Novel Approaches to Transgenesis in the Teleost Medaka (*Oryzias latipes*)

Clemens Grabher

INAUGURAL – DISSERTATION

zur

Erlangung der Doktorwürde

der

Naturwissenschaftlich-Mathemathischen Gesamtfakultät

der

Ruprecht-Karls-Universität

Heidelberg

vorgelegt von

Mag. Clemens Grabher

aus St. Gallen

Tag der mündlichen Prüfung: 15.05.2003

Thema

Novel Approaches to Transgenesis in the Teleost Medaka (*Oryzias latipes*)

Gutachter: PD Dr. Jochen Wittbrodt
PD Dr. Dirk-Henner Lankenau

DISSERTATION

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the

Rupero-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

presented by

Mag. Clemens Grabher

born in St. Gallen

Oral examination: 15.05.2003

Table of Contents

Summary

The aim of thesis was to improve the generation of transgenic medaka fish. General transgenesis including transient expression of reporter genes and germ line integration of reporter genes was improved by application of two novel techniques. In addition, one of these methods allows for the first time efficient enhancer trapping in fish.

First, a transposon-based approach using the artificially reconstructed *Sleeping Beauty* (*SB*) transposon was established.

To address the potential of *SB* for transgenesis, microinjection experiments were performed. Transgenes (*GFP*) and promoter fragments were flanked with the *SB* recognition sequences (inverted direct repeats (IR/DR)) and injected into one-cell stage medaka embryos with or without *SB10* mRNA. Upon injection of a control construct, that lacks *SB* recognition sequences and without transposase, only 13 % of surviving embryos expressed *GFP* uniformly in the entire body. Conversely, when *SB* IR/DRs were included, uniform, promoter-dependent expression was the predominant effect (45 %). The presence of IR/DRs alone strongly enhanced promoter-dependent transgene expression in G0, indicating that *SB* IR/DRs significantly enhance transient transgene expression. G0 expression was a reliable indicator for the efficient selection of transgenic founders. Embryos that exhibit a uniform *GFP* expression in G0 result in the highest yield of transgenic fish. This facilitates an easy selection of putative founder fish for medium- to large-scale approaches. The *SB* system enhanced total transgenesis frequencies to 32 % compared to 4 % resulting from control construct injections. Single copy *SB*-mediated insertion was verified by Southern blot analysis and sequence analysis of flanking genomic sequences. Strikingly, 12 % (21/174) of the transgenics featured typical characteristics of enhancer trap lines, i.e. spatially and/or temporally restricted transgene expression due to regulation imposed by sequences adjacent to the insertion site. Among 21 lines with novel expression patterns, a variety of different patterns ranging from single cell types to whole organs were found. Thus, a set of transgenic lines expressing *GFP* in developmentally important structures/organs can be

established and used without devoting a major effort on the isolation and characterization of promoter elements.

Second, a meganuclease approach was applied. Transgenes of interest were flanked by two *I-SceI* meganuclease recognition sites, and co-injected together with the *I-SceI* meganuclease enzyme into medaka embryos at the one-cell stage. The recognition site comprises 18 bp that are asymmetrically cleaved, rendering it a very rare cutter (app. once in $7x10^{10}$ bp of random sequence).

Upon injection, the promoter-dependent expression was strongly enhanced. Already in G0, 78 % of injected embryos exhibited uniform promoter dependent expression compared to 26 % when injections were performed without meganuclease. The transgenesis frequency was raised to 30.5 % compared to 5-18 % for naked DNA. Even more striking was the increase in germ line transmission rate. In standard protocols it does not exceed a few percent, the number of transgenic F1 offspring of an identified founder fish generated with *I-SceI* reached the optimum of 50 % in most lines, indicating genome insertion events already at the one-cell stage. Southern blot analysis showed that individual integration loci contain only one or few copies of the transgene in tandem. Meganuclease co-injection thus provides a simple and highly efficient tool to improve transgenesis by microinjection.

Introduction

The introduction of genes into the germ line of animal or plant model systems is one of the major technological advances in modern biology. Transgenic animals have been instrumental in providing new insights into mechanisms of development and developmental gene regulation, into the action of oncogenes and into the intricate cell interactions within the immune system. Furthermore, the transgenic technology offers exciting possibilities for generating precise animal models for human genetic diseases and for producing large quantities of economically important proteins by means of genetically engineered farm animals and fish.

The ectopic expression of transgenes in whole animals allows one to study gain-of-function phenotypes. Alternatively, disruption of endogenous genes by random transgene insertion or through targeted homologous recombination allows the study of loss-of-function phenotypes, an approach that allows elucidating the biological role of a gene. Transgenic technology is often used as a tool for identifying mutant genes after they have been mapped to specific chromosomal loci (Antoch et al., 1997). By employing reporter genes under the control of specific regulatory sequences, transgenic techniques facilitate the functional dissection of the *cis*-acting elements responsible for spatial and temporal gene expression patterns. In addition, tissues or cells expressing a reporter transgene can be used in cell lineage analysis and transplantation experiments.

The establishment of methods to introduce exogenous genes into organisms, to transmit these genes to the next generations, and to direct proper transgene expression is one of the basic and indispensable criteria for an organism to be referred to as model organism.

1. Medaka - a Model System for Vertebrate Developmental Genetics

Teleosts, such as the medaka, the pufferfish and the zebrafish, are increasingly popular vertebrate model systems in various fields of biology (Kimmel, 1993; Venkatesh et al., 2000; Westerfield, 1995; Wittbrodt et al., 2002; Yamamoto, 1975). Two major reasons for their popularity are the relative ease of their application in forward genetics and the relatively small genome size (medaka and pufferfish).

Medaka, *Oryzias latipes*, is a small egg-laying freshwater fish native to Asia that is found primarily in Japan (Fig. 1A). The adult fish are about 3 cm long and can tolerate a wide range of temperatures (4-40 °C). The male medaka can easily be distinguished from the female by a clearly dimorphic dorsal fin and, once fertilized, the female spawns a cluster of 20-40 eggs every day. Both eggs and embryos are transparent and encased in a hardy chorion (Fig. 1B); consequently, the morphology of the developing can easily be evaluated. Embryos hatch as young feeding fry 7 days after fertilization and sexually mature within 6-8 weeks under laboratory conditions.

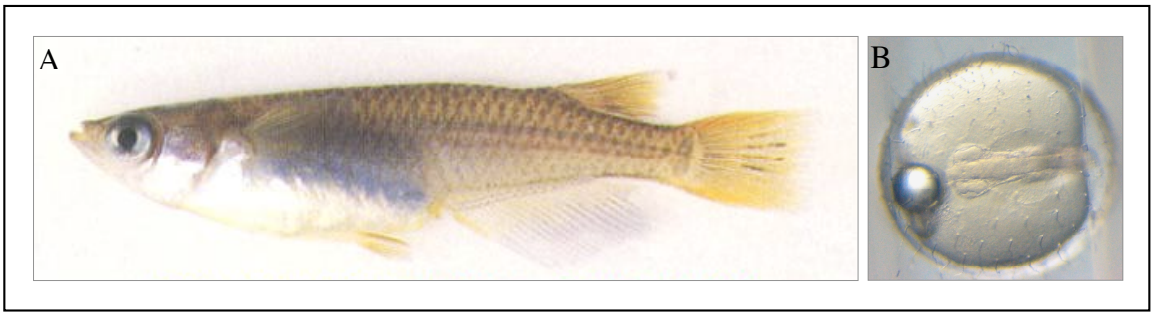


Fig. 1: The medaka fish.
A, Lateral view of an adult, male medaka. B, Dorsal view of a medaka embryo at developmental stage 21 (brain and otic vesicle formation). A and B belong to the inbred Cab strain of the southern population.

The physiology, embryology and genetics of medaka have been studied extensively for the past 100 years (Yamamoto, 1975). Important advances in medaka research include the establishment of inbred strains (Hyodo-Taguchi and Egami, 1985) and the development of transgenesis protocols (Ozato et al., 1986). The development of mutagenesis protocols (Shima and Shimada, 1991) led to the first systematic mutagenesis screens for developmental phenotypes (Loosli et al., 2000) and, in

combination with detailed descriptions of medaka anatomy (Anken and Bourrat, 1998; Ishikawa, 1997; Ishikawa, 2000), they have led to the characterization of many mutant phenotypes that were recovered from these screens. Genomic resources and a detailed linkage map also facilitated cloning of the genes that are responsible for these mutant phenotypes (Fukamachi et al., 2001; Kondo et al., 2001; Loosli et al., 2001), which shows the power and the potential of medaka as a genetic model system.

Considering the evolutionary distance between zebrafish and medaka of about 160 million years to their last common ancestor, which is also reflected in many aspects of their biology, medaka provides an ideal resource for comparative studies. Studies of distantly related vertebrate species permit the identification of conserved and species-specific molecular mechanisms underlying development and evolution.

2. Transgenesis in Fish

The transgenic technology was first applied to fish in the mid- 1980's (Zhu et al., 1985). Since then, transgenic fish have been widely used in both basic and applied research. About 15 years ago it appeared like it would be pretty simple to deliver exogenous DNA to medaka or zebrafish chromosomes; after all, the single fertilized cell was unusually conspicuous, large, and encased in an optically clear chorion permitting rapid microinjection of DNA solution. Alas, transgenic technology in both model organisms has been deceptively difficult. The delivery of DNA into the cell was about as trivial an exercise as predicted, but the DNA generally failed to integrate into the host genome (Hackett, 1993; Iyengar et al., 1996; Westerfield et al., 1992). Consequently, there were attempts to find mechanisms to deliver DNA more efficiently into fish chromosomes. However, microinjection of plasmid DNA has proven to be a reliable means of producing transgenic fish and remains the most widely employed method, partially also due to the relative ease to perform this method and the lack for special prerequisites.

2.1 Methods of Transgene Delivery

Only techniques that yielded significant success in the generation of stable transgenic fish will be discussed in this thesis, including retroviral infection (Lin et al., 1994a), the use of micro-projectiles (Yamauchi et al., 2000; Zelenin et al., 1991) and electroporation (Inoue et al., 1990). Applications that have not been successfully applied for germ line transgenesis in fish like lipofection (Szelei and Duda, 1989) or sperm carriers (Chourrout and Perrot, 1992; Lavitrano et al., 1989), although useful for transient studies, will not be discussed.

2.1.1 Retroviral Infection

Retroviruses use RNA as their genetic material. Following infection, the RNA is transcribed by the virus-encoded reverse transcriptase. The resulting single-stranded DNA is replicated as double-stranded DNA (dsDNA). The dsDNA viral genome then integrates into the host genome in an essentially random fashion (Smith, 2002).

The Hopkins lab developed a pseudotyped retroviral vector that contains a Moloney murine leukaemia virus-based genome packaged in the envelope protein of the vesicular stomatitis virus (Burns et al., 1993; Lin et al., 1994a). Virus pseudotype particles need to be injected into blastula-stage embryos (512-2000 cells) because the viruses are not able to penetrate the chorion. Proviruses are able to integrate into the germ lines of zebrafish at less than 0.1 % efficiency (Allende et al., 1996; Gaiano et al., 1996a) and because of the delay in the delivery of the virus (blastula stage), the embryos obtained by this protocol are highly mosaic. Nevertheless, this strategy has been very successfully used for insertional mutagenesis (Allende et al., 1996; Amsterdam et al., 1997; Gaiano et al., 1996b; Golling et al., 2002). However, there are some drawbacks to these viruses (Ivics et al., 1999). First, because so little volume can be injected, it is necessary to acquire very high titres of virus, which is not trivial. Second, retroviral vectors exhibit difficulties to deliver genes that are stably expressed over several generations although improvements were achieved recently (Linney et al., 1999). Third, the virus must be contained and handled with extreme care. Accordingly, alternative means of gene delivery were needed that would have an early and narrow window of activity to reduce mosaicism, have a high efficiency of integration and be safe and easy to use by any lab.

2.1.2 Electroporation

Electroporation is a process by which high intensity electric field pulses temporarily destabilize cellular membranes. During the destabilization period, DNA molecules present in the surrounding media are able to permeate the cell's external and internal membranes to enter the cytoplasm and nucleoplasm (Lurquin, 1997). The

electroporation process can be equilibrated to yield copy numbers (of integrated transgenes) of between 1 and 20 copies per genome – an advantage compared with microinjection (see below). In addition, large transgene molecules (> 150 kb) can be transferred. The requirement for specialized equipment and extensive optimization procedures for different systems have to be mentioned as major drawbacks.

Since 1990 electroporation has been finding greater favour for transgenesis in fish and some success has been reported (Inoue et al., 1990 ; Ono et al., 1997). However, in recent years electroporation of fertilized fish eggs has been more widely used to perform transient transgenesis than to generate transgenic germ lines (Sussman, 2001; Tawk et al., 2002). To facilitate germ line transgenesis, electroporating sperm before fertilization, represents an interesting variation on the electroporation technique (Muller et al., 1992; Sin et al., 2000). However, integration of the foreign DNA occurs infrequently, and the expression of the foreign genes is poor. The potential of sperm-mediated gene transfer as a routine protocol for mass gene transfer in fish will be dependent on the improvement of integration and expression of the foreign gene.

2.1.3 Particle Bombardment

It is worth considering a fairly novel technique, as an illustration of the many and varied means by which emerging technologies are enabling gene transfer. In particle bombardment, DNA may be adsorbed to spherical tungsten or gold particles and transferred into cells by a particle gun. Once inside the target cell, the DNA is solubilised and may be expressed (Klein and Fitzpatrick-McElligott, 1993). This approach has been originally developed for plant transgenesis but has been shown to be effective for transferring transgenes into animal cells *in vivo* (Cheng et al., 1993).

The procedure has been adapted to fertilized zebrafish eggs when transgenes have been successfully delivered and expressed in the targeted embryos (Zelenin et al., 1991). Only more recently transmission of transgenic green fluorescence protein (*GFP*) to the germ line of medaka embryos has been achieved resulting in true transgenic F1 offspring (Yamauchi et al., 2000). However, the amount of research data presently available is too little to permit definitive comparisons to other techniques.

2.1.4 Embryonic Stem (ES) cells

Homologous recombination was first used in yeast and later in mouse to directly alter the sequence of a known gene, which is known as gene targeting (Capecchi, 1989). It does occur also upon plasmid injection into pronuclei of mouse oocytes (Brinster et al., 1989), but with a very low frequency of 10^{-3} as compared to non-homologous integration events. However, in the mouse system this problem is solved by electroporating thousands of embryonic stem cells. Subsequent positive/negative selection enables identification of proper homologous recombination events (Joyner et al., 1991).

In medaka, ES cells (Mes1 for medaka embryonic stem cells) have been established (Hong et al., 1996) and were found to contribute to organs of all three germ layers in chimeras (Hong et al., 1998). However, generation of stable transgenic fish has not been successful, due to the failure of ES cells to populate the germ line. Although, cell cultures exhibiting characteristics of ES cells have been described in zebrafish, only short-term cell cultures, which must be maintained in the presence of cells from the rainbow trout, have produced germ line chimeras (Ma et al., 2001). It remains to be determined if these cells will contribute to the germ line after long-term culture, which is required for genetic manipulations involving homologous recombination and selection.

2.1.5 Nuclear Transfer

As an alternative to embryonic stem cells, cultured somatic cells offer the possibility of producing cloned animals with targeted genetic manipulations (Lai et al., 2002; McCreath et al., 2000). Since the successful cloning of ‘Dolly’ using a somatic nucleus (Wilmut et al., 1997), several successful cloning experiments using somatic cells have been achieved, including recent reports describing gene-knockout sheep and pigs produced by nuclear transfer from genetically manipulated somatic cells (Lai et al., 2002; McCreath et al., 2000).

Fish nuclei of blastula cells from different genera have been transplanted into enucleated eggs to study the nucleo-cytoplasmic interaction (Zhu and Sun, 2000). Wakamatsu and colleagues have demonstrated that diploid fertile medaka could be produced by nuclear transfer using blastula cells as donors (Wakamatsu et al., 2001). These findings show that nuclei prepared from fresh blastula cells can be reprogrammed in fish to support embryonic and adult development. In 2002, the first cloned zebrafish using long-term cultured cells, amenable to genetic manipulation, was established (Lee et al., 2002). Although, the current success rate of ~2 % does not represent an improvement on transgenesis, the potential availability of cell cultures that can be used for homologous recombination could pave the way for the gene targeting in lower vertebrates.

2.1.6 Microinjection

In 1980, Gordon and co-workers demonstrated that exogenous DNA could be introduced into the mouse genome simply by physical injection of DNA solution into the zygote (Gordon et al., 1980). Subsequently, microinjection has become and remained the most widely used method of germ line transgenesis in several species including fish.

For medaka, a finely drawn glass needle, loaded with DNA solution, is used for the injection. Under a common dissecting microscope, with the aid of a micromanipulator, fertilized eggs are penetrated with the needle. The injection needle is guided through the chorion into the cytoplasm of the cell of an embryo at the one cell stage. Once the tip of the needle has entered the cytoplasm, approximately 1-2 nl of DNA solution containing 10^5 to 10^7 DNA molecules is injected.

The first transgene to be delivered into medaka embryos was the □-crystalline gene of chicken (Ozato et al., 1986). Transient expression of the transgene occurred in a mosaic manner but no germ line transmission was observed. It was only in 1988 when transgenesis by microinjection was successfully performed including transgene expression and transmission to the next generation in a teleost genetic model system (Stuart et al., 1988). Presently, microinjection provides the fastest and simplest means

for germ line transgenesis and transient expression studies in fish (Chou et al., 2001; Lin, 2000). However, techniques widely used in mouse and *Drosophila* such as enhancer- and gene-trapping (Allen et al., 1988; Gossler et al., 1989; Korn et al., 1992; O'Kane and Gehring, 1987; Rubin and Spradling, 1982), although attempted with limited success in zebrafish (Bayer and Campos-Ortega, 1992) and frog (Bronchain et al., 1999), have rarely been used due to low frequency of vector integration into the fish genome. Consequently, technological improvements on transgenesis by microinjection need to be met.

In order to develop enhancing strategies, one has to consider the fate of injected DNA inside a cell. Influencing the fate of a transgene presents an obvious way to alter both, transient transgene expression and integration into the host genome.

2.2 General Fates of Injected Transgenic DNA

When plasmid DNA is injected into medaka or zebrafish embryos, it may meet with several different fates. (1) It may replicate and persist in the cell and its descendants for several cell divisions. (2) It may integrate into the chromosomal DNA of the cell or (3) the plasmid DNA may be lost from the embryo. Commonly, the first two fates lead to embryos that are mosaics with respect to the presence of plasmid DNA. In case of (1) mosaicism, the presence and/or expression of the transgene is due to the uneven distribution and replication of the episomal DNA among daughter cells (Fig. 2A and 10). Nearly all (90-99 %) fish that have integrated transgenes (2) will also be mosaic for its presence and/or level of expression due to integration events occurring later than the one cell stage (Hackett, 1993). Both, transient expression of the transgene and germ line transmission are dependent on the time-point of integration. The later an insertion event occurs the fewer somatic cells will contain an integrated transgene that can be transmitted equally to descendent cells, directly influencing the pattern of transient transgene expression (Fig. 2B, C and 10). Similarly, not all germ cell precursors may have integrated the injected DNA leading to a mosaic germ line. If

the germ line is mosaic, the proportion of transgenic F1 progeny depends on the degree of mosaicism.

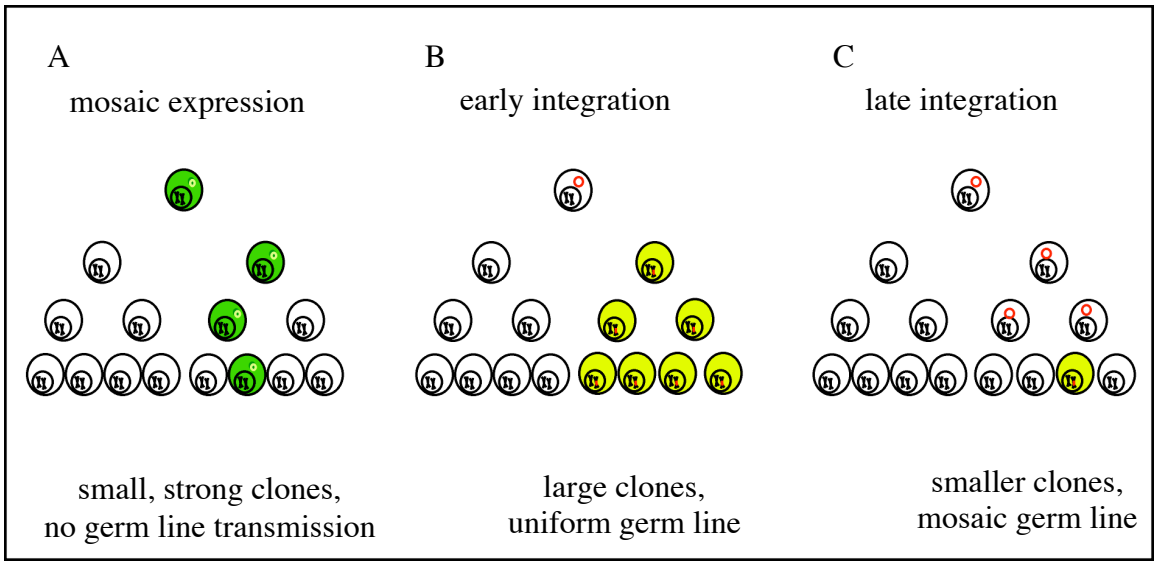


Fig. 2: Fate of injected DNA.
Injected DNA may meet three different fates. A, DNA stays episomal (probably in concatemers) and is expressed in small bright clones, due to uneven segregation. B, DNA integrates early in development (1-2 cell stage). Depending on copy number of inserted transgenes its expression level may vary but the germ line will be uniform, resulting in a large proportion of transgenic F1 progeny. C, DNA integrates at later stages. Mosaicism of both, transgene expression and germ line depends on the time-point of insertion.

Only genomic integration within the one cell stage will undoubtedly lead to a completely transgenic germ line resulting in 50 % of the F1 offspring expected to inherit the transgene (Jowett, 1999).

