reduced aggressiveness of the developing leukemia already in the third generation. This occurs most likely due to Cre induced loss of copy numbers during the breeding through the leaky heat shock promoter and thus requires continued control of the double transgenic line. Instead keeping both constructs in separate lines helps to overcome this problem. The leukemia model of Forrester and colleagues uses the same heat shock inducible system to conditionally activate expression of the human leukemia associated gene NUP98-HOXA9 under the promoter spi1 [185]. Despite its expression in myeloid cells, central nervous system and musculature, they found exclusively myeloid proliferative like malignancies in 23% of 19–23 months old fish. In another zebrafish cancer model the oncogene KRAS<sup>G12D</sup> is expressed ubiquitously under the β-actin promoter using the construct β-actin-LoxP-EGFP-pA-LoxP-KRAS<sup>G12D</sup> [186]. Through crossing with a hsp:Cre line [14], oncogene induction can be timely controlled which leads to increased lethality and development of hyperplasia and multiple tumor types including skeletal muscle tumors, myeloproliferative disorder (MPD), intestinal epithelial hyperplasia and rarely malignant peripheral nerve sheath tumors. Also in this line, background expression of Cre can be observed which leads to tumor development in non-heat shocked fish but at later time points and with lower frequency than in their heat shocked siblings. The MPD developed in this model is not only transplantable, forming the same disease in irradiated recipients but can also be induced through resection of
hematopoietic kidney cells from non-heat shocked double transgenic adults and transplantation after heat shock into irradiated recipients. Hence, both models proved that despite the leakiness, control of Cre expression through a heat shock promoter improves survival of oncogene carriers and allows for time controlled induction of oncogene expression. Additionally, it enables easy transplantation studies with freshly induced tumor cells.

As the leakiness of the heat shock-Cre system provides severe limitations, more reliable methods have been developed. In contrast to the heat shock promoter, fusion of Cre with the estrogen receptor ER\textsuperscript{T2} leads to permanent Cre expression, but strictly tamoxifen dependent Cre activity in the nucleus [187]. Thereby, Cre activity can also be titrated in a concentration dependent way but must be tested for each CreER\textsuperscript{T2} line separately. Until now the tamoxifen inducible Cre system has been successfully applied for developmental studies, mainly for imaging approaches [188, 189]. Additionally, a large number of existing CreER\textsuperscript{T2} lines offers application of this system also for cancer studies. Since its establishment several CreER\textsuperscript{T2} lines have been developed with different global or tissue specific CreER\textsuperscript{T2} expression pattern. Hans et al. developed two different pax2a dependent lines that express CreER\textsuperscript{T2} (1) at the prospective diencephalon of the developing forebrain and (2) at the future rhombomere 3 and 5 of the hindbrain anlage, respectively. The ubi:CreER\textsuperscript{T2} line [69] instead allows general induction of Cre activity at any stage which makes it preferable over most general promoters used such as XIEef1\textalpha\textsubscript{1} (strongest expression during gastrulation and midsomitogenesis) [187], β-actin (low expression in erythrocytes, fins) [96, 190]. These and many more CreER\textsuperscript{T2} expressing lines have been compiled in the database CreZoo (see Box 1 for web site), a large library that contains over 30 different tissue and cell specific CreER\textsuperscript{T2} expressing lines together with well organized molecular and expression data [191]. This database provides a great platform for cancer screen studies when combined with a corresponding set of lox lines for Cre- mediated recombination of tumor suppressors or oncogenes.

Another hormone dependent control mechanism relies on the progesterone analog mifepristone. As it does not cross interact with the tamoxifen receptor ER\textsuperscript{T2} both systems can be used in combination enhancing complexity of transgenic approaches. In zebrafish this system has been tested by Park et al. who combined the tamoxifen inducible CreER\textsuperscript{T2} with the mifepristone inducible DreaPR to the system TAILOR (transgene activation and inactivation through lox and rox) [85] which not only proved the functionality of both systems in zebrafish but also their compatibility. Hence, it can be used for time controlled activation and inactivation of gene expression as well as for sequential transgene activation through lox and rox. Thus, it provides additional flexibility for SSR dependent conditional transgenesis which will not only improve cell lineage tracing, temporal and spatial gene expression control through intersection of partially overlapping promoter activities, but may also facilitate mimicking of consecutive events in cancer development through precise and independent activation or inhibition of oncogenes or tumor suppressor genes, respectively.
Temporal Control of the Gal4/UAS System