2.2.1 Immediate Fate of Injected DNA – Transient Expression

Early developmental processes in injected embryos mainly determine the fate and consequently the expression patterns of exogenous DNA. In fish, the period following fertilization is characterized by a series of rapid cleavages with no distinct growth phases and almost no transcription (Iyengar and Maclean, 1995). Subsequently, the embryo enters the mid-blastula transition (MBT, in medaka at ~1024 cell stage) coinciding with an abrupt loss of cell synchrony, massive upregulation of transcription,

onset of cell motility and elongation of the cell cycle (Aizawa et al., 2003; Andeol, 1994). The unfertilized fish egg contains a large store of proteins, provided by the mother, to support the rapid cell cleavages including chromatin assembly proteins, ligases, polymerases, etc. These proteins are thought to be responsible for the concatemerization (formation of tandem arrays) and the extrachromosomal replication of injected DNA (Vielkind, 1992; Volckaert et al., 1994).

Concatemers arising from circular injected DNA are largely arranged in a head to tail fashion (Vielkind, 1992) while linear DNA is arranged randomly. For both concatemerization has been found to happen very rapidly in medaka (Chong and Vielkind, 1989). This short time renders induction of ligase production unlikely. It is perhaps instead a direct result of the activity of stored ligases within the egg. Replication of exogenous episomal DNA in the early fish embryo correlates with the rapid DNA synthesis during these first cleavage stages. Thus, replication rates are higher than degradation rates causing an increase of injected DNA. Only at post-gastrula stages, a widespread degradation of foreign extrachromosomal DNA is observed (Volckaert et al., 1994).

Expression of injected DNA has been found to be highly mosaic using both, ubiquitous or specific regulatory elements (Stuart et al., 1990; Tsai et al., 1995). This phenomenon is attributed to unequal distribution of transgene copies (Houdebine and Chourrout, 1991) and differential replication. The latter may be the main reason for variable expression observed within multinucleated or polyploid tissues such as muscle cells or the yolk syncytial layer (YSL) (Williams et al., 1996). Transient expression in fish has been found to follow distinct temporal patterns but is almost invariably found to begin after the MBT stage. Repression of transcription prior to the MBT may be due to a large excess of histones for use during the rapid early developmental stages (Prioleau et al., 1994). Later, lengthening of the cell cycle and accumulation of newly synthesized DNA results in a titration of excess histones allowing the onset of transcription. These temporal expression patterns relate clearly to the persistence of the injected DNA. Highest expression levels are observed at gastrula stages because of extensive (also episomal) DNA replication during early cleavages and accumulation of enzymes subsequent to the MBT. Decreasing levels of transient transgene expression in

later rounds of cell divisions result from transgene degradation (Volckaert et al., 1994).

2.2.2 Late Fate of Injected DNA – Stable (Genomic) Expression

Exogenous DNA, after surviving the degradative processes of the cell, may integrate into the host's chromosomal DNA. The mechanism of transgene insertion is largely unknown but is thought to involve a process of end-joining or illegitimate recombination depending on random breaks in the chromosomes (Bishop and Smith, 1989). However, the existence of certain loci, which may be more amenable to transgene insertion, has been proposed as well (Sutherland et al., 1993).

Confirmation of transgene integration has largely been observed by Southern blot analysis. Since the million copies of transgenic copies can ligate to form high molecular weight sequences, today more rigorous proof of integration is demanded. The report of Stuart (Stuart et al., 1990) showed that extrachromosomal DNA could be transmitted to the F1 progeny. Thus, the only unambiguous evidence for transgene integration is given by the demonstration of classic Mendelian inheritance to 50 % of the F2 progeny upon crossing of transgenic F1 to wild type fish. An alternative proof can be obtained by sequencing of junction fragments between transgene and chromosomal sequences.

What is apparent from the literature is that chromosomal integration commonly occurs late in development, resulting in highly mosaic founder fish and low frequencies of germ line transmission (Fig. 2B, C and 11). Although single site insertion is frequently observed (Stuart et al., 1990; Tewari et al., 1992), multiple integration events are also common (Culp et al., 1991). Another characteristic of integrated transgenes represents its concatemerization that may lead to insertion of tandem arrays of up to 2000 in number (Tewari et al., 1992). Disappointingly, integration of transgenes has been found to affect its expression pattern in many unpredictable ways. For example, it has been frequently observed that transgenes are influenced by neighbouring sequences, a phenomenon called position effect. In addition, there have also been problems with the silencing of transgenes as a result of mechanisms such as DNA methylation and heterochromatin formation. It is notable that transgene

integration as an intact single copy as opposed to a tandem array of multiple copies may be desirable, since there is some evidence that tandem arrays may be more readily inactivated by DNA methylation or heterochromatin formation (Dorer and Henikoff, 1002; Garrick et al., 1998; Mehtali et al., 1990).

It is clearly important to avoid position effects if well-regulated expression of a transgene is required (see also section 2.3). The inclusion of introns has been found to enhance expression of transgenes (Clark et al., 1993; Palmiter et al., 1991). Stretches of DNA present at the boundaries of the ~20 kb chicken lysozyme gene locus have been observed to insulate transgenes from position effects in mammalian cells and mice (McKnight et al., 1992; Stief et al., 1989). It has been proposed that the presence of matrix or scaffold attachment regions (MARs or SARs) is responsible for these insulating effects by enabling the looping out of domains (Sippel et al., 1992). Nevertheless, there are examples of insulating regions not consisting of MARs or SARs (Kellum and Schedl, 1992; Noma et al., 2001) suggesting other possible mechanisms.

Silencing of transgenes following passage to F1 and beyond has been noted by several groups (Culp et al., 1991; Stuart et al., 1988; Stuart et al., 1990; Tewari et al., 1992). Transcription of RNA polymerase II genes can be inhibited by methylation of cytosine residues at CpG sites (Eden and Cedar, 1994). Gibbs and co-workers suggested that low levels of expression in F1 and F2 progeny of transgenic zebrafish could be due to methylation of a number of sites within the transgene. They were able to increase the expression levels by treating the embryos with 5-azacytidine, a potent inhibitor of methylation (Gibbs et al., 1994b).

2.3 Strategies to Improve Transgenesis by Microinjection

Major drawbacks, e.g. mosaic transgene expression in G0, low insertion frequency and mosaic germ line distribution have not yet been overcome. Moreover, the transgenesis frequencies upon microinjection are still very variable depending on the vector used and on the skills of the injector. Average stable transgenesis frequencies range within 1-10 % (Collas and Alestrom, 1998; Culp et al., 1991; Lin et al., 1994b;

Stuart et al., 1988; Stuart et al., 1990; Tanaka and Kinoshita, 2001) only exceptionally reaching more than 20 % (Higashijima et al., 1997). Similarly, efficiencies of transient expression of a transgene in the G0 generation vary between 10-50 % (Chou et al., 2001; Higashijima et al., 1997) but are invariably mosaic. Generally, comparison between different reports is difficult due to the differences of promoters and/or vector design. Strategies of vector design affecting the transgenesis quality will be discussed below. Novel strategies that have been applied in this thesis will be introduced in section 2.4 and thereafter.

2.3.1 Nuclear Localization Signal (NLS)

A key to transgenesis lies in efficient uptake of foreign genes by the cell nucleus. Recent studies have shown that a limiting step in fish transgenesis resides in slow nuclear import of DNA (Collas and Alestrom, 1997; Collas and Alestrom, 1998), a situation likely to favour late and mosaic transgene integration into the germ line (Culp et al., 1991; Stuart et al., 1988). Improvements in nuclear uptake of DNA have resulted from the use of protein-DNA complexes. Non-covalent attachment of DNA to karyophilic proteins including NLS peptides (CGGPKKKRKVG-NH_2) has been shown to enhance nuclear import and expression of DNA in cultured mammalian cells and zebrafish (Collas and Alestrom, 1997; Fritz et al., 1996; Kaneda et al., 1989), whereas covalent cross-linking of NLS peptides to DNA proved ineffective (Sebestyen et al., 1998). In medaka and zebrafish reports on the use of non-covalent NLS peptide applications are somewhat conflicting. Although Collas and co-workers reported enhanced integration and expression, for other authors the effect of NLS peptide remained elusive (Higashijima et al., 1997). However, the rationale behind this approach appears logic and even if enhanced genomic integration of NLS-DNA complexes is not necessarily a consequence of nuclear uptake, a more even segregation of the foreign DNA resulting in more uniform, transient expression might be expected. In conclusion, the NLS approach has not become a commonly used method so far, indicating that improvements on both transient expression and transgenesis frequency have not been sufficient.

2.3.2 Restriction Endonuclease Mediated Integration (REMI)

Another promising technology that has been applied successfully in *Dictyostelium* (Kuspa and Loomis, 1992) and *Xenopus* (Kroll and Amaya, 1996; Kroll and Gerhart, 1994) involves the use of restriction endonucleases. In frogs, linearized plasmid DNA and a restriction endonuclease, usually a 6 bp cutter (i.e. *XbaI,* etc), are incubated together with sperm nuclei. Nuclei are then swelled and partially decondensed in an interphase egg extract. By this means, plasmid DNA is introduced into the nuclei and eventually integrates into the sperm genome utilizing endogenous DNA repair mechanisms. Single nuclei are then transplanted into unfertilized eggs. A variable fraction of these eggs (5-40 %) develops normally to the tadpole stage (Kroll and Amaya, 1996). Transgenic embryos from gastrula to tadpole stages express the transgene non-mosaically. In these experiments, the concentration of a frequently cutting restriction enzyme is a very critical parameter. Amounts of enzyme that do not exhibit deleterious effects due to fragmentation of the genome failed to result in any significant improvement of transgenesis in medaka (J. Wittbrodt unpublished data).

A variation of this strategy, although based on a different mechanism, is the application of extremely rare cutting enzymes (meganucleases) that do not cleave within the genome but in the transgene vector only. By co-injection into fertilized medaka embryos, thus skipping the transplantation step, this technique provides a powerful means for fish transgenesis that will be presented in this thesis.

2.3.3 Boundary Regions

To avoid position effects, DNA methylation or changes of chromatin state affecting transgene expression it is desirable to supply the transgene with regulatory sequences. The importance of the inclusion of enhancer elements, introns and appropriate polyadenylation (pA) signals is generally accepted (Iyengar et al., 1996). Moreover, genes are thought to be organized on chromosomes as contiguous but independent units known as expression domains (Elgin, 1990). These expression domains are believed to remain insulated from neighbouring sequences by boundary

regions. A feature commonly linked with such boundary elements is the ability to protect against mechanism that affect gene expression as mentioned above (Kellum and Schedl, 1991; Noma et al., 2001; Stief et al., 1989).

Inverted terminal repeats (ITR) from adeno-associated virus (AAV) have been used to improve transient transgene expression and insertion in mammalian cell culture, frog, zebrafish and medaka (Chou et al., 2001; Fu et al., 1998; Hsiao et al., 2001; Philip et al., 1994). AAV type 2, a non-pathogenic human virus, has a single-stranded DNA genome with characteristic ITRs (Srivastava et al., 1983). Each ITR consists of 145 nucleotides of which the terminal 125 bp form palindromic hairpin structures that serve as primers for AAV DNA replication. These hairpin structures also play a role in DNA integration (Philip et al., 1994). It has been suggested that inverted terminal repeats direct the injected DNA to the nucleus and thereby facilitate equal distribution of extra chromosomal DNA to daughter cells (Fu et al., 1998; Hsiao et al., 2001). In addition, Noma and co-workers identified inverted repeats acting as barriers for heterochromatin spreading in fission yeast (Noma et al., 2001). However, using this technique, DNA tends to remain episomal and the vectors are difficult to prepare due to the presence of potentially deleterious repeated sequences (Chou et al., 2001).

2.3.4 Transposable Elements

Transposable elements are discrete segments of DNA capable of moving from one locus to another in their host genome or between different genomes. They are distributed across the living world and play a fundamental role as motors of genome plasticity in all three classical biological kingdoms.

One of the major distinguishing features of transposable elements is whether their transposition relies exclusively on DNA intermediates or includes an RNA stage (Tab. 1). DNA elements (transposons and insertion sequences) can be found in both prokaryotes and eukaryotes, whereas those with RNA intermediates (viral and non-viral retrotransposons) are restricted to eukaryotic genomes.

2.3.4.1 RNA Elements

RNA elements can be divided into those that carry long-terminal repeats (LTR) (viral retrotransposons), those that do not (non-viral retrotransposons) and retroviruses (Fig. 3). This classification is based on their mode of amplification and their general characteristics. The common feature of all retroelements is their ability to generate multiple DNA copies that can be integrated into new chromosomal positions (Andersson et al., 1998).

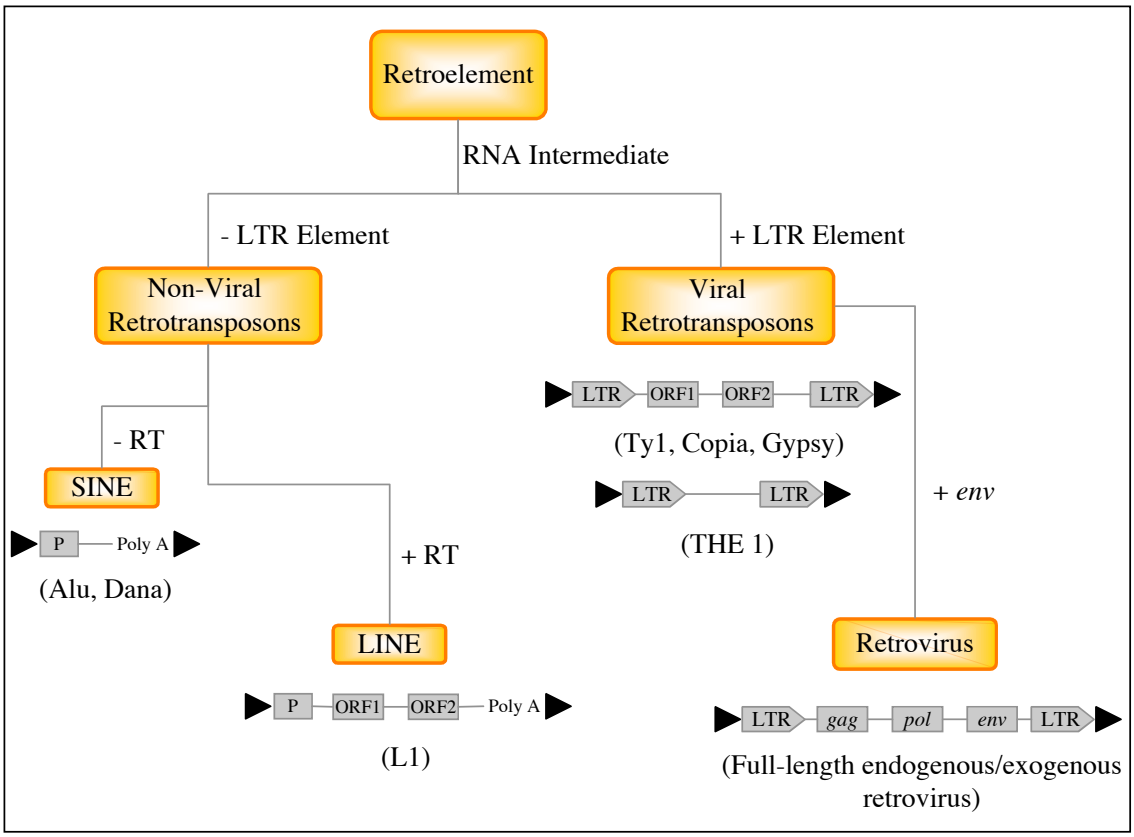


Fig. 3: The structure and relation ship between retroelements.
Triangles represent the short direct repeats flanking the retroelements. Internal promoters (P), LTR elements, open reading frames (ORF), genes and poly A tails are indicated. (adapted from (Andersson et al., 1998)

The presence of envelope (*env*) genes distinguishes retroviruses from viral retrotransposons. Some truncated endogenous retroviruses can be classified to belong to the large family of viral retrotransposons because they lack functional *env* genes. In some species, viral retrotransposons have been extensively characterized, like the

human *THE-1* element, the *Copia* elements in *D. melanogaster* and the *Ty1* of *S. cerevisiae*. The sequence similarity of viral retrotransposons to retroviruses, the presence of reverse transcriptase (RT) and LTRs, imply that they might be the predecessors of retroviruses or alternatively, truncated endogenous retroviruses.

Non-viral retrotransposons lack both, functional *env* genes and LTRs. They are subdivided, based on the presence or absence of RT into long interspersed elements (LINEs) and short interspersed elements (SINEs). None of the retrotransposons has been used in transgenesis approaches yet. As mentioned in previous sections (2.1.1 Retroviral Infection), only retroviruses have been applied successfully in mutagenesis screens (Golling et al., 2002).

2.3.4.2 DNA Elements

DNA transposons are characterized by the fact that their transposition does not involve an RNA intermediate, but occurs in form of DNA. All DNA transposons contain inverted repeats (IRs) at either end flanking a central region encoding for a transposase, which catalyzes transposition. The transposase is expressed at very low levels to strictly control the transposition to low frequencies. A transposition frequency too high would cause a high rate of insertional mutagenesis resulting in a lower evolutionary fitness of the host and therefore of the transposon.

If the transposase is functional, the transposon is termed autonomous. If due to mutations the transposase is inactive, the transposon is termed non-autonomous, as it is dependent on other transposons to provide a functional transposase for its transposition. Most DNA transposons move via a non-replicative cut-and-paste mechanism, but some exceptions transpose replicative via a DNA copy (Tab. 1).

Eukaryotic transposons move in a non-replicative manner via excision and integration (cut-and-paste mechanism). For transposition, the transposase binds to the IRs and cleaves the DNA, thereby precisely excising the transposon. Additionally the transposase binds to the target DNA introducing a staggered cut leading to protruding single strands. The DNA repair machinery accomplishes integration into the host genome, resulting in target site duplications directly adjacent to the transposon ends. The length of the direct repeats (DRs) is characteristic for the specific transposon.

The P-element of *Drosophila melanogaster* is the most famous DNA transposon, used for generating P-element insertion lines. P-element insertions can lead to regulated reporter gene expression through an adjacent enhancer (Bellen et al., 1990; O'Kane and Gehring, 1987) and can cause mutations (Cooley et al., 1988; Sentry and Kaiser, 1992). Genomic regions flanking the insertion can be cloned by plasmid rescue (Bellen et al., 1990). In the *Drosophila* P-element system, other factors than the transposase are involved in the transposition event.

Type	Structural Features	Mechanism of Movement	Examples
DNA-Mediated Transposition			
Bacterial Insertion Sequences (IS)	~50 bp inverted repeats flanking the transposase and/or resolvase	Excision or copying of DNA and its insertion at target site	IS1, IS10
Bacterial Transposons	Central antibiotic resistance gene flanked by IS-elements	Copying of DNA and its insertion at target site	Tn9
Eukaryotic Transposons	Inverted repeats flanking coding region	Excision of DNA and its insertion at target site	P-elements, TcEs, Ac-and Dc-elements
RNA-Mediated Transposition			
Viral Retrotransposons	~250 to 600 bp direct terminal repeats (LTRs) flanking reverse transcriptase, integrase, and retroviral-like Gag protein	Transcription into RNA from promoter in left LTR by RNA polymerase II followed by reverse transcription and insertion at target site	Ty elements, Copia elements
Non-Viral Retrotransposons	Of variable length with a 3' AT-rich region; full-length copy encodes a reverse transcriptase	Transcription into RNA from internal promoter; folding of transcript to provide primer for reverse transcription followed by insertion at target site	F and G elements, LINE and SINE elements, Alu sequences

Tab. 1: Major types of mobile DNA elements. (adapted from Lodish *et al.* 2001)

Therefore, attempts to use the P-element for transgenesis of non-drosophilid insects, zebrafish and mammalian cells have been unsuccessful (Gibbs et al., 1994a; Handler et al., 1993; Rio et al., 1988). In medaka, the DNA-transposon *Tol2* has been found which is a member of the hAT superfamily (hobo of *Drosophila melanogaster*, Ac of maize and Tam3 of the snapdragon (Calvi et al., 1991)). This transposon has inserted into the tyrosinase gene of the *albino* mutant and has been shown to be active during medaka embryogenesis (Koga et al., 1995; Koga et al., 1996). Due to its activity, it is not a candidate tool for transgenesis in medaka. However, it has been successfully applied in zebrafish (Kawakami et al., 2000).

The application of an artificially reconstructed member of the Tc1/mariner family of DNA transposons (Ivics et al., 1997) for transgenesis in medaka, together with a more detailed introduction into this transposon family, will be presented below.

2.4 Novel Strategies to Improve Transgenesis by Microinjection

Two different strategies to improve transgenesis in medaka have been tested in this thesis. The first technique is based on the use of a very rare cutting restriction endonuclease (meganuclease). This approach has been successfully applied to enhance both, transgene expression in the G0 generation and transgenesis frequency. The second technique involves *Sleeping Beauty (SB)*, an artificially reconstructed member of the Tc1/mariner family of transposons. This approach again improved transient and stable transgenesis in medaka. In addition, *SB*-mediated transgenesis allows efficient enhancer trapping resulting in transgenic fish that exhibit spatially and temporally restricted patterns of *GFP* expression. These transgenic fish lines provide valuable tools to biologists interested in the development and/or function of specific organs or cell populations.

2.4.1 The I-SceI Meganuclease

I-SceI is a homing endonuclease encoded by the mobile group I intron of the large rRNA gene of *S. cerevisiae* (Beylot and Spassky, 2001; Jacquier and Dujon, 1985). This family of enzymes mediates the propagation of the intron by cutting intron-less genes at the site of the intron insertion. Like restriction enzymes, homing endonucleases cleave double-stranded DNA with high specificity in the presence of divalent metal ions. However, they differ from restriction endonucleases in their recognition properties and structures, as well as in their genomic location (Belfort and Roberts, 1997). In particular, whereas restriction enzymes have short recognition sequences (3-8 bp), homing endonucleases, despite their small size, recognize long DNA sequences (12-40 bp). They have been classified into four families based on both their sequence motifs and DNA cleavage mechanism (Mueller et al., 1993). The protein *I-SceI* is a member of the largest class of homing enzymes, characterized by the presence of either one or two conserved 12 amino acid residue sequence motifs (LAGLIDADG).

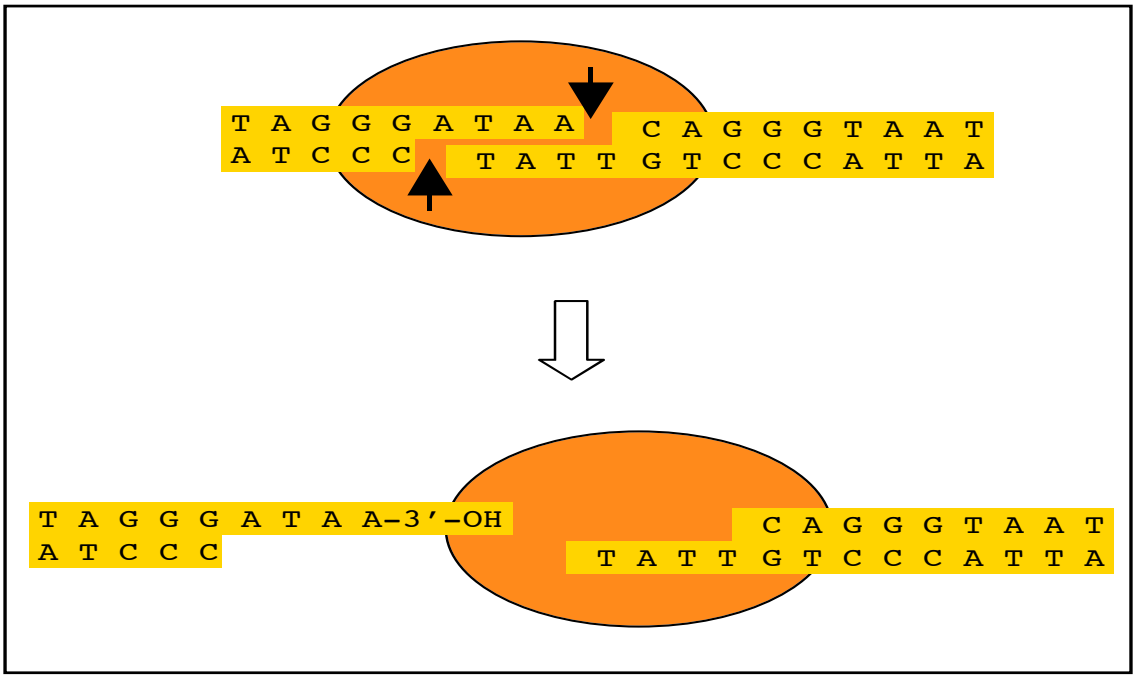


Fig. 4: Mechanism of cleavage of the *I-SceI* meganuclease.
The meganuclease acts in monomeric form, recognizes and cleaves the DNA in an asymmetrical fashion. It exhibits a low turnover, due to its strong association to the larger half-site.

Most of these proteins, like *I-SceI*, carry the motif in duplicate and are endonucleases. *I-SceI* has been purified as a monomeric globular protein of 235 amino acids (Monteilhet et al., 1990). Its endonuclease activity requires Mg^{2+} or Mn^{2+} to cleave DNA within its recognition sequence (TAGGGATAACAGGGTAAT) and leaves a four bp overhang presenting a 3'-hydroxyl terminus (Monteilhet et al., 1990). The enzyme displays a low turnover, probably because of its strong affinity for one of the products of the cleavage reaction (Perrin et al., 1993) (Fig. 4).

Other than REMI, that carries the intrinsic risk to fractionate the genome, extremely rare cutting meganucleases (*I-SceI*) could be employed for transgenesis, acting only on sites introduced into the insertion construct.

In mammalian cell culture, preliminary experiments have shown that co-transfection of plasmids bearing meganuclease recognition sites with expression vectors encoding the corresponding meganuclease, efficiently led to stably transfected cell lines with single copy integrations (A. Choulika, unpublished results). The 18 bp recognition site of *I-SceI* is expected to be found only once in $7x10^{10}$ bp of random sequence. Consequently, such a site has not been found in the medaka genome (10^9 bp). The *I-SceI* meganuclease does not cut the genomic DNA, but acts solely to digest the injected DNA.

Here, I report, in collaboration with the research group of J. -S. Joly, the efficient generation of stably transgenic medaka strains by co-injection of the *I-SceI* protein with reporter vectors flanked at both ends by the corresponding recognition sites. We show that co-injection leads to a strong enhancement of the promoter dependent expression already in G0 and an increased transgenesis frequency. Our data suggests that this is due to integration at the one-cell stage.

2.4.2 The Sleeping Beauty Transposon System

The *Sleeping Beauty* transposon is a member of the Tc1/mariner superfamily of DNA transposons (Plasterk et al., 1999). It was synthetically reconstructed from several inactive transposable elements in salmonid fish (Ivics et al., 1997), which had been transpositionally inactive due to the accumulation of mutations in the transposase coding region and, in most cases, additionally in the inverted repeats (vertical inactivation).

Twelve partial salmonid-type TcE sequences from eight species were aligned and conserved protein and DNA sequence motifs were identified that are supposed to correspond to functionally important domains. The inactivating mutations were eliminated based on the majority rule consensus sequence, thereby *quasi*-inverting 10 million years of divergence and accumulation of mutations, and reconstructing the original functional transposon, or a similar one.

Members of the Tc1/mariner superfamily of transposons are, unlike the P-element from *Drosophila melanogaster*, independent of specific host factors for transposition (Vos et al., 1996). This is supported by the fact that these elements are extraordinarily widespread in nature, ranging from single-cellular organisms to humans (Plasterk, 1996). Consistently, *SB* transposase has been shown to efficiently mediate transposition in cells from fish (carp), mouse, and human in culture (Ivics et al., 1997) and has been used as a genetic tool in the mouse *in vivo* (Dupuy et al., 2002; Dupuy et al., 2001; Fischer et al., 2001; Horie et al., 2001). It does not seem to interfere with endogenous elements in these species.

SB belongs to the IR/DR subgroup of Tc1-like elements (TcEs) that contain IRs of 210-250 bp flanking the central transposase coding region. Each IR contains two direct repeats (DRs), one at either end, which constitutes the cores of the binding sites for the transposase. The transposase binds to both DRs and adjacent sequences, but only the outer DRs are utilized for cleavage and excision of the transposon. The left and right IR are imperfect, with a match less than 80 % at the centre, but perfect in the DRs. The DNA binding domain of the transposase that provides specificity for salmonid-type IR/DRs spans the N-terminal half and includes a bipartite nuclear localization signal (NLS). The C-terminal half comprises a glycine-

riche sequence of unknown function and the DD (34) E domain that catalyses transposition, termed after the highly conserved amino acid residues Asp, Asp, Glu, in which the latter are separated by 34 amino acids (Fig. 5).

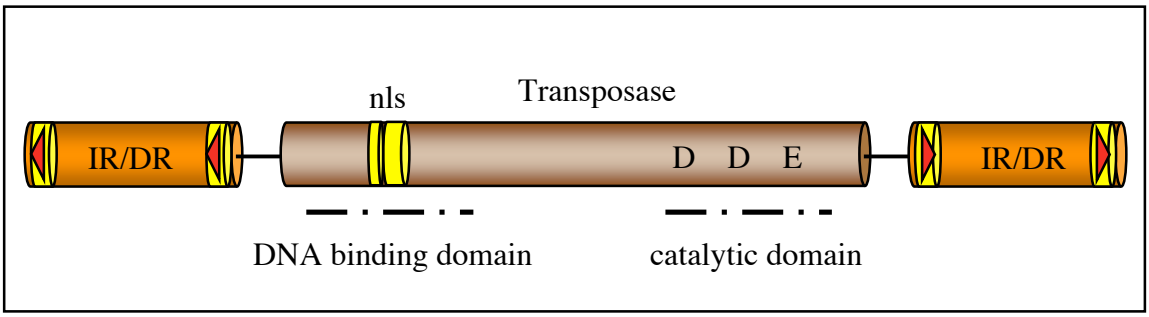


Fig. 5: Schematic drawing of an autonomous DNA transposon of a Tc1-like element.
The transposase (brown) mediates transposition. The transposase gene is flanked by inverted direct repeats (IR/DRs), each of them containing two direct repeats (red arrowheads).

Transposons of the Tc1/mariner superfamily integrate into TA target dinucleotides that become duplicated upon integration (Fig. 6). The presence of TcEs in many different species indicates that, in contrast to P-elements, they are independent of any host factors and therefore might be used for germ line transformation in many different species. This hypothesis is corroborated by the fact that recombinant Tc1 and mariner transposase are able to catalyze transposition *in vitro* without help of any other factor (Lampe et al., 1996; Vos et al., 1996). Furthermore, vectors based on *minos*, a TcE endogenous to *Drosophila hydei*, were successfully used for germ line transformation of the fly *Ceratis capitata*. (Loukeris et al., 1995) and the mariner element from *Drosophila mauritiana* was capable of undergoing transposition in the protozoan *Leishmania* (Gueiros-Filho and Beverley, 1997).

Raz and colleagues used Tc3 from *C. elegans* to stably introduce a reporter construct containing *GFP* into zebrafish embryos by co-injection of the Tc3 transposase RNA together with the reporter flanked by inverted repeats (Raz et al., 1998). In one line, they could show transposon-mediated integration, expression of the reporter construct and germ line transmission.

Pilot experiments performed in our lab indicated the *Sleeping Beauty* transposon system could be useful for medaka transgenesis (Henrich, 1999). Here I systematically investigated the potential of *SB* to perform transient and stable

transgenesis. Furthermore, the applicability for enhancer traps, repeated transposition and combination with other systematic tools used in developmental biology will be evaluated.

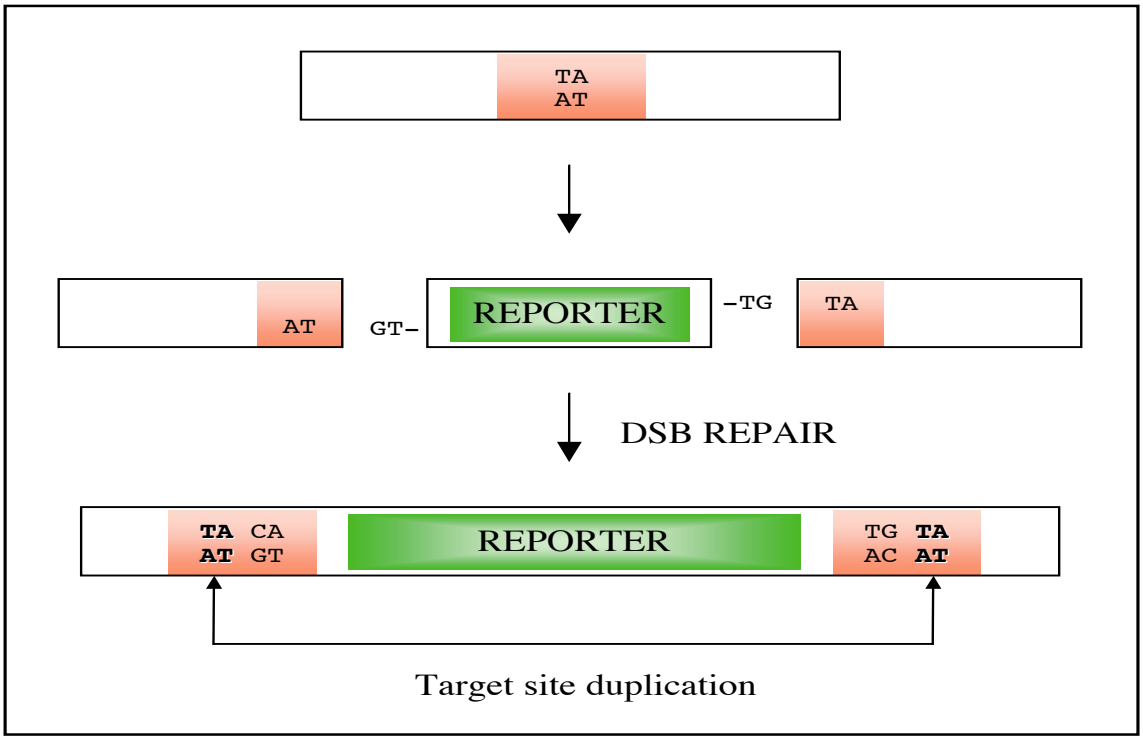


Fig. 6: Mechanism of transposition.
The TcE is excised by transposase mediated staggered double-strand breaks. Repair of DNA results in a target site duplication at the insertion site.

It will be shown that the *Sleeping Beauty* transposon system efficiently generates transgenic medaka lines with a high proportion of lines exhibiting novel, spatially and temporally restricted *GFP* expression patterns. In addition, transient, promoter dependent expression of the transgene is strongly enhanced using the *SB* system.

Aims of the Thesis

The primary goal of this thesis included the development of novel or enhanced transgenesis technologies to facilitate studies in developmental biology using the medaka fish, *Oryzias latipes*. The availability of fast and simple techniques to introduce foreign DNA into fish embryos allowing the expression of transgenes is a prerequisite for the modern developmental biologist to investigate the crucial processes of embryogenesis.

Both, transient expression of episomal transgenes as well as stable integration into the genome and expression in subsequent generations are equally important. Moreover, generation of transgenic fish exhibiting novel tissue- or cell-specific patterns of reporter transgene expression (enhancer trap) provides useful tools for the analysis of these tissues or cell populations and the transcriptional regulation of the trapped regulatory units during embryonic development under various conditions.

So far, microinjection of DNA into fish embryos at the one cell stage provides the fastest and easiest approach. Unfortunately, the quality of transient promoter-dependent expression of the transgene and the efficiency to generate stable transgenics are very low (Collas and Alestrom, 1998; Culp et al., 1991; Lin et al., 1994a; Stuart et al., 1988; Stuart et al., 1990; Tanaka and Kinoshita, 2001; Westerfield et al., 1992). In addition, the frequency of enhancer traps in fish is even lower which prevented the routinely use of this approach in fish (Bayer and Campos-Ortega, 1992).

I applied two different technologies to overcome these previous limitations.

First, co-injection of DNA and the meganuclease *I-SceI* (Jacquier and Dujon, 1985) enhanced transgene expression in the G0 generation by integration into the host genome as early as the one cell stage. Thereby, allowing reliable promoter studies in G0 and at the same time increasing the transgenesis frequency to ~30 %.

Second, by application of the artificially reconstructed transposon system *Sleeping Beauty* (*SB*) (Ivics et al., 1997) transient transgene expression in G0 was improved by equal segregation of episomal DNA. Insertion of the transgene into the host genome at later stages of development led to a transgenesis frequency of ~32 %.

Thus, these approaches represent the most potent tools to perform transgenesis by microinjection in fish up to date. Moreover, the *SB* approach revealed its potential to perform enhancer trap experiments, ~12 % of the transgenics featured typical characteristics of enhancer-trap lines, i.e. spatially and/or temporally restricted transgene expression due to regulation imposed by sequences adjacent to the insertion site.

Results

3. Application of the *I-SceI* Meganuclease in Medaka

3.1 Co-injection of Reporter Gene and *I-SceI* Meganuclease Leads to Uniform Promoter Dependent *GFP* Expression in G0

In collaboration with Violette Thermes, I established the meganuclease protocol for transgenesis in medaka. To assay the potential of meganuclease-mediated transgenesis in medaka, we used two types of reporter constructs in which *GFP* is driven by two different promoters: the moderate cytoskeletal-actin promoter of *X. borealis* (pCSKAGFPS-I) for ubiquitous expression in the entire embryo (Condie et al., 1990); and the muscle specific α-actin promoter of zebrafish (pαact-GFPI2) for tissue specific expression (Higashijima et al., 1997). These two reporter constructs were flanked at both ends by *I-SceI* meganuclease recognition sites.

Circular and linear plasmid-DNA was tested for transient expression in G0 by injection into one-cell stage embryos (stage 2a, (Iwamatsu, 1994)) with or without *I-SceI* meganuclease. When co-injecting *I-SceI* and plasmid DNA, a DNA concentration of 10 ng/μl leads to efficient expression without significantly increasing the mortality of the injected embryos. No further deleterious effects on embryo survival were observed. *GFP* expression was then examined at a number of distinct developmental stages using fluorescence binocular microscopy. *GFP* expression levels were similar with both vectors; only those obtained with the muscle specific zebrafish α-actin promoter (pαact-GFPI2) will be presented in detail (Thermes* et al., 2002).

Embryos were scored after three days of development (stage 31) when the muscular α-actin *GFP* expression was easily detectable in striated cells. The embryos were grouped according to the intensity of fluorescence in order to determine quantitatively the level and distribution of transgene expression in each experiment

(Fig. 7) by my collaborator Violette Thermes. Upon injection of circular or linear plasmid without *I-SceI* enzyme, about fifty percent of the surviving embryos showed no muscular fluorescence and were classified as negative (Tab. 2). In the other embryos, the amount of *GFP* positive cells ranged from a few cells (classified as weak, 24 % with circular plasmid and 29 % with linear plasmid), to an almost ubiquitous labelling of muscle cells (strong, 5 % with circular plasmid and 3 % with linear plasmid). When expression was detected in a large domain of the muscles, it was qualified as moderate (21 % with circular plasmid and 19 % with linear plasmid). Expression in individual muscle cells was always strong enough to be easily detected, and no ectopic expression was observed (Tab. 2).

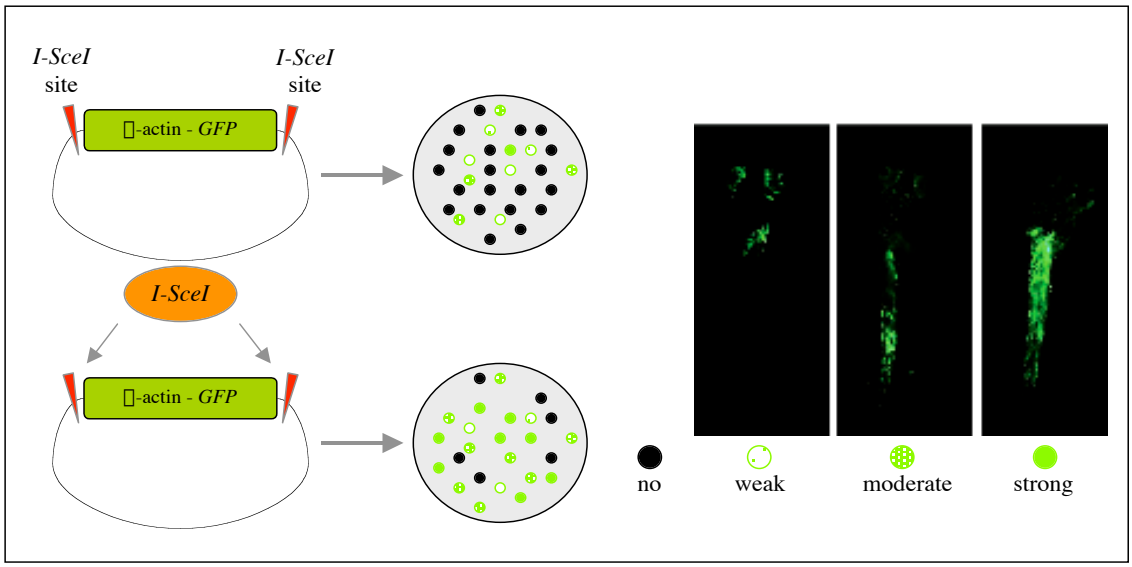


Fig. 7: Transient expression of *GFP* driven by the zebrafish -actin promoter in medaka embryos (see (Thermes* et al., 2002)).
Co-injection of plasmid DNA containing the transgene flanked by *I-SceI* recognition sites and meganuclease leads to an increased number of embryos with a promoter dependent muscular expression pattern.
Left panel: Schematic drawing illustrating transient *GFP* expression levels in G0 upon injection of naked DNA or co-injection of *I-SceI*, respectively. Injected embryos are represented as coloured spots indicating the level of fluorescence, as shown in the right panel.
Right panel: Embryos were grouped according to the level of *GFP* expression observed in injected embryos at stage 31 (3 days post fertilization). The different expression levels are illustrated with a colour code.

In contrast, when circular p act-GFPI2 was co-injected with *I-SceI* meganuclease, the *GFP* muscular expression was dramatically improved. About 76 % of the embryos exhibited a moderate or strong expression in the trunk musculature, as

compared to 26 % when injection was performed without enzyme. Specificity of this effect was further validated by application of mutated or deleted meganuclease sites (Tab. 2). Thus, transient transgene expression in G0 fish is readily and efficiently improved applying the meganuclease protocol (Thermes* et al., 2002).

		I-SceI	Negative (%)	Weak (%)	Moderate (%)	Strong (%)
pβact-GFPI2	Linear	-	49	29	19	3
	Circular	-	50	24	21	5
	Circular	+	16	8	37	39
pβact-GFPMI	Circular	+	72	14	8	6
	Linear	+	48	31	19	2
pβact-GFPDI	Circular	+	69	14	12	5
	Linear	+	50	35	15	0

Tab. 2: Distribution of G0 *GFP* expression (see (Thermes* et al., 2002)).
Frequencies of different muscular *GFP* expression levels observed in embryos injected with pβact-GFPI2 alone or upon co-injection with *I-SceI*. Surviving embryos were scored after 3 days of development (stage 31). Embryos were grouped according to the *GFP* muscular expression pattern. Circular and linear pβact-GFPMI (mutated *I-SceI* site) and DI (deleted *I-SceI* site) were injected with meganuclease to test a hypothetical NLS activity of *I-SceI*.

Using pβact-GFPI2, the transient *GFP* muscular expression persisted in adult fish. An important point was then to investigate if the transgene was transmitted to the next generation more frequently upon co-injection of the meganuclease. In order to analyze the transmission of the transgene to the progeny and, in particular, whether the improved rate of *GFP* expressing G0 fish also resulted in a higher germ line transmission rate, we generated families of transgenic fish.

3.2 Generation of Germ Line Transmitting Fish by *I-SceI* Meganuclease Co-injection

The two plasmids bearing the expression cassettes flanked by the *I-SceI* sites were co-injected with and without *I-SceI* meganuclease (see above). All injected fish exhibiting no *GFP* expression turned out to be negative for germ line transmission of the transgene (data not shown). Injected *GFP* expressing G0 fish were selected as putative founder fish, grown to sexual maturity, mated to wild-type partners and tested for germ line transmission. Fish transmitting the functional transgene to the progeny, as judged by *GFP* expression, were then selected as real founders. The *GFP* fluorescence in their three-day-old progeny was then scored to estimate the germ line transmission rate.

In control experiments, entire linear and circular p□act-GFPI2 plasmids were injected without the meganuclease. Of injected *GFP* expressing G0 fish only 5.9 % and 15.6 % respectively, were transgenics (Tab. 3). These transgenic founder fish also showed a highly mosaic germ line. For most p□act-GFPI2 G0 transgenic fish derived from linear or circular plasmid injection only few F1 offspring did express *GFP* (Tab. 3). We observed an average germ line transmission rate of 15.1 % for the linear form, and 17.6 % for the circular form, and a high standard deviation of 17.7 % and 22.2 % (Tab. 3), respectively. This low rate is indicative of a late integration event after several cell cleavages, taking place in only a fraction of the blastomeres contributing to the germ line (Thermes* et al., 2002).