The Gal4/UAS system has proved its value in a variety of cancer studies [8–10, 43, 45, 192]. Although no cancer model in zebrafish applied temporally controllable Gal4/UAS-systems yet, several other studies revealed the advantageous characteristics of temporally controllable Gal4/UAS as they allowed studying gene transcription effects at later time points of development. Coupling Gal4 to the heat shock promoter allows to activate Gal4 expression in the whole organism fast and easy through an incubation at 38 °C for 30 min or spatially restricted through laser exposure [193, 194].

Several heat shock controlled Gal4 lines that have been developed by Asakawa and Kawakami can be found in the zebrafish line database zTrap [195]. However, heat shock promoters driving Gal4 expression exclude the use of tissue specific promoters, possess leakiness and permit only limited control of expression kinetics. Hormone bound Gal4 expression instead retains the possibility of tissue specific Gal4 expression and provides dose dependent kinetic control. While glucocorticoid receptor regulated Gal4 has already been published in 1998 [196], further applications were only described in 2013 using ER\textsuperscript{T2} fused Gal4 [197, 198]. In these lines, tamoxifen could induce Gal4 expression in larvae (already 3–4.5 h post treatment) and in adults. Even though slight leakiness was observed, tissue specific Gal4 expressing lines could be generated and allowed to study for example tissue specific effects of elevated Notch expression on notochord development.

The Gal4/UAS system has also been combined with optogenetic tools. Placing optogenetic activators and sensors under the UAS promoter allowed Umeda et al. to take advantage of existing Gal4 lines and to control channel protein expression in specific neural cell types [199]. Non-channel protein systems such as the CRY2-CIB1 instead have been used to control Gal4 activity through fusion of the Gal4 components AD and DBD to CRY2 and CIB1, respectively which allows for light controlled Gal4 activation through light induced fusion of the two components [179].

Besides combination with temporal control tools, the Gal4/UAS system has also been connected with the Cre/lox system such as in the promising zebrobow model developed by Pan and colleagues. In this model lox flanked fluorescent proteins are under UAS control which allows Gal4 dependent tissue specific expression in contrast to the ubiquitously expressing \textit{ubi:zebrabow} line [68].

Another combination of several tools is referred to as MAZe (Mosaic Analysis in Zebrafish) [200, 201]. This system combines a heat shock controlled Cre and a promoter controlled Gal4 which is inhibited by a lox\textsuperscript{P} flanked stop codon. As the heat shock promoter activates Cre only inefficiently, this combination results in mosaic expression of Gal4 which allows studies on sparse somatic cells or in a highly restricted spatial control of target gene activation by laser induced heat shock. Together with the previously descried Gal4 combination tools and the wide range of UAS encoding effector lines and constructs available, this system could provide new possibilities for detailed studies on somatic expression of oncogenes.
Mifeprestone/LexPR

The mifeprestone inducible LexPR system allows both temporal and spatial control of gene expression. Similar to the Gal4/UAS system, the LexPR system consists of a transactivator, which can be set under tissue specific promoter control but additionally functions only upon binding of mifeprestone, and an operator that controls expression of a subsequent gene. The transcription factor LexPR is a fusion of the DBD of the bacterial LexA protein, a truncated ligand binding domain of the human progesterone receptor and the AD of the human p65 protein. The operator promoter LexOP consists of a synthetic LexA operator fused to the minimal 35S promoter from Cauliflower Mosaic Virus [202]. Emelyanov and Parinov were the first to adopt and test this system in zebrafish [202]. For a proof of principle, they developed a line containing both, the LexPR transactivator and the gene of interest under control of the LexOP. In this line they showed that the LexPR system is not toxic to zebrafish, allows tight control of gene expression without basal expression, can reliably and concentration dependent be turned on by mifeprestone and turned off again through withdrawal of mifeprestone by easily moving fish into fresh water. Mifeprestone itself was tested to be not toxic in larvae at concentrations sufficient for target gene expression and to be applicable to adult females to gain maternal expression in embryos. Additionally, mifeprestone is a relatively cheap drug and safe to use due to the low concentrations required for system activation [202].