		I-SceI	Rate of G0 transgenics (%)	Average transmission rate	Standard deviation (□)
p□act-GFPI2	Linear	-	2/35 (5.9)	15.1 %	17.7 %
	Circular	-	5/32 (15.6)	17.6 %	22.2 %
	Circular	+	11/36 (30.5)	48.4 %	9.1 %
pCSKAGFPS-I	Circular	+	20/65 (30.7)	49 %	10.3 %

Tab. 3: Transgenesis frequency and germ line transmission rates.
Rates of transgenic fish and transgene transmission to F1 progeny after injection into one-cell stage medaka embryos. Circular plasmids and *I-SceI* were co-injected, and resulting fish were tested for their ability to transmit the transgene to the F1 offspring.

When *I-SceI* meganuclease was co-injected with circular plasmids (pβact-GFPI2 and pCSKAGFPS-I), 30.5 % and 30.7 % respectively, of *GFP* positive G0 fish had F1 offspring expressing *GFP*. Moreover, 48.4 % and 49 %, respectively, of the F1 fish expressed the transgene with a low standard deviation of 9.1 % and 10.3 %, respectively.

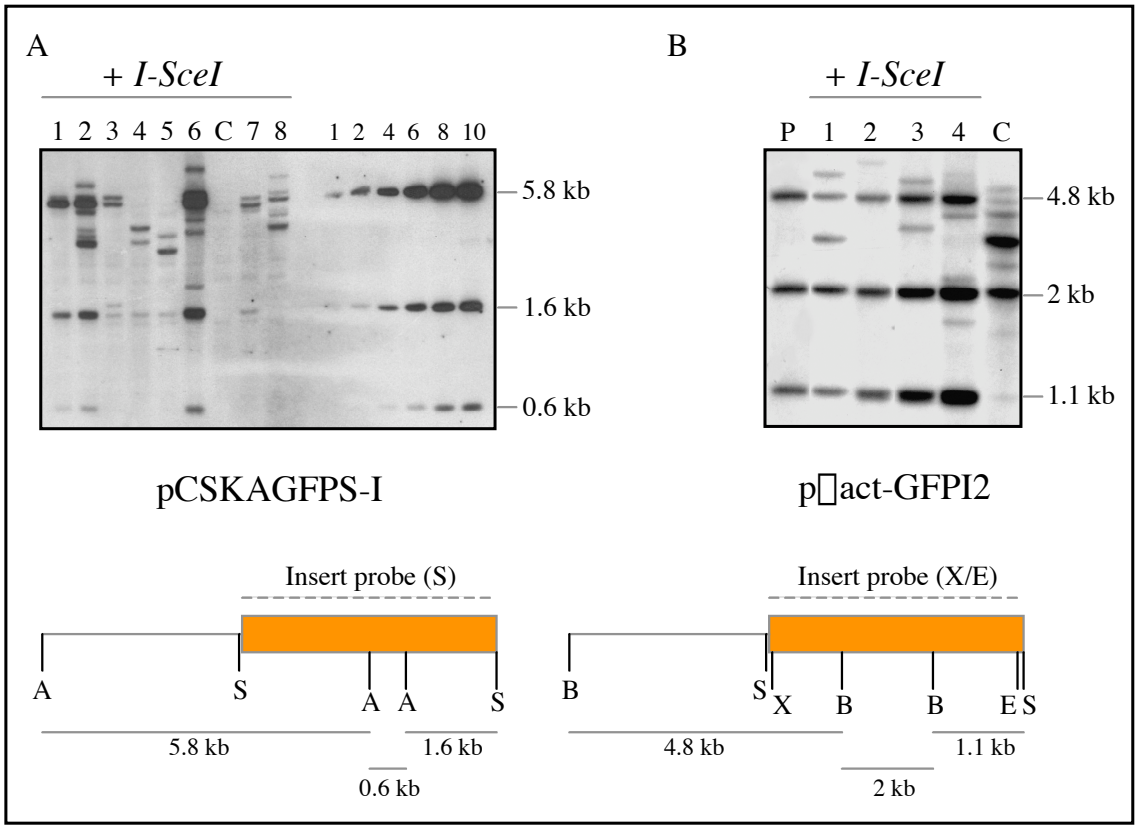


Fig. 8: Southern blot analysis of transgenic lines.
Southern blot analysis of genomic DNAs (10μg/lane) isolated from adult medaka F1 fish resulting from outcrosses of G0 transgenic fish obtained by co-injection of pCSKAGFPS-I (A) or pβact-GFPI2 (B) with *I-SceI* meganuclease. Genomic DNA was digested with *AflIII* (pCSKAGFPS-I) or *BamHI* (pβact-GFPI2) (Thermes* et al., 2002), and hybridized with corresponding insert probes, resulting from *I-SceI* and *XhoI*/*EcoRV* digestions, respectively (see schematic representation). The copy number of integrated concatemers was estimated in pCSKAGFPS-I transgenics, using a standard array of plasmid DNA. Lanes (1-3 and 6-7 in A; 1-4 in B) show integrations of less than eight copies of the whole linearized plasmid in tandem as demonstrated by the presence of the expected fragment sizes (5.8kb, 1.6kb, 0.6kb for pCSKAGFPS-I or 1.1kb, 2kb, 4.8kb for pβact-GFPI2). Restriction patterns of ten out of twelve investigated lines (lanes 1, 3, 4, 5 and 7 in the left panel; lanes 1-4 in the right panel) suggest low copy numbers insertions at a single genomic locus.
Left panel, 1-10: plasmid standards. Quantities indicated on top of each lane correspond to one to ten copies of plasmid integrated in the genome; C: control lane with uninjected wild type fish. Right panel, lane P, pβact-GFPI2 digested with *BamHI*; C: control lane with pβact-GFPI2 embryos injected without *I-SceI*. Abbreviations: A, *AflIII*; B, *BamHI*; E, *EcoRI*; S, *I-SceI*; X, *XhoI*.

The rate of co-injected embryos that turned out to be germ line transmitters is higher than in the control experiment, and these transgenic fish transmitted the transgene to about half of their progeny compared to only 15 percent in most control transgenics. The improved expression in G0 obtained by co-injection with meganuclease was thus positively correlated with both an enhanced transgenesis frequency and an increased germ line transmission rate. Germ line transmission rates close to 50 % are due to a single integration of the transgene at the one-cell stage, leading to non-mosaic heterozygous transgenic fish.

To determine the nature of DNA integrations (the number of insertion loci and the length of concatemers), we performed Southern blot analyses on genomic DNA from independent F1 transgenics of fish co-injected with *I-SceI* (Fig. 8). Genomic DNA was first digested with restriction enzymes that cut the insert and the plasmid (*AflIII* for pCSKAGFPS-I; *BamHI* for pβact-GFPI2). Blots were hybridized with insert probes obtained by digestion with either *I-SceI* for pCSKAGFPS-I (Fig. 8 left panel, schematic diagram) or *XhoI/EcoRV* for pβact-GFPI2 (Fig. 8 right panel, schematic diagram).

Southern blot analysis of nine independent transgenic lines revealed insertions of the entire plasmid in tandem arrays (Fig. 8 left panel, lanes 1, 2, 3, 6 and 7; Fig. 8 right panel lanes 1, 2, 3 and 4), as demonstrated by the presence of the expected fragments (5.8 kb, 1.6 kb and 0.6 kb for pCSKAGFPS-I; and 1 kb, 2 kb and 4.8 kb for pβact-GFPI2) (Fig. 8). This pattern also suggested that *I-SceI* sites were still present in the genome of all these transgenics, which was thereafter confirmed by Southern blots on DNA digested by *I-SceI* (data not shown). The copy number of integrated concatemers was estimated in pCSKAGFPS-I transgenics, using a standard array of plasmid DNA. Copy numbers range from only one or two copies of the injected constructs (Fig. 8 left panel; lanes 3, 4, 5, 7 and 8) to a maximum of eight copies in all of the lines analyzed, significantly lower than reported for standard transgenesis in fish, where up to 2000 copies were reported to integrate in tandem clusters (Hackett, 1993; Iyengar et al., 1996).

In other cases (Fig. 8 left panel, lane 4, 5), the absence of the internal fragment found in concatemers (5.8 kb fragment lane 4 and 5) indicates a single copy insertion. The restriction pattern is consistent with insertions at more than one locus as more than two junction fragments were detected. Segregation analysis however suggested

integration of the functional copy at a single site. For *I-SceI* mediated integration, two junction fragments are expected for both vector ends as found in lanes 1, 3, 4, 5, 7 (Fig. 8 left panel) and lanes 1-4 (Fig. 8 right panel), indicating single- or double-sided *I-SceI* mediated insertion of the entire plasmid in tandem at a single locus. Hybridisation bands in lanes 2, 6 and 8 indicate several insertion loci, as demonstrated by the presence of several junction fragments. Taken together, in the majority of the lines analyzed, the functional reporter integrated as single copy element or low copy number tandem repeat into mostly single sites within the genome, in striking contrast to the high copy numbers of tandem repeats reported to integrate in standard transgenesis approaches (Hackett, 1993; Iyengar et al., 1996).

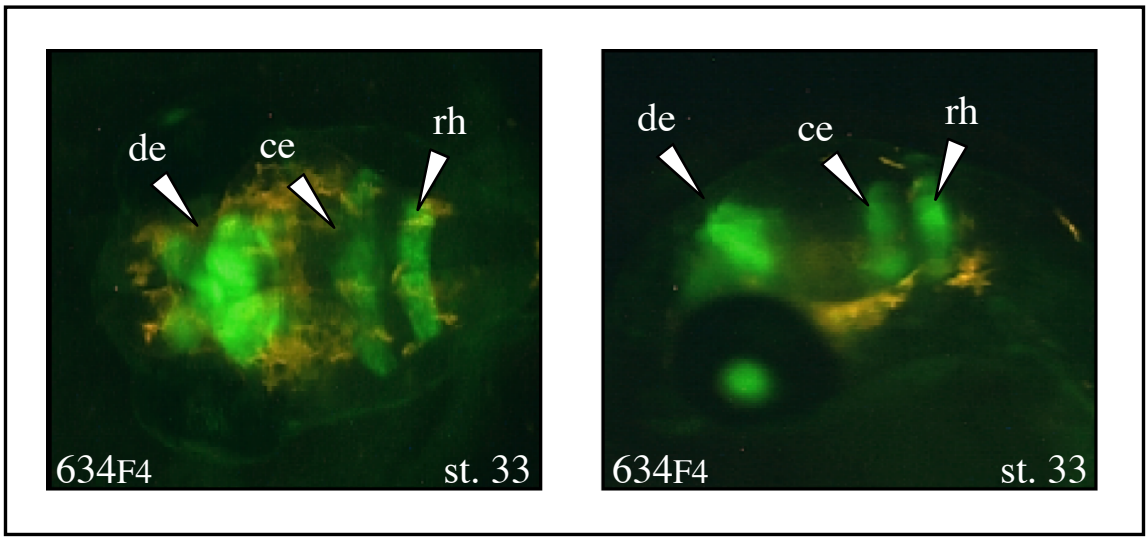


Fig. 9:Pattern in F4 medaka embryos upon co-injection of pCSKAGFPS-I and meganuclease.
Dorsal (A) and lateral (B) view of a transgenic embryo (line 634) at stage 33 (organogenesis), anterior is to the left. *GFP* expression is detected in the diencephalons (de), cerebellum (ce) and rhombomere 3 (rh).

In addition to the promoter dependent expression obtained in the transgenic lines, interestingly, one of twenty independent transgenic lines obtained with pCSKAGFPS-I, exhibited an intriguingly specific and stable expression pattern (Fig. 9) in the F1 generation. Transgenic line 634 shows expression of *GFP* in regions of the fore- and hindbrain. At stage 21 (brain and otic vesicle formation) *GFP* expression is induced in this line. During further development of these brain structures, *GFP*

expression is confined to precise morphological areas. Anteriorly, the lenses and the dorsal to medial diencephalon are *GFP* positive including the epiphysis and the *habenulae*. In the anterior part of the hindbrain, *GFP* can be detected in the most ventral part of the cerebellum, which is referred to as rhombomere 1. Rhombomere 2 is negative while *GFP* is expressed again in the entire rhombomere 3.

This indicates that this reporter construct was sensitive to position effects and may be useful for enhancer or gene trapping strategies.

3.3 Improvement in G0 Transgene Expression is Not Linked to a Nuclear Targeting Activity of the Meganuclease

Three explanations for the increased efficiency of *I-SceI* transgenesis can be contemplated. First, cleavage of the transgene by the enzyme promotes rapid integration, probably by counteracting the endogenous ligase activity fusing the transgene into multimers and thereby sustaining a high number of transgene copies as short linear fragments. Second, *I-SceI* cleaves the host genome and thereby promotes integration of the transgene through endogenous non-homologous end joining. Third, *I-SceI* binds to the transgene and promotes nuclear localization of the transgene. Although it is not known if *I-SceI* has an implicit nuclear localization activity, it is known that *I-SceI* binds its recognition site tightly, both before and after cleavage (Jacquier and Dujon, 1985).

Since *I-SceI* cleavage is expected randomly only once in $7x10^{10}$ bp, the second possibility is very unlikely. Furthermore, Southern blot analyses of the transgenic lines excluded this possibility, as it provided no evidence for a unique, reappearing integration site. In addition, in mammalian cell cultures, no evidence for preferential integration sites was found when flanking regions of several tenths of insertions were analyzed (A. Choulika, unpublished data).

To assess whether the increased efficiency of transgenesis is due to a nuclear localization activity of the *I-SceI* meganuclease, linear or circular control constructs bearing deleted (p□act-GFPDI) or mutated (p□act-GFPMI) recognition sites (both are

bound, but not cleaved by the meganuclease were co-injected with or without the meganuclease (Tab. 2) (Colleaux et al., 1988)(Thermes* et al., 2002).

Advantage was taken of the tight correlation of uniform promoter dependent expression in G0 and the germ line transmission rate to quantitatively address this question. Following injection, embryos were grouped according to the criteria described above (Fig. 7). The distribution of the transient *GFP* expression in G0 embryos injected with circular and linear p□act-GFPMI, containing a modified meganuclease site that allows enzyme binding but not cleavage, or with circular and linear p□act-GFPDI in which the meganuclease site has been deleted was determined. *GFP* expression data using those two constructs were highly reminiscent of those obtained in control experiments involving the injection of the circular and linear p□act-GFPI2 in the absence of the enzyme (Tab. 2). Thus, in none of the cases enhanced transgenesis rates were observed, indicating that a putative NLS located in the enzyme is not sufficient to mediate efficient integration. It is still possible however that it contributes to the translocation of the digested DNA to the nucleus.

Taken together the primary reason for the increased efficiency of transgenesis is the cleavage of the transgene by the *I-SceI* meganuclease that mediates efficient integration.

4. The *SB* Transposon System in *Oryzias Latipes*

4.1 Application of the *SB* System Results in Increased Numbers of G0 Embryos Uniformly Expressing *GFP*

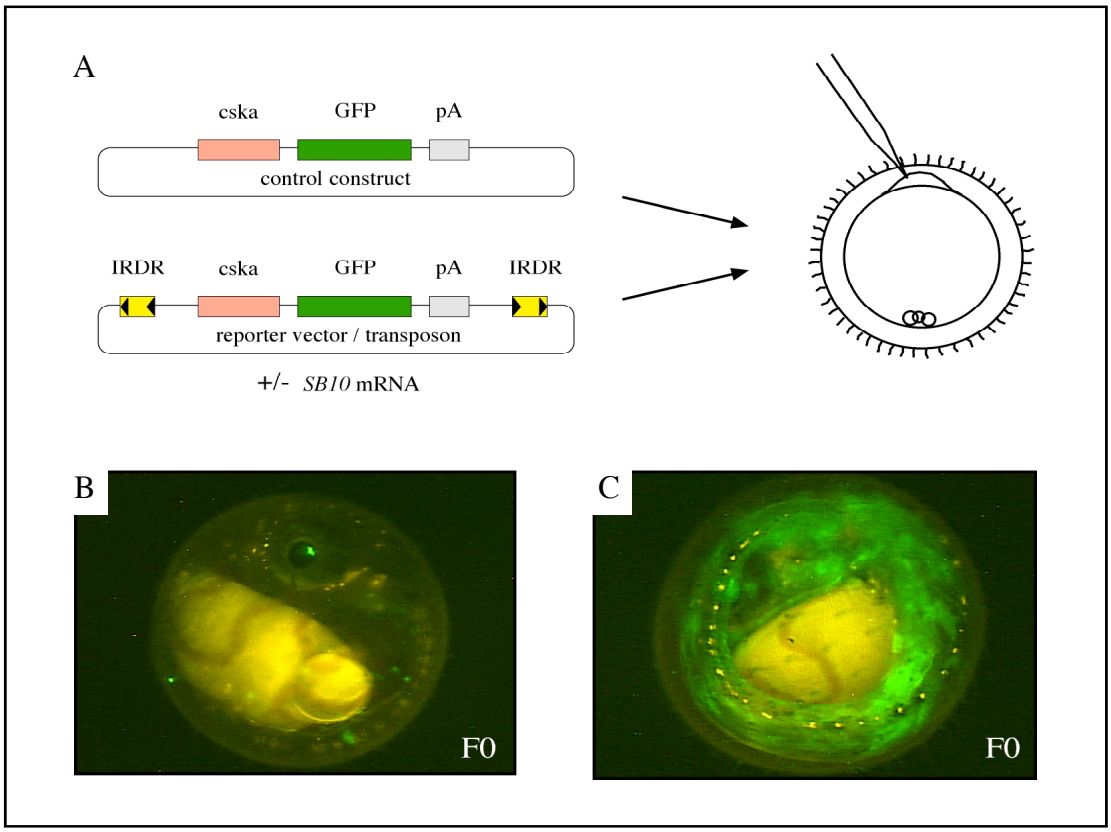


Fig. 10: Generation of founder fish by microinjection of circular DNA into one-cell stage medaka embryos.
A, Circular plasmid DNA containing an expression cassette with the cytoskeletal-actin promoter (pink box) of *X. borealis* driving *GFP* (green box), and a SV40 poly A signal (grey box) was injected into one-cell stage medaka embryos as control construct. To test the *SB* transposon system, this expression cassette was flanked by *SB* recognition sequences (inverted direct repeats, IRDR, transposon, yellow box) and injected with or without *SB10* transposase mRNA. B, Medaka embryo 3 days after injection showing mosaic expression of *GFP*. C, Medaka embryo 3 days after injection showing promoter dependent ubiquitous expression of *GFP*.

To efficiently use an animal model system for transposon-mediated transgenesis, it needs to be free of (cryptic) endogenous elements that could be mobilized by the newly introduced transposase. The absence of such sequences in the medaka genome was verified functionally by over-expression of the *SB* transposase in wild type medaka upon injection of *SB10* mRNA into embryos. Even at the highest concentration, no effects on embryonic survival or development were observed (data not shown). In addition, low stringency Southern blot analysis using the inverted terminal repeats (IR/DRs) that serve as recognition elements for the *SB* transposase as a probe did not yield a signal, indicating that the medaka genome is free of endogenous elements capable of transposition by the *SB* transposase.

To address the potential of the *SB* transposon and the *SB10* transposase for transgenesis, early one-cell stage embryos were co-injected with a reporter vector (transposon) and mRNA encoding the *SB* transposase (Fig. 10).

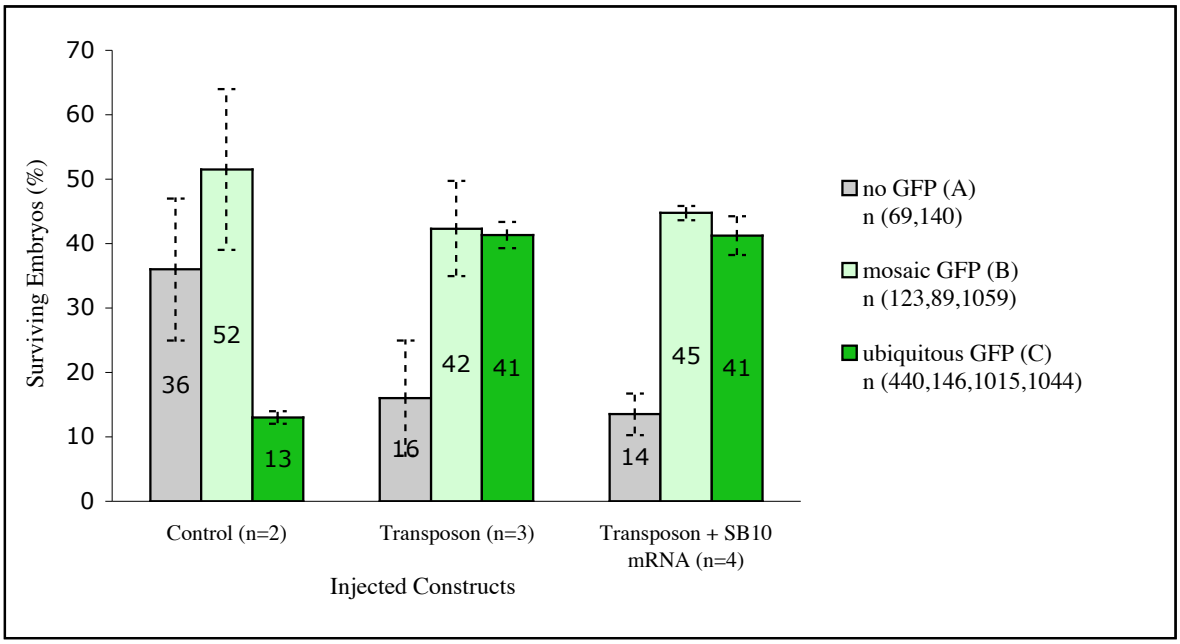


Fig. 11: The *SB* transposon system enhances transient *GFP* expression.
GFP expression was monitored in G0 upon injection of control construct (Control) or reporter vector (transposon) and upon co-injection of transposon with *SB10* mRNA (transposon + *SB10* mRNA). After 3 days of development, embryos were grouped in expression groups according to the level of transgene expression (A-C). Percentages of surviving embryos exhibiting *GFP* expression are indicated within the bars. n, number of experiments (amount of embryos per experiment is indicated within brackets).