Using a two line system composed of a driver line that express the transactivator LexPR and an effector line encoding the LexOP controlled gene of interest, Emelyanov and Parinov also demonstrated the possibility to control gene expression in trans, to control expression of two independent genes in the same line and to efficiently express oncogenic KRAS under the cytokeratin promoter krt8 which caused strong developmental abnormalities. One drawback of the system, as observed also in other binary gene expression system, is the diversity in expression strength not only between larvae of the same clutch but also between cells of the same larva, thus caution is required with quantitative analyses. Nguyen et al. took advantage of the system to timely control oncogene expression in a zebrafish liver cancer model using an activator line expressing LexPR under control of the promoter fabp10 and an effector line expressing krasG12V under LexOP control [203]. This approach allowed not only to recapitulate the phenotype of the previous not-inducible model [204] but also to delay the oncogene expression to 1 month of age at which 100% of fish treated with 2 μM mifeprestone developed tumors after 1 month of treatment, while lower concentrations induced tumors later and less frequently. Additionally, the reversibility of the system allowed investigators to turn off krasG12V expression and demonstrated oncogene addiction of tumors. As the LexPR and the Gal4 system do not cross react, it is possible to combine them in the same line to develop more sophisticated models [202].

Table 1 shows an updated list of cancer models developed through conditional methods in zebrafish.
### Table 1 Application of the conditional transgenic toolbox in zebrafish cancer models

<table>
<thead>
<tr>
<th>Spatial control</th>
<th>Temporal control</th>
<th>Cancer</th>
<th>Promoter</th>
<th>Oncogene</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gal4/UAS</td>
<td>–</td>
<td>Melanoma</td>
<td>kita</td>
<td>GFP-HRAS&lt;sub&gt;G12V&lt;/sub&gt;</td>
<td>[9]</td>
</tr>
<tr>
<td>Gal4/UAS</td>
<td>–</td>
<td>Leukemia</td>
<td>fi1.1</td>
<td>GFP-HRAS&lt;sub&gt;G12V&lt;/sub&gt;</td>
<td>[43]</td>
</tr>
<tr>
<td>Gal4/UAS</td>
<td>–</td>
<td>Chordoma</td>
<td>twhh</td>
<td>GFP-HRAS&lt;sub&gt;G12V&lt;/sub&gt;</td>
<td>[13]</td>
</tr>
<tr>
<td>Gal4/UAS</td>
<td>–</td>
<td>Glioma</td>
<td>ptf1a</td>
<td>marhAKT1</td>
<td>[12]</td>
</tr>
<tr>
<td>Gal4/UAS</td>
<td>–</td>
<td>Glioma</td>
<td>krt5</td>
<td>mCherry-KRAS&lt;sub&gt;G12V&lt;/sub&gt;</td>
<td>[44]</td>
</tr>
<tr>
<td>Gal4/UAS</td>
<td>–</td>
<td>Pancreas</td>
<td>ptf1a</td>
<td>GFP-KRAS&lt;sub&gt;mut&lt;/sub&gt;</td>
<td>[10, 45, 192]</td>
</tr>
<tr>
<td>Gal4/UAS</td>
<td>–</td>
<td>Multiple</td>
<td>ptf4</td>
<td>smoa1-GFP myrhAKT1</td>
<td>[8]</td>
</tr>
<tr>
<td>Cre</td>
<td>– or Heat Shock</td>
<td>Leukemia (T-ALL)</td>
<td>rag2</td>
<td>GFP-mMyc</td>
<td>[14, 15]</td>
</tr>
<tr>
<td>Cre</td>
<td>Heat Shock</td>
<td>MPN</td>
<td>spi</td>
<td>NUP98-HOXA9</td>
<td>[185]</td>
</tr>
<tr>
<td>Cre</td>
<td>Heat Shock</td>
<td>Multiple</td>
<td>b-actin</td>
<td>KRAS&lt;sub&gt;G12D&lt;/sub&gt;</td>
<td>[186]</td>
</tr>
<tr>
<td>–</td>
<td>Heat Shock</td>
<td>Leukemia</td>
<td>hsp70</td>
<td>AML1-ETO</td>
<td>[141]</td>
</tr>
<tr>
<td>–</td>
<td>Heat Shock</td>
<td>Leukemia like lesions</td>
<td>hse</td>
<td>Mycn</td>
<td>[142]</td>
</tr>
<tr>
<td>–</td>
<td>Heat Shock</td>
<td>Ewing’s sarcoma</td>
<td>hsp70</td>
<td>EWS-FLI1</td>
<td>[143]</td>
</tr>
<tr>
<td>–</td>
<td>Tet-On (Dox)</td>
<td>Glioma</td>
<td>krt5</td>
<td>mCherryKRAS&lt;sub&gt;G12V&lt;/sub&gt;</td>
<td>[44]</td>
</tr>
<tr>
<td>–</td>
<td>Tet-On (Dox)</td>
<td>Liver</td>
<td>fabp10</td>
<td>Xmrk</td>
<td>[160]</td>
</tr>
<tr>
<td>–</td>
<td>Tet-On (Dox)</td>
<td>Liver</td>
<td>fabp10</td>
<td>Myc</td>
<td>[161]</td>
</tr>
<tr>
<td>–</td>
<td>Tet-On (Dox)</td>
<td>Liver</td>
<td>fabp10</td>
<td>GFP-kras&lt;sub&gt;G12V&lt;/sub&gt; mCherry-rhoA&lt;sub&gt;(WT)&lt;/sub&gt; mCherry-rhoA&lt;sub&gt;(DA)&lt;/sub&gt; mCherry-rhoA&lt;sub&gt;TDN&lt;/sub&gt;</td>
<td>[162]</td>
</tr>
<tr>
<td>–</td>
<td>Tamoxifen</td>
<td>Leukemia (T-ALL)</td>
<td>rag2</td>
<td>mMyc-ER</td>
<td>[147]</td>
</tr>
<tr>
<td>–</td>
<td>Tamoxifen</td>
<td>Liver (zebrafish liver cell line)</td>
<td>deltaRaf1-ER</td>
<td></td>
<td>[146]</td>
</tr>
<tr>
<td>LexPR</td>
<td>Mifepristone</td>
<td>Liver</td>
<td>fabp10</td>
<td>GFP-kras&lt;sub&gt;G12V&lt;/sub&gt;</td>
<td>[203]</td>
</tr>
</tbody>
</table>

Collection of available cancer models in zebrafish that are based on conditional transgenic methods stating the conditional methods applied, the type of cancer induced and the promoter and oncogene used

**Abbreviations:** DA dominant active, DN dominant negative, Dox doxycycline, MPN myeloproliferative neoplasia, T-ALL T-cell acute lymphatic leukemia, WT wildtype
Future Perspectives

The development of conditional transgenic approaches in zebrafish is already offering a broad variety of tools for diverse expression controls. Moreover, new tools are continuously developed adding even more flexibility to the already large tool box, including the Gal4/UAS analog systems TrpR/tUAS and QF/QUAS, the novel gene silencing shRNA-miRNA, the dre/rox system, CRISPR/Cas9 knock-in system and optogenetic approaches. Additionally, the possibility of a synergistic application of the large variety of tools will expand exponentially the repertoire of conditional approaches in zebrafish cancer models. For example, the CRISPR/Cas9 system is a promising easy tool to generate attP landing sites at desired loci either in non-coding regions with limited effect on adjacent genes or directly adjacent to enhancers or in coding regions in order to generate a library of lines that allows locus independent comparison of gene activities, enhancer studies or endogenous gene labeling/disruption, respectively. Additionally, these attP sites can be used to integrate sequences encoding for (1) stable or reversible, (2) light, hormone or heat tunable (3) single or multicomponent systems. Current studies in zebrafish using mi-siRNA also promise to achieve targeted silencing of genes at will [47]. Thus, the field of zebrafish cancer studies will soon make full use of an immense set of tools for controlling gene expression specifically in organs, tissues, cells or time, or all at once.

Acknowledgements We thank World Wide Cancer Research (grant no: 11_0624) and the “Studienstiftung des deutschen Volkes” for their financial support.

We thank Addgene for the plasmid CMV:cre (31308) [205] and Martin Distel, Reinhard W. Köster, Markus Affolter and Alexander Schier for the transgenic lines shown in Figs. 1 and 2.

References


M. Mayrhofer and M. Mione


Cancer and Zebrafish
Mechanisms, Techniques, and Models
Langenau, D.M. (Ed.)
2016, XII, 552 p. 85 illus. in color., Hardcover
ISBN: 978-3-319-30652-0