The reporter contained an expression cassette with the cytoskeletal-actin promoter/enhancer (cska) of *X. borealis* (Condie et al., 1990) to drive moderate, ubiquitous expression of *GFP*. This expression cassette was flanked by the terminal inverted repeats of *SB* containing the binding sites of the transposase. Injections were performed with or without *SB10* mRNA. A similar vector, lacking *SB* recognition sequences (control construct), was used for control injections. In contrast to mosaic expression due to transient transcription of non-integrated plasmids (Winkler et al., 1991), an early integration event leads to the transmission of the transgene to all of the daughter cells, and thus results in *GFP* expression in large clones of cells (Fig. 2 and 10). To distinguish between these two possibilities, embryos were screened for *GFP* expression at day three of development.

The injected embryos were scored and grouped according to the degree of mosaicism in *GFP* expression: (A) no *GFP* expression, (B) mosaic expression only and (C) ubiquitous *GFP* expression (Fig. 10B, C, 11). Upon injection of control construct in the absence of *SB* recognition sequences and transposase, only 13 % of surviving embryos expressed *GFP* uniformly in the entire body. Almost 40 % did not show any fluorescence and about one half expressed *GFP* in mosaic cell clones of variable size, in accordance with results previously reported for DNA microinjection (Chou et al., 2001). Conversely, when *SB* IR/DRs were included, uniform, promoter dependent expression was the predominant result. This was observed in 41 % of the surviving embryos co-injected with the transposon and *SB10* mRNA or injected with the transposon only (Fig. 11). Thus, the presence of IR/DRs strongly enhanced promoter dependent transgene expression in G0, indicating that *SB* IR/DRs, similar to the ITRs of adeno-associated virus (AAV) (Chou et al., 2001), significantly enhance transient transgene expression.

4.2 Establishment of Stable Transgenic Lines Using the *SB* System

To further examine to whether *GFP* expression in G0 positively correlates with germ line transmission of the transgene, I analysed the F1 offspring of all three expression groups for inheritance of the transgene. G0 fish were mated to wild type and the F1 offspring was screened for *GFP* expression (Fig. 12).

This analysis showed that G0 expression was a reliable indicator for efficient transgene transmission. Application of the *SB* transposon alone or in combination with *SB10* mRNA enhanced total transgenesis frequencies to 31 % (174/560) compared to 4 % (3/70) resulting from control construct injections. Strongly expressing G0 fish injected with the *SB* transposon (group C), transmitted *GFP* expression in 39 % (118/305) or 45 % (33/74) with or without transposase, respectively. On the other hand, of the fish injected with the control construct and showing uniform G0 expression (13 % of injected survivors, expression group C), only 14 % (2/14) stably transmitted *GFP* to the next generation. From the mosaic G0 fish (expression group B), 14 % (4/28) or 24 % (12/50) founded stable transgenics with transposon only or in combination with *SB10* mRNA, respectively, while only 3 % (1/34) transmitted a functional transgene to the F1 generation when the control construct had been injected. *GFP* negative G0 fish (expression group A) transmitted a *GFP* expressing transgene to the next generation in 9 % (7/77) of the analysed fish only upon injection of the complete *SB* system.

These numbers indicate a positive correlation between the expression of the transgene in G0 and the frequency of stable transgenics. Embryos that exhibit a uniform *GFP* expression in G0 result in the highest yield of transgenic fish. This facilitates an easy selection of putative founder fish for medium- to large-scale approaches.

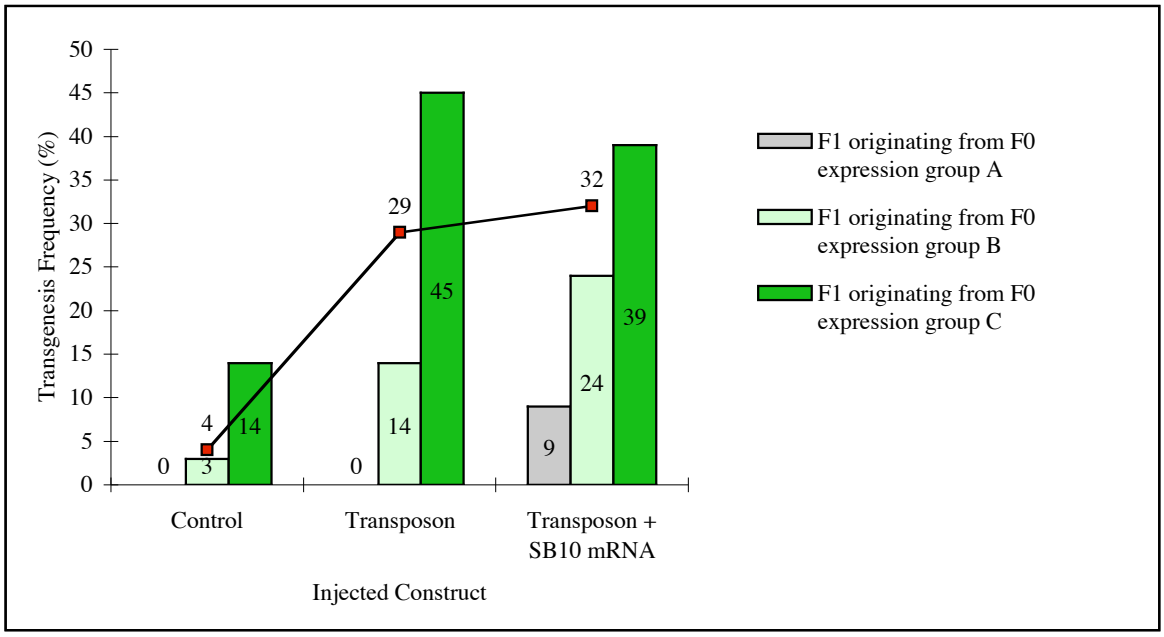


Fig. 12: Transgenesis frequencies are enhanced upon application of the *SB* transposon system.
Bars indicate transgenesis frequencies of each G0 expression group (A-C). Line indicates total transgenesis frequency, including screened embryos of all F0 expression groups. Percentages of transgenic F1 embryos per G0 expression group are indicated within the bars. Percentages of transgenic F1 embryos per injected construct (irrespective of transient *GFP* expression) are indicated above the line.

4.3 Genomic Integration of Single or Multiple Copies

To investigate the molecular nature of the integrations, I performed Southern blot analysis on *GFP* positive F1 fish using *GFP* coding sequence as a probe. Five lines are shown, including three transgenics with novel expression patterns (Fig. 13A; lanes 2, 4-5), and two lines expressing *GFP* in a promoter dependent manner (Fig. 13A; lane 1, 3).

The copy numbers of the integrated transgenes were determined to range from a single copy (lane 7) to more than 10 copies (lane 10). The transgenic line SV (lane 3) revealed a banding pattern consistent with a single copy insertion mediated by *SB10*. In order to verify the proposed mechanism of *SB10*-mediated transposition in medaka, the SV insertion was analysed in more detail by cloning and sequencing flanking genomic regions. Comparison to the wild type locus revealed TA target site duplications

flanking the inserted *SB* transposon, a molecular hallmark of transposition of Tc1-like elements (Henrich, 1999; Ivics et al., 1997; Plasterk et al., 1999) (Fig. 13B).

Most of the other lines revealed plasmid tandem head-to-tail insertions (Iyengar et al., 1996) either at a single locus or at multiple independent sites (Fig. 13A and data not shown). This is likely due to the strong DNA ligation and replication activity in early fish embryos (Hackett, 1993). I cannot exclude the possibility that transgene concatemers can be inserted into the genome by the proposed mechanism (Tab. 4).

In contrast to reports that a concatemeric array of transgenes induce silencing of the transgenic locus (Garrick et al., 1998) we find stable expression of the transgenes in tandem arrays over so far up to 7 generations. This hints at a transgene activity-stabilizing function of *SB* IR/DR sequences. Recently, similar results have been reported by Noma and co-workers who identified inverted repeats, shielding euchromatic regions of the mating-type locus of the fission yeast, thereby acting as barriers for heterochromatin spreading (Noma et al., 2001).

I could show that multiple insertions can be segregated by selective screening in subsequent generations. Already in siblings of the F2 generation the number of independent insertions in F1 (at least 4; Fig. 13C lane 1) of line DE was reduced (lane 2-3). In addition, I observed a reduction of ubiquitous *GFP* expression in some specific transgenic lines upon selective screening over several generations, likely due the decrease of multiple independent insertions.

The degree of mosaicism of the germ line is indicative of the time point of transgene integration. A single integration event at the one-cell stage results in a non-mosaic heterozygous fish that transmits the transgene to 50 % of its offspring, while greater fractions indicate multiple independent insertions. I analyzed the frequency of germ line transmission from identified G0 founders to their F1 progeny. Between 8 % and 60 % of the offspring showed *GFP* expression indicating single or multiple insertion events between the 1-8 cell stages (Tab. 4). The transgenes were subsequently inherited in a Mendelian fashion over many generations without any alteration of expression level or pattern.

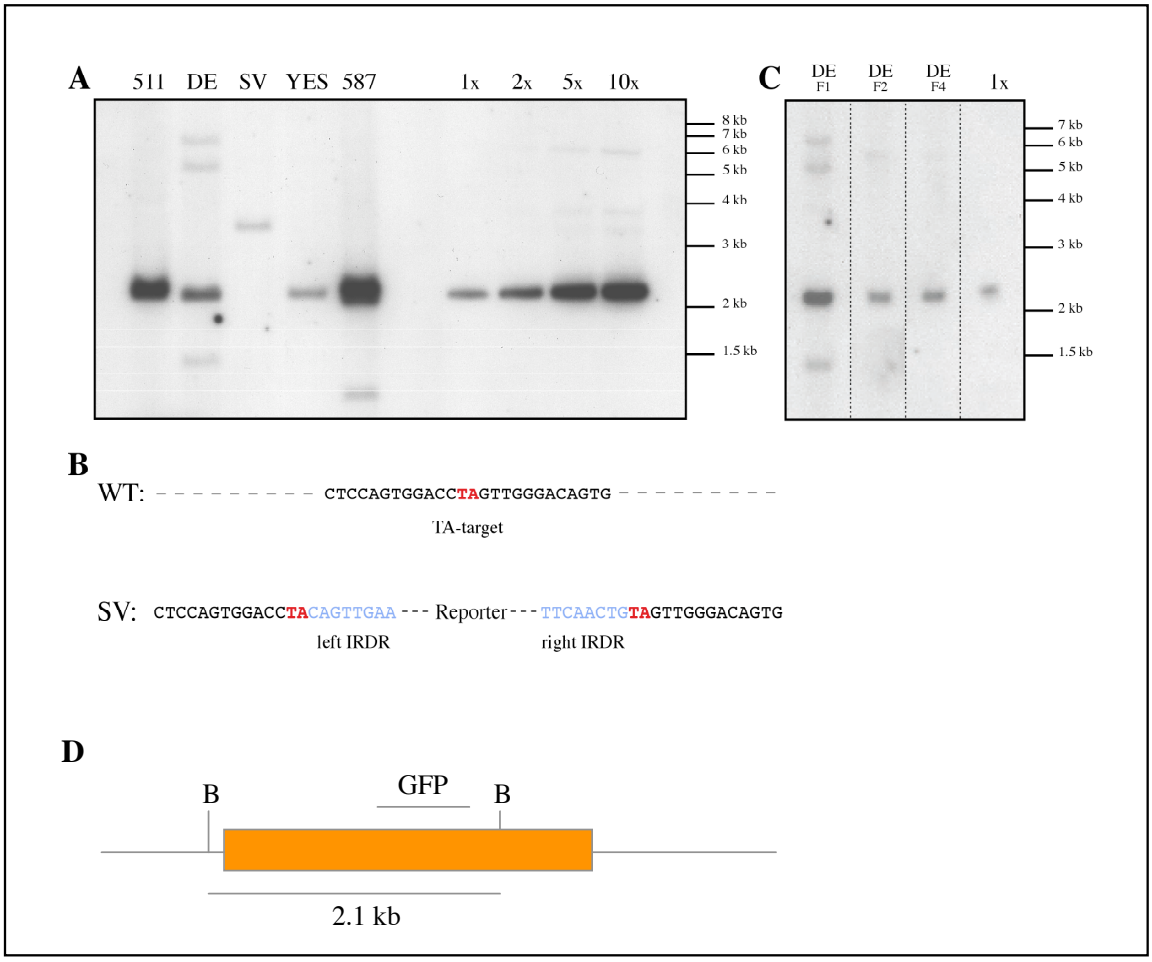


Fig. 13: Southern analysis of transgenic lines and genomic locus of *SB10*-mediated insertion in transgenic line SV.

A, Southern blot on F1 offspring of five independent transgenic lines. Genomic and plasmid DNA was digested with *BamHI* and separated by gel electrophoresis (0.8 %). Lines 511 and 587 (lanes 1, 5) show insertions of multiple copies (more than 10) of the entire plasmid in tandem array. Line 587 (lane 5, see Fig. 14) reveals an independent integration of a *GFP* fragment in addition. Similarly, line YES (lane 4; see Fig. 14) resembles whole plasmid integration with reduced copy number (1-2). Transgenic line DE (lane 2; see Fig. 14) shows integration of entire plasmid in arrays and at least two independent insertions at different loci. Line SV (lane 3) shows a banding pattern that is in accordance with a *SB10*-mediated insertion of a single copy of the expression cassette. Copy number of insertions was determined by comparison to linearized plasmid, ranging from 1 to 10 copies (lanes 7-10). B, Wild type locus (WT) of SV insertion. Target site of *SB10* transposase (TA dinucleotide) is marked in red. *SB10*-mediated insertion of a single copy of the expression cassette flanked by IR/DRs (light blue) of transgenic line SV (SV) leads to the predicted duplication of the TA target site (red). C, Segregation of multiple independent insertions was analyzed with transgenic line DE. At least 4 independent insertions in F1 (lane 1) were reduced in generations F2 (lane 2) and F4 (lane 3) while specific *GFP* expression was retained. Copy number was determined as in A (lane 4); *GFP* coding sequence was used as probe. D, Schematic drawing of the injected vector. *GFP* coding sequence was used as a hybridization probe.

4.4 High Frequency Generation of Spatially and Temporally Restricted Expression Patterns in F1 Progeny

We screened 560 fish that were injected with the *SB* transposon only (128) or co-injected with the *SB10* transposase (432) and established 174 independent transgenic fish lines expressing *GFP*. F1 offspring of the majority of the founders exhibited ubiquitous *GFP* expression according to promoter activity. Strikingly, however, 12 % (21/174) of the transgenics featured typical characteristics of 'enhancer-trap' lines, i.e. spatially and/or temporally restricted transgene expression due to regulation imposed by sequences adjacent to the insertion site. I showed previously that the relatively weak cska promoter/enhancer element of *X. borealis* might be used to generate novel patterns of expression (section 3.2). Accordingly, I found one novel pattern among seventy fish injected with the control construct. Injections of transposon without transposase resulted in ~2 % (2/128) novel patterns. Noticeably, *SB10* transposase augmented pattern formation almost 3-fold: nineteen out of 432 screened fish injected with the complete *SB* system exhibit novel *GFP* expression patterns.

Among those 21 lines with novel expression patterns, we found a variety of different patterns ranging from single cell types to whole organs (Fig. 14A-I and Tab. 4). A number of lines expressed *GFP* in ectodermally derived organs, but we also identified specific expression in tissues derived from other germ layers. For several lines, inheritance of appropriate transgene expression was observed now up to the 7th generation.

Thus, among the transgenic lines we have established up to now, novel *GFP* patterns are generated in an unbiased fashion. We found lines specifically expressing *GFP* in the anterior retina (Fig. 14A), telencephalon and mid-hindbrain boundary (Fig. 14B), otic vesicles (Fig. 14C) or in lens and specific cells of the hindbrain (Fig. 14D). Among others, we identified transgenic lines with *GFP* expression in olfactory bulbs, olfactory neurons and *habenulae* (Fig. 14E), along the notochord (Fig. 14F), in ventricles of the heart (Fig. 14G), the yolk (Fig. 14H), somites and trunk muscles (Fig. 14I)(Grabher et al., 2002; Henrich, 1999).

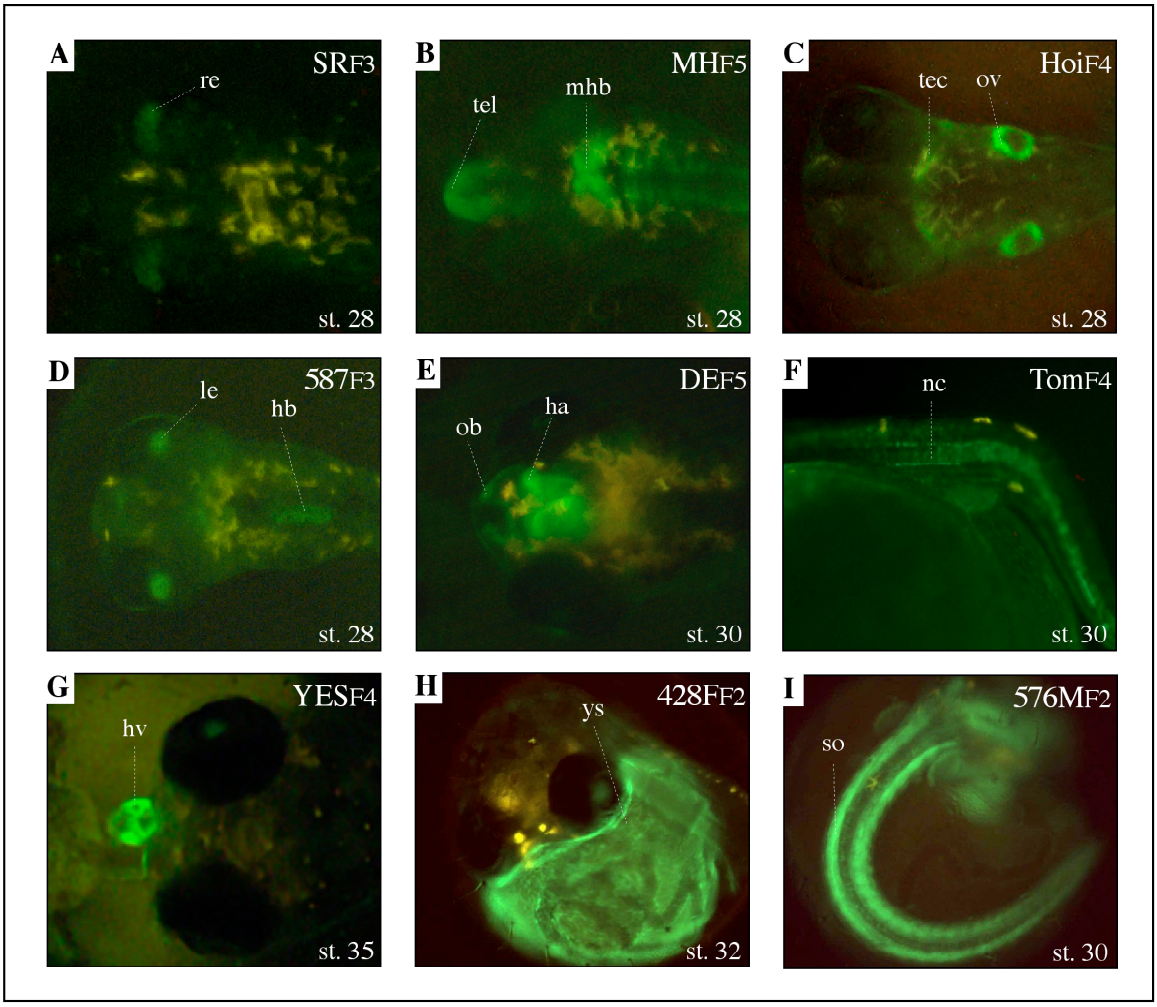
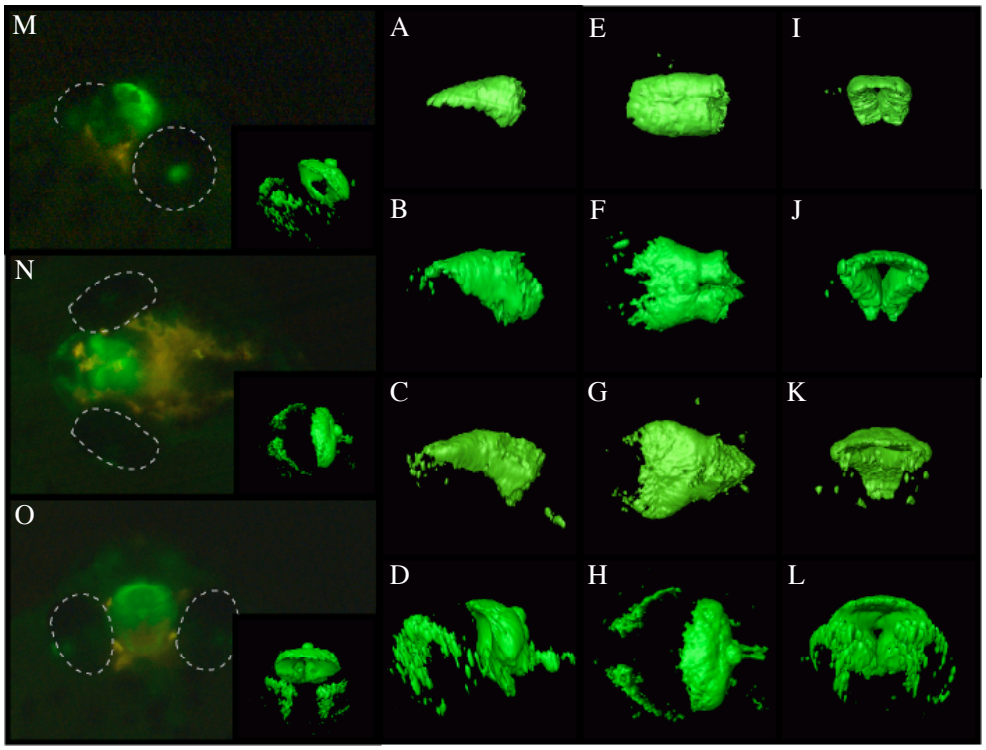


Fig. 14: Stable transgenic lines with spatially and temporally restricted *GFP* expression patterns.
Collection of nine specific transgenics (A-I) showing expression of *GFP* in various tissues or organs; developmental stages are indicated; anterior is to the left (A-G). A, SR, *GFP* expression is enhanced in anterior cells of the retina. B, MH, strong *GFP* expression in telencephalon, mid-hindbrain boundary and along the spinal chord. C, Houichi (Hoi), otic vesicles and tectum show expression of *GFP*. D, 587, *GFP* is enhanced in lens and a row of cells in the ventral hindbrain. E, DE, olfactory bulbs, olfactory neurons and *habenulae* exhibit *GFP* expression. F, Tomoe (Tom), cells of the notochord express *GFP*. G, YES, ventricles of the heart are *GFP* positive. H, 428F, *GFP* expression in epidermis of the yolk, frontal view. I, 576M, strong *GFP* signal in somites, dorsal view. ha, *habenulae*; hb, hindbrain; hv, heart ventricles; le, lens; mhb, mid-hindbrain boundary; nc, notochord; ob, olfactory bulbs; ov, otic vesicles; re, retina; so, somites; tel, telencephalon; tec, tectum; ys, yolk.

The *GFP* expression pattern of the transgenic line DE was analyzed in more detail using confocal microscopy and 3D-surface reconstruction. Medaka embryos are covered with a 2-layer chorion that has a hard inner layer and a soft outer surface. The chorion interferes with the confocal optics reducing the image quality. Therefore,

embryos dechorionated with medaka hatching enzyme were used for confocal analyses of living embryos. The hatching enzyme is a protease, which is secreted from the hatching gland and dissolves the inner layer of the chorion (Lee et al., 1994). The remaining soft outer layer can then be removed with sterile forceps. Dechorionated embryos develop normally and at the same rate as embryos with a chorion. However, more careful handling is necessary as dechorionated embryos are fragile. Living embryos at various developmental stages were embedded in agarose and subjected to confocal microscopy.



Fig, 15: The pattern of *GFP* expression of the transgenic line DE in development.
The figure shows a medaka embryo at various stages of development (A-D, M lateral, E-H, N dorsal, I-L, O frontal). At st. 23 (A, E, I) expression is already strong in the presumptive diencephalon. In subsequent stages (st.25: B, F, J; st 28: C, G, K) this area is expanding and reaches the olfactory placodes. At stage 30 *GFP* expression appears restricted to the olfactory bulbs, olfactory neurons and the epiphysis (D, H, L, M-O). Additionally, *GFP* covers the dorsal part of the *habenulae* and more ventrally several pretectal and diencephalic nuclei (see also Fig. 16).

Series of optical section were thereafter used to perform computational 3D-surface reconstruction of the *GFP* positive tissue. DE shows specific expression of *GFP* in the telencephalic/diencephalic region (Fig. 16). Expression is already strong at stage 23 (41 hpf/12 somites) in the dorsal brain region between the eyes, which will later give rise to the diencephalon. In subsequent stages, this area is expanding in size reaching more ventral tissues until it becomes visible in the developing olfactory placodes. At stage 30 (82 hpf/35 somites) *GFP* expression appears restricted to the olfactory bulbs and neurons and the epiphysis. Additionally, *GFP* covers the dorsal part of the *habenulae* and probably the anterior optic tectum. Expression of *GFP* includes several pretectal and diencephalic nuclei that correspond to one or two layers of the *torus semicircularis*, which is a caudal and dorsal part of the mesencephalon. Its homologue in mammals is called the *inferior colliculus* and it is an important centre for processing auditory information.

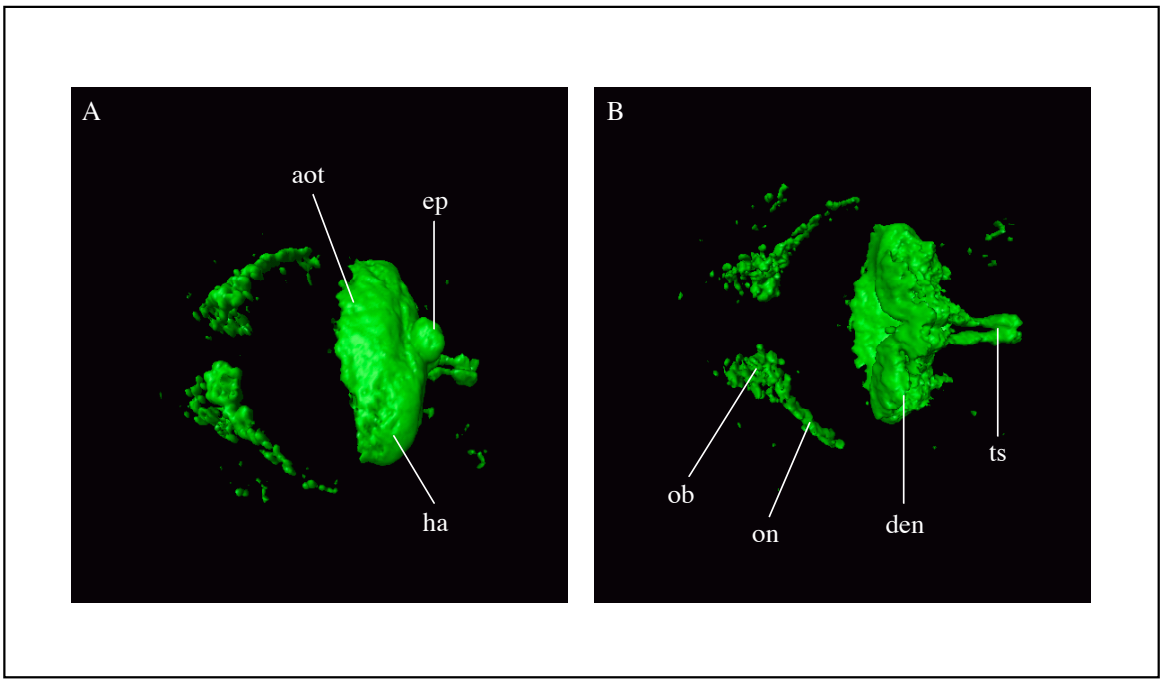


Fig. 16: Close-up of a 3D-surface reconstruction of a transgenic DE embryo.
DE embryo at stage 30. Anterior is to the left. A, Semilateral view of *GFP* expressing tissue; anterior optic tectum (aot), epiphysis (ep) and *habenulae* (ha). B, Ventral view of the same embryo; olfactory bulbs (ob), olfactory neurons (on), diencephalic nuclei (den) and *torus semicircularis* (ts)

Line	Injection	Germ line transmission frequency	Type of insertion	Integration mechanism	*GFP* expression pattern
381 #1 (SV)	Transposon + SB10	nd	Insert only	SB10	Ubiquitous
381 #2 (YES)	Transposon + SB10	nd	Complete tandem array	Illegitimate SB10	Heart ventricles
381 #3 (587)	Transposon + SB10	nd	Complete tandem array	Illegitimate SB10	Lens, *habenulae*, pineal gland, posterior tectum, central row of cells in ventral hindbrain
381 #4 (DE)	Transposon	8.1 %	Complete tandem array	Random plasmid integration	Olfactory pits, diencephalon, epihysis
381 #5 (576M)	Transposon + SB10	27.7 %	Partial array	Random plasmid integration	Somites, muscles
381 #6 (428F)	Transposon + SB10	55.5 %	Partial array	Random plasmid integration	Yolk sac
381 #7 (Tom)	Transposon + SB10	10.3 %	Partial array	Random plasmid integration	Anterior neural tube, posterior notochord
MH	Transposon + SB10	23.7 %	nd	nd	Telencephalon, mid-hindbrain boundary, CNS
Hoi	Transposon + SB10	6.1 %	nd	nd	Otic vesicle
294 #1 (SR)	Control	15.8 %	Complete tandem array	Random plasmid integration	Retina

Tab. 4: Summary of presented transgenic medaka enhancer trap lines.
The enhancer trap lines presented in this thesis are summarised here. Type of injection, germ line transmission frequency, type of insertion, the proposed integration mechanism and a description of the *GFP* expression pattern are described where determined. nd, not determined

Transgenic lines exhibiting such a specific marker gene expression are perfectly suited to sort *GFP* positive cells (FACS) and to generate tissue specific cDNA libraries to be used for the identification of specifically expressed genes. In addition, such lines and their libraries facilitate the determination of gene expression variations in a mutant background.

Discussion

5. *I-SceI* Meganuclease and the *SB* Transposon System Mediate Highly Efficient Transgenesis in Fish

In this work, I aimed for the development of novel or enhanced transgenesis technologies to facilitate studies in developmental biology using the medaka fish, *Oryzias latipes*. The availability of fast and simple techniques to introduce foreign DNA into fish embryos allowing the expression of transgenes is a prerequisite for biologists to investigate the crucial processes of embryogenesis.

Both, transient expression of episomal transgenes as well as stable integration into the genome and expression in subsequent generations are equally important. Moreover, generation of transgenic fish exhibiting novel tissue- or cell-specific patterns of reporter transgene expression (enhancer trap) provides useful tools for the analysis of these tissues or cell populations and the transcriptional regulation of the trapped regulatory units during embryonic development under various conditions.

Microinjection of DNA into fish embryos at the one cell stage provides the fastest and easiest approach. Unfortunately, the quality of transient promoter-dependent expression of the transgene and the efficiency to generate stable transgenics are very low (Collas and Alestrom, 1998; Culp et al., 1991; Lin et al., 1994a; Stuart et al., 1988; Stuart et al., 1990; Tanaka and Kinoshita, 2001; Westerfield et al., 1992). In addition, the frequency of enhancer traps in fish is even lower which prevented the routinely use of this approach in fish (Bayer and Campos-Ortega, 1992).

I applied two different technologies to overcome these previous limitations. I established a meganuclease-mediated approach and a transposon-based approach enabling rapid and stable integration of transgenes into the genome of medaka. Both techniques strongly enhance transgenesis in fish at several levels.

Application of the meganuclease results in an improved G0 expression (transient expression) of an injected DNA construct (Fig. 7): mosaic expression of the reporter gene in injected embryos is greatly diminished, overcoming one of the main

pitfalls of transgenesis in fish. This result opens the way to easy and reliable promoter studies in fish already in G0 without the immediate need to establish stable transgenic lines.

Second, this technique significantly increases the frequency of positive founder fish in G0, thus improving transgenesis frequency (Tab. 3). I also observed that expression of reporter gene in adult injected fish correlates with the transmission of the transgene to the next generation. Therefore, the tools presented here greatly simplify the time- and space-consuming selection of transgenics by identifying potential founder fish already in the G0 generation.

Third, and maybe most strikingly, the germ line transmission rates in transmitting families reach about 50 % in *I-SceI* experiments (Tab. 3). Thus, a few G0 fish selected by *GFP* expression are sufficient to establish a transgenic line with limited screening effort in F1. Southern blot analysis indicated single to few integration events in individual lines. Germ line transmission rates of close to 50 % may be due to a single integration of the transgene into one-cell stage embryos, leading to non-mosaic transgenic fish. Equally possible are multiple independent insertions in different cells at later stages, creating a mosaic germ line by which the different insertions are inherited to different F1 fish. In both cases, transgenic carriers are easily identified. However, the observed tight correlation of uniform *GFP* expression in G0 with high germ line transmission rates and the segregation analyses in cases with several integrations consistently argue for early integration event(s) in meganuclease injected embryos leading to uniform G0 patterns and transgenesis.

Transgene integration never occurred as long concatemers, a feature otherwise encountered in transgenic fish (Hackett, 1993; Iyengar et al., 1996), and known to eventually result in gene silencing in vertebrates (Garrick et al., 1998). Instead, I found that the transgene was integrated in short repeats and consequently not silenced (Fig. 8). In these two independent series of stable transgenic lines, I detected integrations of repeated units of inserts linked to plasmid still bearing *I-SceI* sites (Fig. 8). The reason why *I-SceI* sites are still present in the transgene insertions of these fish remains unclear. One likely hypothesis is that *I-SceI* is unable to fully overcome the strong ligase activity present in fish egg cytoplasm, by cutting the concatemers. The enzyme indeed only cuts iso-stoechiometrically and remains linked to the longer half of the

recognition site after cleavage (Colleaux et al., 1988). In addition to the requirement for cleavage by *I-SceI*, it is possible that the continued stable binding of the enzyme to the DNA ends plays a role in the high frequencies of transgenesis, possibly by protecting the linear monomers from degradation.

The very long recognition site of *I-SceI* renders the meganuclease a very rare cutter (once in $7x10^{10}$ bp). Therefore, it appears unlikely that *I-SceI* cuts the medaka genome (10^9 bp). Furthermore, all investigated transgenic lines showed different insertion patterns in Southern analysis. I thus think that the improved transgenesis efficiency by the meganuclease is mediated by a mechanism different from that described for REMI (restriction endonuclease mediated integration).

Taken together, the co-injection of the integration construct with the *I-SceI* meganuclease that cuts only flanking to the insert leads to early integration of a functional insert already at the one-cell stage. The co-injected enzyme likely counteracts the endogenous ligase activity, preventing the generation of long concatemers found upon the injection of circular or linear DNA and thus provides more recombinogenic ends that facilitate highly efficient integration. Due to this early integration, transgenic founders are easily identified in G0. Furthermore, the early integration warrants a very high germ line transmission rate of nearly 50 %.

The second technique introduced in this thesis, the *SB* transposon system, also facilitates the generation of transgenic medaka fish. Transgenesis is strongly enhanced by the presence of the *SB* recognition sites alone (transposon), even in the absence of *SB10* transposase. Other than in "conventional" transgenesis, the expression of the transgene is stable in subsequent generations even if it is integrated in tandem arrays.

In contrast to control injections, the mosaicism of reporter gene expression was greatly reduced in G0 fish injected with the *SB* transposon. Efficient promoter-dependent expression in G0 depends on the presence of IR/DRs. Although ~45 % of these G0 fish are transgenic founders, the other half was not transmitting *GFP*, indicating that, in contrast to the meganuclease approach, in those fish ubiquitous expression was not mainly due to early integration. Thus, widespread *GFP* expression in G0 appears to be due to an equal distribution of extrachromosomal plasmid DNA. A similar mechanism, the direction of the injected DNA to the nucleus, has been

suggested for the ITRs of AAV in *Xenopus* and zebrafish (Weitzman et al., 1996). This allows an even segregation of transient, non-integrated reporter DNA. Consequently, the *SB* transposon also provides a useful tool for transient expression studies upon injection of G0 embryos.

Furthermore, the uniform expression of the transgene in G0 embryos injected with transposon only or the complete *SB* system is a reliable marker for the efficient selection of transgenic founders, limiting a time- and space-consuming screening effort in F1. It was shown that *SB10* acts with the proposed mechanism integrating a single copy insert sequence into a TA dinucleotide in the transgenic line SV. However, transgenesis frequencies using the transposon are comparable with or without transposase (32 or 29 %, respectively) and thus appear relatively independent of the *SB10* transposase. Do IR/DRs influence transgene integration of inserted DNA? For AAV-ITRs, palindromic sequences that fold into hairpin structures and function as origins of replication, it has been suggested that they also play a crucial role in the process of integration (Cooley et al., 1988). Mammalian cell culture experiments revealed an improvement of integration frequency when reporter genes were flanked by AAV-ITR sequences over control reporters (Balague et al., 1997). The IR/DRs of *SB* may perform similar functions, although such structural properties remain to be shown.

Transgenesis by direct microinjection of DNA is the most convenient and cost effective technique for many vertebrate prolific species (fish, mouse, rat, rabbit, pig, cow). However, the rate of foreign gene integration used to be very low and constituted the major limitation for transgenesis. In these species the meganuclease-mediated approach or the *SB* transposon system technique, which are both simple and efficient, have many potential applications in basic research and biotechnology.

Although, different routes to enhance transgenesis have been followed since the first report of a transgenic fish (Stuart et al., 1988) (Müller et al., 2002), only little progress was achieved in establishing ‘enhancer-trap’-like technologies so far (Bayer and Campos-Ortega, 1992), mainly due to low transgenesis frequencies.

Here I present that both techniques introduced above can be used as highly efficient tools for the generation of transgenic medaka fish with an intriguing potential to randomly generate novel patterns of *GFP* expression.

Transgenesis mediated by *I-SceI* leads in one out of twenty cases to temporally and spatially specific expression patterns most likely due to the insertion of the transgene in the vicinity of enhancer elements (Fig. 9). This indicates that the reporter construct is sensitive to position effects and may be useful for enhancer or gene trapping strategies. Most of the meganuclease-mediated transgenics showed promoter-dependent expression patterns. I suggest that the single transgenic line with a specific pattern resulted randomly due to the high general transgenesis frequency, in contrast to any specific role that could be appointed to the meganuclease itself.

However, using the same reporter construct flanked by IR/DRs, the *SB* system is more efficient in generating transgenic fish that exhibit novel specific gene expression patterns. These patterns range from single cell types to larger tissues or organs (Fig. 14), and do not show a bias for any germ layer.

Strikingly, compared to control injections or injections of transposon alone, in the presence of *SB10* transposase, novel expression patterns were found at high frequency, indicating a preference of integration adjacent to regulatory sequences that is facilitated by the transposase. *SB10* transposase co-injection uniquely increases the generation of novel random patterns of reporter gene expression almost three-fold compared to injections of control constructs or transposon in the absence of transposase. This result is not only due to the high integration frequency, because techniques such as the meganuclease approach, that reaches comparable transgenesis efficiencies, do not result in a comparable trapping activity (Thermes* et al., 2002). Approaches using viral constructs in fish, although leading to efficient integration and mutagenesis, have not been shown to be useful as enhancer trapping tools so far. This renders the *SB* transposon system also a promising tool to perform gene trap experiments.

Similar to the P-element of *Drosophila* (Tsubota et al., 1985), *SB* integration in medaka might be favoured in the vicinity of transcriptional regulatory elements. The target site for *SB* transposition is part of a palindromic AT-rich sequence (ataTAtat) (Vigdal et al., 2002). Transgene insertion may be directed to AT-rich scaffold attachment regions or matrix-attached regions by the *SB10* transposase even if the integration event itself is not mediated by the transposase.

This could account for the increased frequency of novel expression patterns obtained by transposon/transposase co-injections. These regions are frequently comprised of control elements that maintain independent realms of gene activity (Vigdal et al., 2002). Fractions of which might also cohabit with transcriptional enhancers or silencers (Boulikas, 1995; Sandmeyer et al., 1990).

The enhancement of transgenesis frequencies combined with the random generation of novel expression patterns using the *SB* transposon system enables the fast and simple generation of a wide range of random *GFP* expression patterns. For an average sized lab 5000 injections per month are a feasible goal and, enhanced by the *SB* system, could give rise to more than 900 transgenic lines, of which 110 are expected to show a differential expression pattern. Thus, a set of transgenic lines expressing *GFP* in developmentally important structures/organs can be established and used without devoting a major effort on the isolation and characterization of promoter elements.

These enhancer trap lines are useful for a variety of applications relevant for basic or applied research. Transgenic fish showing specific spatial or temporal expression patterns can be used as tissue specific molecular markers especially in the context of mutant analysis. Crossing a specific *GFP* transgenic to a mutant of interest allows investigation of the *GFP* expressing tissue during development in the mutant background that will result in a fast elucidation of subtle changes of organ or cell development *in vivo*. Specific expression patterns within the brain provide interesting tools for neuroanatomists to identify developmentally or functionally related structures in the embryonic and adult brain. Moreover, analysis can be performed using confocal microscopy, time-lapse and 3D-reconstruction to obtain a highly realistic view of embryonic development in all dimensions (see supplementary CD). *GFP* expressing cells can easily be isolated by fluorescence-activated cell sorting (FACS), which in turn allows the establishment of pure cell cultures and the construction of tissue or cell type specific cDNA libraries. In the context of mutant analysis, such cDNA libraries enable subtractive approaches to compare gene expression profiles between wild type and mutant cells or tissues. Many more applications can be realised depending on the nature of the transgenic line available. Finally, yet importantly, these lines can be used to

identify the regulatory elements that are responsible for the spatially and temporally restricted gene expression. For the rapid isolation of genomic DNA flanking the insertion, splinkerette PCR (Devon et al., 1995) was successfully applied for the transgenic line SV. However, independent insertions need to be segregated by successive out-crossing prior to cloning attempts, which is possible within one generation as shown for the transgenic line DE. Problems with tandem arrays can be overcome by pre-selection of the flanking fragments after Southern blotting. The ongoing sequencing of the medaka genome (Shima et al., 2003; Wittbrodt et al., 2002) and genomic resources available are crucial prerequisites for the isolation and identification of DNA sequences near insertion sites on a larger scale and for the characterisation of the genes controlled by the trapped regulatory elements *in vivo*.

5.1 Future Aspects

5.1.1 Gene Targeting

The *I-SceI* meganuclease has been used previously in mammalian cells (Johnson and Jasin, 2001) and *Drosophila* (Lankenau et al., 2003; Rong and Golic, 2000; Rong and Golic, 2001; Rong et al., 2002) for gene targeting. Gene targeting is the modification of an endogenous gene by recombination between an exogenous DNA fragment and a homologous endogenous target gene, mediated by a DNA repair mechanism of the host. This allows specific ablation of genes to study their function in a complete loss-of-function approach. So far, the only vertebrate amenable to this technique is the mouse, taking advantage of the ES cell technology.

Cells have evolved numerous repair pathways to contend with various types of DNA damage (Friedberg et al., 1995). The significance of DNA repair is apparent, as defects in repair mechanisms are linked to diseases and malignancy (Vogelstein and Kinsler, 1998). One type of rare but severe lesions, a DNA double-strand break (DSB), poses a particular threat to genomic integrity. In bacteria and yeast, homologous recombination (HR) has long been known to be a major mechanism for the repair of DSBs. In addition to non-homologous end joining (NHEJ), a pathway exclusively dedicated to the repair of DSBs, HR has also been shown to be a major DSB repair pathway in mammalian cells (Rouet et al., 1994). By NHEJ, DNA ends are joined with little or no base pairing at the junction and the end joining may be associated with insertions or deletions. In order to repair a DSB by HR, a second DNA molecule with homology to the region to be repaired must be available to serve as template. In this process, called gene conversion, the information from the donor sequence is copied into the broken locus, making the repaired locus an exact copy of the donor sequence.

The mouse model offers the advantage of ES cells that are transfected with a linearized donor vector and subsequently screened by positive/negative selection for successful homologous recombination events. Linearization of the donor vector provides DNA containing DSBs, by that activating the host DNA repair machinery (Muller, 1999).

In *Drosophila,* lacking ES cell technology, it is difficult to introduce a linear DNA molecule into germ cells. Recently, a method to generate such a linear fragment *in vivo* has been reported, accompanied by the demonstration of gene targeting. For targeting, the FLP recombinase and *I-SceI* meganuclease expression are induced to generate the DSBs that stimulate HR. Additional studies have also shown that in principle gene targeting in *Drosophila* could also be achieved by using *I-SceI* alone, although at lower efficiencies (Gong and Golic, 2003). These low efficiencies have been coupled with a highly efficient repair of *I-SceI* mediated DSBs that in the design of this particular experiment performed by Gong and Golic results in low amounts of donor vector.

The results I obtained on transgenesis frequency applying the meganuclease in medaka suggest that *I-SceI* is actively participating in an integration event. Not only the linearization step itself, producing DSBs, promotes integration into the genome as injection of *in vitro* linearized DNA fragments resulted in lower transgenesis frequencies. This is indicative of an additional function performed by the meganuclease as discussed earlier.

I-SceI meganuclease thus provides potential to be used for gene targeting also in medaka. The situation for such an approach resembles more the situation given for *Drosophila* than for the mouse model as medaka ES cells so far failed to contribute to the germ line (Hong et al., 1998). Also lacking small reporter genes that could be used to screen for successful recombination events, the screening procedure in medaka will demand other methods. Injection of hundreds to thousands of medaka embryos is a manageable task, however PCR screening of G0 fish and F1 offspring would overload the capacity of fish facilities of most labs. Thus, efficient screening in G0 is mandatory. In medaka, the green fluorescent protein (*GFP*) provides an alternative to mini marker genes. Detailed knowledge of target gene sequence and genomic structure is a prerequisite for gene targeting. A donor vector should contain sequence of the first or second exon, in frame disrupted by a promoter-less *GFP* and flanked by two *I-SceI* recognition sites (Fig. 17).

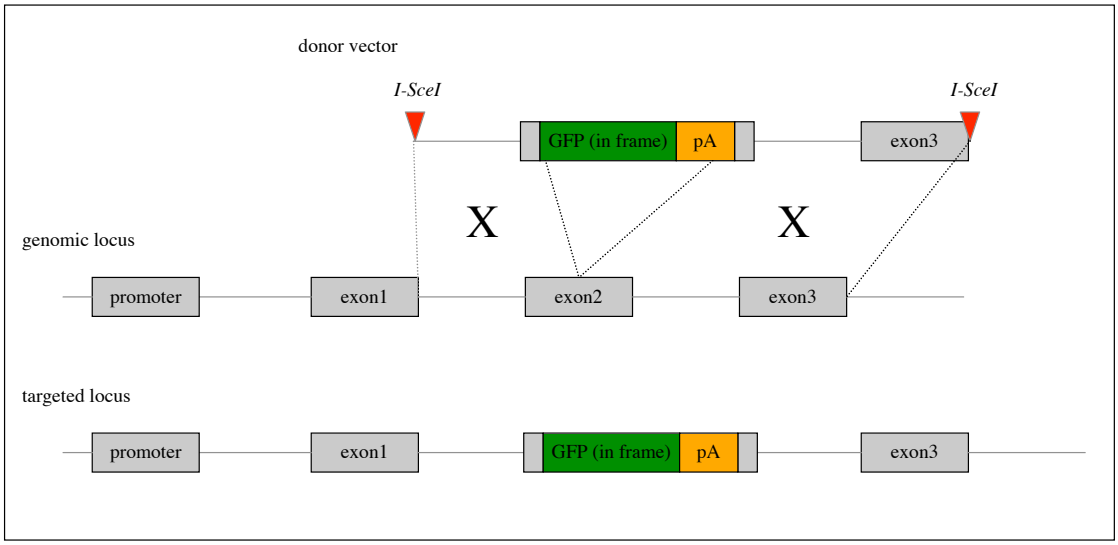


Fig. 17: Replacement gene targeting.
The donor vector contains a *GFP* insertion in the second exon of a target gene. The homology region includes 5' and 3' intron sequences and the third exon flanked by two *I-SceI* sites. The donor vector is linearized *in vivo* or *in vitro* by *I-SceI* co-injection and digestion. After homologous recombination, the second exon of a target gene is disrupted by an in frame *GFP* insertion followed by a poly adenylation signal and several stop codons. Successful HR events are scored upon *GFP* expression according to endogenous promoter control.

The homology region on either end of the *GFP* open reading frame should be as large as possible (0.5 to 10 kb). It will be important that the 5' homology region does not contain functional promoter sequence. Ideally, co-injection of meganuclease with circular or *in vitro* linearized vector results in ends-out or replacement targeting. The donor vector would replace the endogenous exon; subsequently successful HR is scored by *GFP* expression controlled by the endogenous promoter of the target gene (Fig. 17). Only *GFP* expressing embryos will then be raised and tested for germ line transmission.

To achieve homology regions large enough to enhance HR probability also intronic sequence may be used albeit the use of isogenic sequence should then be aimed for (te Riele et al., 1992). The design of such a donor vector is currently under development in our laboratory.

5.1.2 Genetic Transposition

Besides the potential of transposons to enhance the initial integration of exogenous DNA into the genome of a host as discussed above, transposons offer more possibilities.

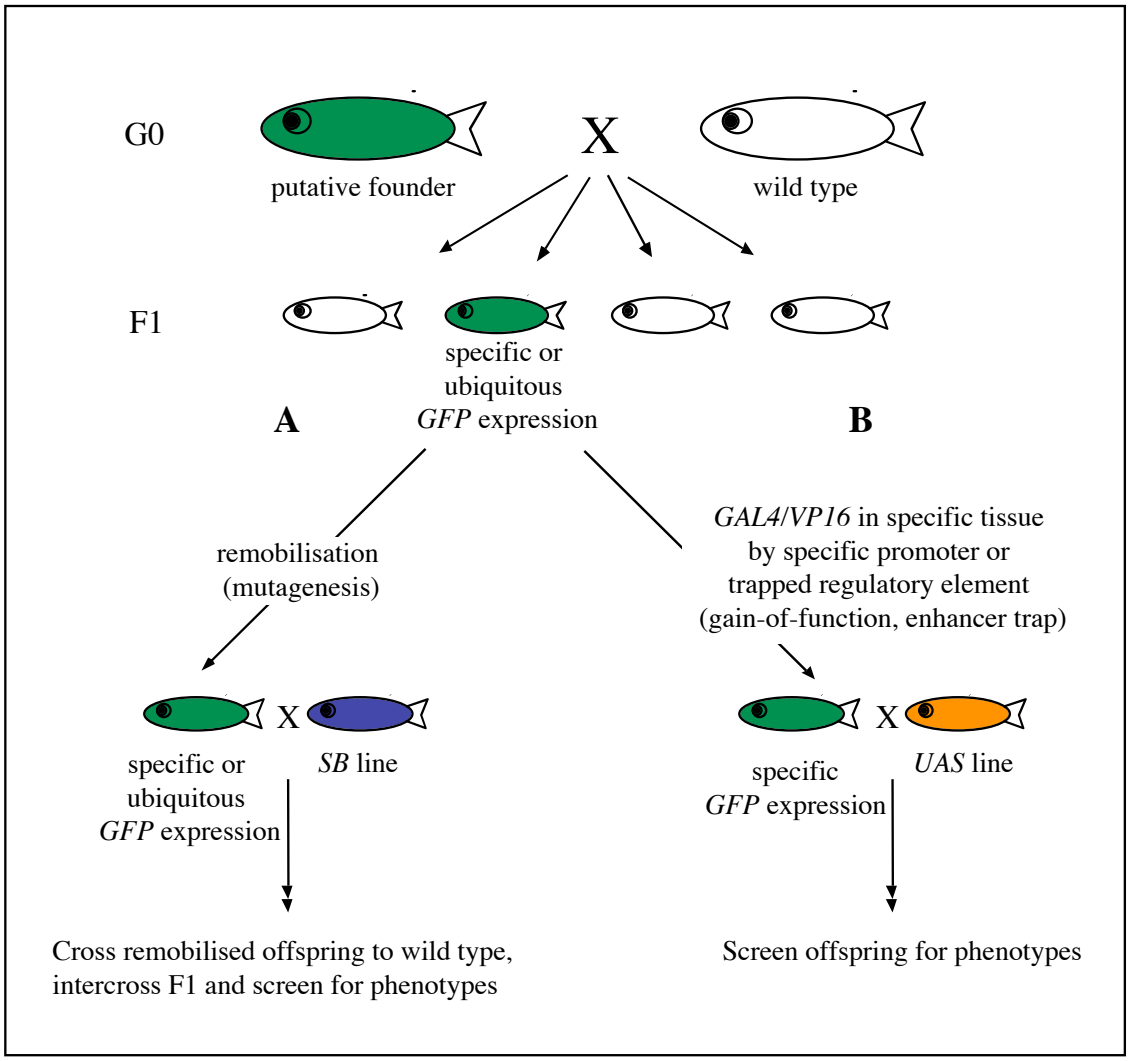


Fig. 18: Scheme for a transposon-mediated genetic insertional mutagenesis or enhancer trap screen.
A, After initial transgenesis, transgenic F1 offspring (blue) carrying the transposon may be crossed to a transposase source line (green) harbouring the *SB* transposase under control of a germ line specific (*vasa*) or heat-shock inducible promoter. Offspring thereof may be screened for mutant phenotypes. B, Alternatively, transgenic F1 offspring carrying a transposon (with a specific or heat-shock inducible promoter driving *GAL4/VP16* and *UAS/CFP* as an internal control) (blue) may be crossed to a *UAS* effector line harbouring a gene of interest under *UAS* control (orange).

Transposons are able to move actively within a genome. This feature renders transposon systems an attractive candidate for insertional mutagenesis or enhancer- and gene-trapping experiments. In contrast to insertional mutagenesis by retroviral vectors that can integrate into the host genome only once, transposons may be used for repeated insertions. One initially integrated transposon can be activated to remobilize by its transposase. This is especially interesting concerning transposons that move in a cut-and-paste mechanism. Providing a conditional transposase source within the germ line, reporter gene-containing transposons can be remobilized to identify insertions that in the next generation result in mutant phenotypes linked to the reporter gene expression (Fig. 18A).

In theory, a single fish with a transposon insertion mated to a fish containing a transposase source could be used to perform a permanent F1 mutagenesis screen. Insertional mutagenesis bears the advantage of providing a marker of known sequence to the mutagenized locus. Thus, the identification of the affected gene is highly facilitated, and the adjacent genomic sequences can be determined by inverse PCR. This resembles the principle applied in *Drosophila*, where females carrying the gene encoding for the transposase can be simply crossed with males carrying a non-autonomous P-element, resulting in remobilization of the P-element in the offspring and insertion into other loci (Robertson et al., 1988). The mandatory prerequisite, however, is a highly efficient transposon system.

Similarly, initially integrated enhancer- or gene-trap vectors may be remobilized to isolate novel expression patterns. The injection of hundreds or thousands of single embryos as described previously (see section 4) could be omitted by application of a stable genetic transposase source. The *SB* system, independent of host factors, may provide the prerequisites to apply sophisticated technologies developed in invertebrates for the use in a vertebrate like medaka.

5.1.3 The GAL4/UAS *System in Medaka*

The most common way to analyze the function of any gene cloned in fish is to mis-express its wild type product, or an altered variant of it, by mRNA injection. This method rapidly yields insights into the developmental function of a gene, but it is also hampered by some disadvantages. Thus, mRNA injections are unspecific with respect to the tissue and developmental stage of expression. This makes it difficult to determine the function of a gene product in any given process. In addition, if the gene product plays a role during early stages of embryogenesis, the phenotypic consequences may obscure the effects on later stages. So far, no methods are available for the mis-expression of a gene product in medaka in a directed stage- and tissue-specific manner.

In *Drosophila,* on the other hand, one such method – the *GAL4/UAS* system (Brand and Perrimon, 1993) – is routinely used to analyze the function of developmental genes (e.g. (Brand Andrea et al., 1995)). The technique is based on two different kinds of transgenic strains, called activator or driver and effector lines. In an activator line the gene for the yeast transcriptional activator *GAL4* is placed under the control of a specific promoter, while in the effector line the gene of interest is fused to the DNA-binding motif of *GAL4* (Upstream Activating Sequences, *UAS*). The effector gene will be transcriptionally silent unless animals carrying it are crossed to those of an activator line. In the progeny of this cross, expression of the effector gene will reflect the pattern of expression of *GAL4* in the activator, which is ultimately dependent on the promoter that has been used to control it. This, of course, allows controlled ectopic expression of the effector gene. The establishment of the *GAL4/UAS* method for targeted gene expression in medaka is highly desirable for several reasons (Scheer and Campos-Ortega, 1999). A steadily increasing number of genes cloned from medaka could then be analyzed in more detail than is feasible with mRNA injections. Expression of *GAL4* under the control of a heat-shock promoter in a variety of stages and tissues through activation by UV or IR laser in single cells or tissues could provide a meaningful tool for developmental studies (Halloran et al., 2000). In combination with a transposon system like *SB,* various driver lines could be randomly generated in an enhancer like fashion by transposition (Fig. 18B).

Appendix to Discussion

7. Appendix A: Application of the *I-SceI* Meganuclease in Medaka

7.1 Establishing a Heat-Shock Inducible *GAL4* Driver Line Reveals a Major Toxicity Upon Over-expression

In order to investigate the function of known genes within specific tissues in an organism I combined the meganuclease system and the *SB* system with the *UAS/GAL4* system used in flies (Brand and Perrimon, 1993).

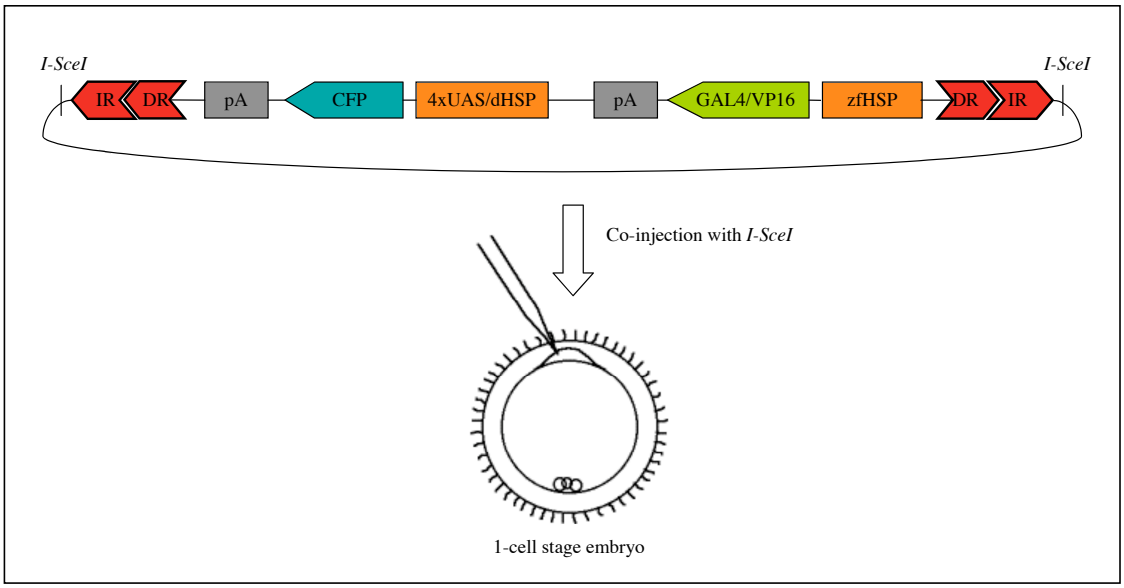


Fig. 19: Generation of a heat-shock inducible *GAL4/VP16* driver line.
I-SceI recognition sites flank the insert to facilitate genomic integration. *SB* inverted repeats are placed at the inner flanks to allow subsequent remobilization of the insertion. A zebrafish heat-shock promoter drives expression of the *GAL4/VP16* fusion protein. *CFP* is placed under control of four *UAS* elements on the same vector as an internal control of *GAL4/VP16* expression. Circular plasmid (pCG 6.0Sce) was co-injected with *I-SceI* meganuclease and surviving embryos were raised to sexual maturity. F1 offspring of injected G0 x wild type matings were screened for *CFP* expression by a 2 minute heat-shock at 42 °C in a waterbath.

This would allow establishment of specific driver lines expressing *CFP* as internal marker driven by *GAL4/VP16*. By crossing the driver lines to effector lines containing genes of interest under *UAS* control specific miss-expression studies could be performed.

Moreover, the transposon system would eventually allow the remobilization of insertions, combining the *GAL4/UAS* system with a genetic enhancer trap screen (Fig. 18). A heat-shock inducible driver line was successfully established (Fig. 19, 20).

Although control experiments involving transient expression of *GAL/VP16* were promising, stable transgenics revealed a major toxicity of *GAL4/VP16* when systemically expressed from within the genome. This leads to general retardation of *GAL4/VP16* expressing embryos compared to heat-shocked control embryos (Fig. 20).

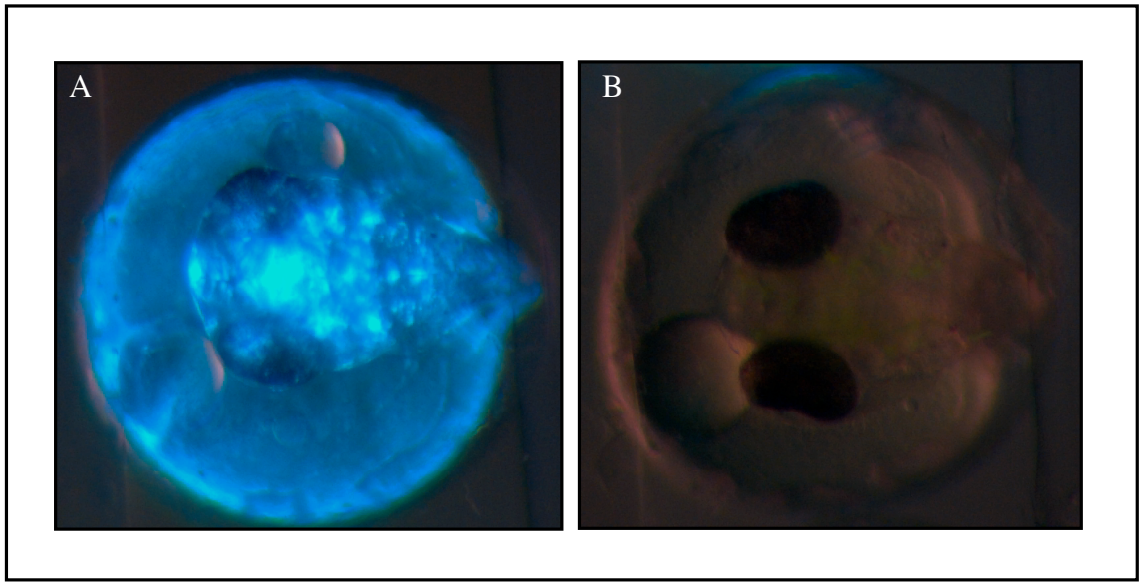


Fig. 20: Genomic expression of *GAL4/VP16* results in a retardation phenotype.
Transgenic medaka embryos were heat-shock induced for 5-30 minutes at 42 °C. Dorsal view; anterior is to the left. Expression of *GAL4/VP16*, monitored by *CFP* expression, leads to retardation of embryonic development. The severity of the phenotype is dose dependent. Short induction (5 minutes) shows less effect on embryogenesis but does not result in detectable *CFP* expression within the entire embryo. Long induction (>5 minutes) results in high *CFP* expression leading to strong retardation and embryonic lethality. A, transgenic embryo after 10 minutes at 42 °C. B, wild type sibling after 10 minutes at 42 °C as a control.

The degree of toxicity is dependent on the expression level. This dose dependence suggests that the *GAL4/VP16* fusion protein sequesters the basal transcription machinery in excess, thereby resulting in a retardation phenotype. Although, I could show that the *GAL4/UAS* system is principally usable in medaka, transactivation by the *GAL4/VP16* fusion protein appears to be too strong.

The potential applicability of this transgenic line for induction of *GAL4/VP16* within single cells or groups of cells by IR or UV (Halloran et al., 2000) laser is under investigation in collaboration with a Japanese research group. Laser induction might allow more subtle induction that in addition will not affect the entire embryo. Furthermore, application of wild type *GAL4* protein, *mini-GAL4* proteins (Wu et al., 1996) exhibiting lower transactivation potential or identification of the minimum duration of a heat-shock resulting in sufficient but non-toxic *GAL4* activation might solve this problem.

8. Appendix B: The *SB* Transposon System in *Oryzias Latipes*

8.1 Repeated Germ Line Transposition Does Not Occur Upon Genetic Transposase Induction

Owing to their inherent nature to move from one chromosomal location to another within and between genomes, transposable elements have been exploited as genetic vectors for genetic manipulations in several organisms (Bellen et al., 1989; Jaenisch, 1988). Transposon tagging is a well-established technique in which transposons are mobilized to jump into genes, thereby inactivating them by insertional mutagenesis. In the process, the inactivated genes are tagged by the transposable element, which then can be used to recover the mutated allele. Although, insertional mutagenesis is less efficient and less random than chemical treatment, it is such a powerful technique for the generation of recoverable mutations that it will undoubtedly be useful in medaka developmental genetics.

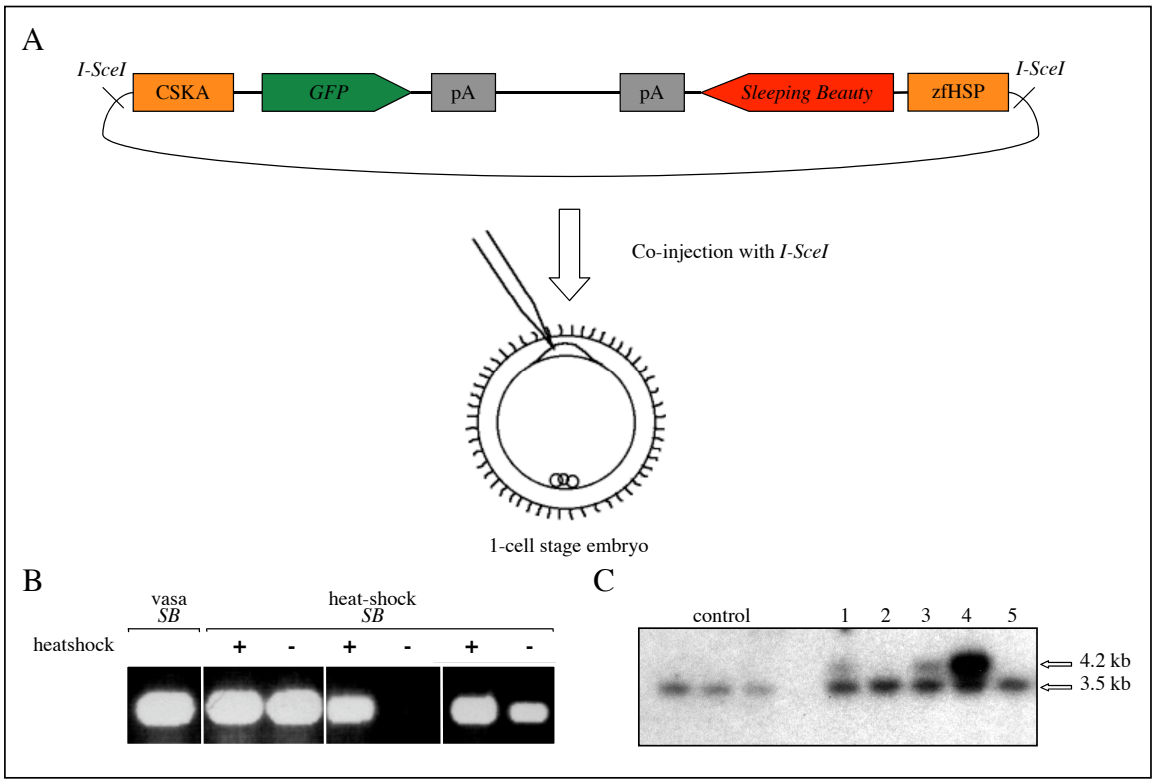


Fig. 21: Efficient repeated *SB*-mediated transposition only occurs in somatic cells.
A, circular plasmid (pzHSPSBGFPS-I) was injected into 1-cell stage medaka embryos. Putative founder fish were selected based upon their *GFP* expression, rose to sexual maturity and crossed to wild type fish. B, *GFP* expressing F1 offspring was collected and tested for *SB* inducability by RT-PCR. Similarly, embryos of a transgenic line expressing *SB* under control of the germ line specific *vasa* promoter were tested by RT-PCR. Three independent transgenic heat-shock lines and the *vasa-SB* line have been tested in B, showing constitutive or inducible *SB* transcription. C, the *vasa-SB* line was crossed to the transgenic Yes line (harbouring the transposon as tandem repeats of the entire vector). Offspring was raised to adulthood and tested for transposition events by Southern analysis. *GFP* coding sequence was used as a probe. Genomic DNA was digested with *BamHI*. Control lanes show the *GFP* signal of 3 individuals of the original Yes line. Lanes 1-5 show 5 independent F1 individuals of *vasa-SB* x Yes matings. Lanes 1,3 and 4 clearly show an additional *GFP* signal at about 4.2 kb, indicative for a transposition event. Offspring of these individuals was again investigated by Southern analysis. The novel *GFP* signal was not transmitted to the next generation (data not shown) suggesting that the initial transposition occurred in somatic cells only.

Transposon tagging can also be used in enhancer trap screens. A marker gene, such as *GFP*, can act as a reporter for genomic transcriptional enhancer-like elements located sufficiently close to the inserted transposon (see section 4.4). The same procedures that generate insertional, loss-of-function mutants or trap regulatory elements can be used to deliver genes that will confer new phenotypes to cells (see sections 3, 4 and 7). Application of transposase by mRNA injection in single embryos is too work intensive to be used for genetic screens.

The advantages of transposons take effect only if they are applicable in a genetic manner, by crossing transgenic lines harbouring a transposase source to a target line harbouring the transposon (Fig. 18). I investigated the remobilization potential of *SB* using genetically stable (constitutive) or conditional (inducible) transposase sources (transgenic lines). A germ line specific driver line (*vasa SB*) (Shinomiya et al., 2000) was already available; a heat-shock inducible driver (*HSP SB*) (Halloran et al., 2000) was established in addition (Fig. 21A, B).

Upon crossing to target lines (SV, Yes) I got promising results in G0 with both driver lines, 30 % (5/15) of the investigated fish showed an additional reporter signal on Southern blots (Fig. 21C). Suspiciously, they all were of the same size, suggesting hot spots of insertions at a nearby locus. Unfortunately, this additional signal was not transmitted to further generations. Within 35 F1 and 30 F2 individuals, none showed transposition within the germ line as validated by Southern analysis (data not shown) but only somatic transposition in G0 occurred with high efficiencies. These results led to the conclusion that the activity of *SB* is not sufficient to perform a reasonable genetic mutagenesis screen. The transposon I used was flanked by a left and a right inverted direct repeat (IR/DR). The two original recognition sequences are not identical and may reduce the transposition efficiency. Meanwhile, cell culture experiments have shown that usage of two identical left repeats does enhance the transposition frequency at least two-fold (Z. Ivics, personal communication). Application of the *SB* system with enhanced IR/DRs therefore may yield higher efficiencies also in medaka.

Materials and Methods

9. Materials

9.1 Buffers and Media

All buffers not specifically described in this section were prepared according to standard protocols (Sambrook et al., 1989) using highly deionized water (millipore), unless indicated differently. Sterilization was achieved by autoclaving.

LB Medium (Luria-Bertani Medium)

10 g	Tryptone
5 g	Yeast extract
10 g	NaCl
ad 1 l	Deionized H_2O.

pH adjusted to 7.0 with 5 N NaOH; sterilized. If necessary, 100 µg/ml ampicillin were added.

LB Agar

15 g agar were dissolved in 1 l LB medium, allowed to cool down to 60° C, supplemented with antibiotics if necessary, and poured into 9 cm diameter petri dishes.

ERM (Embryo Rearing Medium)

0.1 % (w/v)	NaCl
0.003 % (w/v)	KCl
0.004 % (w/v)	$CaCl_2x2H_2O$
0.016 % (w/v)	$MgSO_4x7H_2O$

10x Yamamoto Ringer Solution

7.5 % (w/v)	NaCl
0.2 % (w/v)	KCl
0.2 % (w/v)	$CaCl_2x2H_2O$

adjusted to pH 7.3 with HCl

SSC

3 M	NaCl
0.3 M	Sodium citrate

TE

1 mM	EDTA pH 8.0
10 mM	Tris/HCl pH 8.0

TAE

40 mM	Tris base
20 mM	Acetic acid
1 mM	EDTA pH 8.0

TEN9

100 mM	Tris/HCl pH 8.5
10 mM	EDTA
200 mM	NaCl
1 %	SDS

Denaturing Solution

1.5 M	NaCl
0.5 M	NaOH

Neutralising Solution

1.5 M	NaCl
0.5 M	Tris/HCl pH 7.5

Embryo Injection Plates

1.5 % agarose was dissolved in water (1/3 dH_2O + 2/3 tap water) and poured into 9 cm petri dishes. Before the agarose had solidified, a plastic mould was put on top to form troughs in the agarose. Finally, the mould was removed and the troughs were used to align and orient the embryos for injection.

Gel loading buffer (6x)

15% (w/v) ficoll (type 400, Pharmacia)
0.05% (w/v) bromphenol blue
0.05% (w/v) xylene cyanol FF

9.2 Enzymes and Standards

- Shrimp alkaline phosphatase, Roche
- Klenow fragment of DNA polymerase I, Roche
- LaTaq DNA polymerase, TaKaRa Biomedicals
- T4 DNA ligase, Roche
- Restriction enzymes, Roche or New England Biolabs (NEB)
- GeneRuler 100 bp DNA ladder, ready-to-use, MBI Fermentas
- 1 kb DNA ladder, Stratagene

9.3 Kits

- QIAquick Gel Extraction Kit, QIAgen
- QIAquick PCR Purification Kit, QIAgen
- QIAfilter Plasmid Maxi Kit, QIAgen
- QiaPrep Spin Miniprep Kit, QIAgen
- RNeasy Mini Kit, QIAgen
- Ambion mMessage machine SP6 Kit, Ambion
- Ambion mMessage machine T7 Kit, Ambion

- Topo TA Cloning Kit, Invitrogen
- Gene Images Alkphos Direct Labelling and Detection System, CDP-Star, Amersham
- Megaprime DNA labelling systems, Amersham
- Rapid hyb buffer, Amersham

9.4 Chemicals

All chemicals not listed were supplied by Sigma-Aldrich or Merck.

- Agarose, ultraPure, GibcoBRL
- Ampicillin, Sigma-Aldrich
- BSA, New England Biolabs
- dNTPs, Stratagene
- (α^{32}-p)-dCTP 3000 Ci/mmol, Amersham
- Phenol equilibrated, stabilised/Chloroform/Isoamylalcohol (25:24:1), AppliChem
- X-gal, BTS Biotech Trade & Service
- IPTG, Roche
- Polyethylene glycol (PEG) 8000, Merck

9.5 Bacteria

- DH10B, Stratagene
- XL1-Blue, Stratagene
- TOP10F', Invitrogen

9.6 Vectors

These vectors were used as staring material for DNA cloning or templates for *in vitro* transcription (numbers in brackets (X) designate the lab-internal plasmid stock number, if existent).

pCS2+ (221) as *in vitro* transcription vector (constructed by D. Turner and R. Rupp, 1993)

pBSKS+ (22) cloning vector, Stratagene

pCRII-Topo (from Topo TA Cloning Kit, Invitrogen)

pCSKA *GFP* Wuerzb. (294) contains humanised *GFP* under control of the *Xenopus borealis* cytoskeletal-actin promoter with a SV40 polyadenylation signal (kind gift of M. Schartl)

***pECFP*-N1**, Clontech. *ECFP* bears 6 aa substitutions (compared to wild type *GFP*) to shift excitation/emission spectrum, with the emission maximum in blue, to enhance brightness and solubility. More than 190 silent mutations are introduced to adjust the codon usage to preferred human codons

pBS□700 *Sleeping Beauty* (380) (kind gift of Z. Ivics) containing a left and a right *Sleeping Beauty* IR

pBSSK/*SB10* (362) (kind gift of Z. Ivics) *in vitro* transcription vector for *SB* transposase (kind gift of Z. Ivics)

pzHSP70/4prom (574) containing a 1.5 kb fragment of the zebrafish HSP70 promoter (kind gift of J. Warren)

pCGGal (731) a GAL4/VP16 fusion construct is driven by the cska promoter and followed by a SV40 pA signal

***ISceI*-pBSII SK+** (765) pBSII SK+ backbone with the MCS flanked by two inverted *I-SceI* recognition sites

pCG3.0C (739) contains a *UAS-CFP-SV40* pA cassette flanked by a left and right *SB* IR/DR

9.7 Equipment

- Incubators
  - Gallenkamp Plus II (for 50 ˚C, 55 ˚C, 65 ˚C)
  - Heraeus, Karlsruhe (for embryo breeding at 28 ˚C)
  - Hybaid Micro-4 spinning wheel incubator, MWG Biotech
- Hybridisation oven BFED 53, Fischer, Schwerte, Germany
- Microcentrifuge Eppendorf centrifuge 5417C
  Gel electrophoresis chamber Hoefer HE Mini Submarine electrophoresis unit, Pharmacia Biotech, USA
- PCR Multicycler PTC 200, Biozym
- Electroporator: BioRad GenePulser II
- Electroporation cuvettes: BioRad Gene Pulser cuvette 0.1cm, BioRad
- UV crosslinker: UV Stratalinker 2400
- Needle puller P-30, Sutter Instrument Co, USA
- Microinjector Eppendorf 5242
- Micromanipulator Leica
- Stereo microscopes
  - Stemi 2000, Leica, Wetzlar
  - MZ FLIII fluorescence stereomicroscope, Leica
  - with a 370 nm to 420 nm excitation filter and a 455 nm LP emission filter and a Jenoptik ProgRes C14 UV Camera

  Both with transillumination by a Schott/Leica KL 1500 electronic cold light source

9.8 Additional Materials

- Nylon membranes: GeneScreen Plus Hybridization Transfer Membrane NEF 976, NEN Life Science Products; Hybond N+, Amersham
- Nitrocellulose filters (for dialysis) 0.025µm, white, 13 mm, millipore
- Borosilicate glass capillaries with filament GC-100F10, Clark Electromedical Instruments
- Films Kodak BioMax Light-1, Eastman Kodak Co
- Films Kodak BioMax MR-1, Eastman Kodak Co
- Whatmann 3MM chromatography paper
- Vacutainer SST II Plus, BD Bioscience

9.9 Medaka Stocks

Wild type medaka (*Oryzias latipes*) from a closed stock at EMBL-Heidelberg were kept as described (Köster et al., 1997).

10. Methods

10.1 Isolation of Genomic DNA from Adult Fish

Adult fish were anaesthetised in ice water for at least 10 min and ground up in a pre-cooled mortar filled with liquid nitrogen. The resulting powder was suspended in 5 ml TEN 9 buffer. 5 mg proteinase K and 250 µl 20 % SDS, were added for lysis, and the suspension was thoroughly mixed by pipetting up and down with a 10 ml glass pipette. The sample was poured into 15 ml falcon tubes and incubated overnight at 50 °C in a spinning wheel incubator. Proteins were removed by phenol / chloroform extraction. For this, the solution was cooled down to RT, transferred to a vacutainer, supplemented with 3 ml phenol / chloroform / isoamylalcohol (25:24:1) and mixed by

shaking. To allow the phases to separate, the solution was left for 30 min and subsequently centrifuged for 1 h at 3000 rpm and rt. The upper, aqueous phase, now separated from the lower organic phase by the wax-layer of the vacutainer, was transferred to a fresh vacutainer, mixed with another 3 ml of phenol / chloroform / isoamylalcohol and centrifuged for 20 min at 3000 rpm and rt. The upper, aqueous phase, containing the purified genomic DNA was transferred to a new 15 ml Falcon tube, and the DNA was precipitated by adding 0.6 vol. (3 ml) of isopropanol. The DNA became visible as filaments and could be collected with a metal hook and transferred to an Eppendorf tube filled with 500 µl 70 % ethanol to be washed by gentle mixing. After sedimentation in a centrifuge for 5 min at 4000 rpm, removing of the ethanol supernatant, and air-drying, the DNA pellet was resuspended in 400 µl TE buffer (pH 8.0) at 40 ˚C overnight.

10.2 Southern Blot Hybridisation

20 µg of genomic DNA were digested o/n with 100 Units of restriction enzyme and an aliquot of 8-10 µg was separated by gel electrophoresis on a 0.8 % agarose gel at 65 V for 5 h. (~ 2.5 V/cm electrode distance). After staining with ethidium bromide for 20 min the gel was examined under UV illumination to ensure proper separation. The gel was rinsed with water to remove ethidium bromide, incubated in 300 ml 0.2 N HCl for 10 min with gentle agitation to depurinate the DNA by acidic hydrolysis for better transfer. The HCl was removed and the gel was rinsed several times. The gel was incubated in 250 ml denaturing solution for two times 10 min each with gentle agitation. The denaturing solution was removed, and the gel was incubated for 15 min in 250 ml neutralising solution. The gel was placed upside down on a wick of Whatmann 3MM paper on a glass plate above a reservoir of 20 x SSC buffer, with the ends of the wick hanging into this reservoir. A nylon membrane of a size 3 mm less than the gel in length and width was equilibrated for 1 min in water and for 10 min in 20 x SSC and placed on top of the gel. A stack of Whatmann paper, again 7 mm less than the nylon membrane in length and width was put on top of the membrane, with the lowest sheet of Whatmann paper wet in SSC before. Up most a 3 cm stack of paper

towels and a glass plate was put to pin down the setup, and the reservoir was covered with plastic wrap to prevent evaporation during the transfer. Any air bubble trapped between the glass plate, the lower Whatmann paper, the gel, the nylon membrane, and the wet upper Whatmann paper had to be removed to ensure complete and even transfer. With this setup, the transfer of DNA fragments to the nylon membrane was driven o/n by capillary forces. After the transfer was completed, the nylon membrane was marked on the upper side with the positions of the gel slots and dried on a Whatmann paper. The DNA fragments were crosslinked to the membrane by UV exposure (UV Stratalinker 2400, autocrosslink).

Probe labelling was performed using the Megaprime DNA labelling system (Amersham) to obtain a final concentration of 2-5 ng per ml hybridisation buffer according to the manufacturers protocol.

Hybridisation was performed using the Rapid-hyb buffer according the manufacturers protocol for high stringency hybridisations.

Autoradiography was performed for 2 h to several days at –80 degree Celsius using Kodak BioMax MR-1 films.

10.3 Sequencing

Sequencing was performed by the EMBL sequencing core facility.

10.4 Microinjections

10.4.1 Meganuclease

Medaka embryos and adults of the inbred Cab strain were used in all experiments. Fertilized eggs were collected immediately after spawning (at the onset of light) and placed in pre-chilled Yamamoto's embryo rearing medium (Yamamoto, 1975). For injection, one-cell stage embryos were transferred to 4 °C to arrest development. In all experiments, a pressure injector (Eppendorf 5242) was used with

borosilicate glass capillaries (GC100T(F), Clark Electromedical Instruments). Capillaries were backfilled with the injection solution (DNA: 10 ng/μl; commercial meganuclease buffer (Roche or New England Biolabs Buffer): 0,5x; meganuclease *I-SceI*: 0.2 units/μl). DNA was prepared using a Qiagen Maxiprep kit. DNA was injected through the chorion into the cytoplasm of the one-cell stage embryos. Embryos were raised to sexual maturity and outcrossing to wild type fish identified transgenic carriers. Rates of germ line transmission of identified transgenic founder fish were then established to determine the percentage of transgenic F1 offspring.

10.4.2 Sleeping Beauty

Capillaries were backfilled with the injection solution (DNA: 50-100 ng/μl; Yamamoto buffer: 1x; *SB10* mRNA: 100 ng/μl). To test for functional endogenous *SB* recognition sequences (IR/DRs) *SB10* mRNA was injected in concentrations up to 400 ng/ μl. DNA was prepared using a Qiagen Maxiprep kit and dialysed using nitrocellulose filters. DNA was injected through the chorion into the cytoplasm of the one-cell stage embryos. Embryos were raised to sexual maturity; transgenic carriers were identified by out crossing to wild type fish.

10.5 Epifluorescence Microscopy

Embryos were observed and scored using a MZFLIII dissecting microscope (Leica) with a 370 nm to 420 nm excitation filter and a 455 nm LP emission filter.

10.6 DNA cloning

DNA digestions, Klenow reactions, ligations and dephosphorylations were performed as described (Koester 1998). Fragments were purified using the QIAgen Nucleotide Removal Kit or the QIAgen QiaQuick Gel Extraction Kit if fragments were separated by electrophoresis. PCR products were cloned using the TA cloning kit from

Invitrogen. Ligations were transformed by electroporation as described (Dower et al., 1988; Köster et al., 1997). All kits were used according to the manufacturers protocol.

10.6.1 Cloning of #381 SB Reporter Vector

A 1.9 kb cassette containing the cytoskeletal actin promoter of *X. borealis*, a humanised version of *GFP* and a SV40 pA signal was removed from plasmid #294 by an *ApaI/Ecl136II* double digest. The fragment was purified and cloned into plasmid #380, resulting in the *SB* Reporter vector #381 (5.6 kb) (Henrich, 1999).

10.6.2 Cloning of the Meganuclease Vectors

The p□act-GFPI2 (7.9 kb) was generated by introducing two *I-SceI* recognition sequences in a plasmid bearing the eGFP cDNA reporter gene driven by a zebrafish □-actin muscle specific promoter and a BGH pA signal (p-G-BS, gift from Dr S.I. Higashijima (Higashijima et al., 1997). «Megalinkers» were generated by annealing complementary oligonucleotides containing the *I-SceI* recognition site (TAGGGATAACAGGGTAAT) flanked by free ends compatible with either of the *EcoRI* or *KpnI* digest products. «□Megalinkers□ were inserted at the *EcoRI* and *KpnI* sites, located at both ends of the □-actin/GFP/SV40polyA cassette in the pBluescript polylinker, and verified by sequencing. A construct with a single *I-SceI* linker at the former *KpnI* site was digested by *I-SceI* to generate a linearised control for stable transgenesis experiments.

Several other constructs were obtained by inserting different linkers at the *KpnI* site: p□act-GFPI with only one *I-SceI* recognition site, p□act-GFPDI with a shortened recognition site (GGGTAATATA), and p□act-GFPMI containing a mutated (TAGGGtTAACAGGGTAAT) version of the *I-SceI* site. The *I-SceI* meganuclease binds these latter two sites but does not cleave (Thermes* et al., 2002).

Similarly, the pCSKAGFPS-I vector (7.7 kb) was constructed. An *I-SceI* backbone vector was created by insertion of a double strand oligonucleotide containing two *I-SceI* sites interrupted by the pBSIISK+ MCS at the *BssHII* sites of pBSIISK+ and

verified by sequencing. A reporter cassette containing a *GFP* reporter gene driven by the cytoskeletal actin promoter of *X. borealis* and followed by a SV40 pA signal was inserted into the *I-SceI* backbone vector at the *Bsp120I* site resulting in a 7.7 kb plasmid.

10.6.3 Cloning of the heat-shock inducible GAL4/VP16 vector (pCG 6.0Sce)

Vector #739 was linearised by *ApaI*/*XbaI* digestion and ligated to the cska-*GAL4*/*VP16*-SV40 pA cassette that was isolated from vector #731 with the same enzymes. A 3' truncated HSP70 promoter fragment from zebrafish that was isolated by digestion with *XbaI*/*EcoRI* from vector #743 replaced the cska promoter, which was removed by the same enzymes. Finally, a single *I-SceI* recognition site was introduced between the outer edges of the *SB* IR/DRs. The resulting vector pCG6.0Sce (801) contains a HSP70 promoter fragment driving *GAL4*/*VP16* to activate transcription of *CFP* by binding to the *UAS* elements.

10.6.4 Cloning of the heat-shock inducible SB vector (pzHSPSBGFPS-I)

The 1.5 kb HSP70 promoter from zebrafish was isolated from plasmid #574 by a *SmaI*/*ClaI* double digest. Vector #362 was linearised with *XhoI*, filled in and digested with *ClaI*. The purified promoter fragment was then ligated to to #362 resulting in the vector pzHSPSB (743). In this vector *SB* is driven by the HSP70 promoter and followed by the globin 3' UTR of *X. laevis*. This vector was linearised with *Asp718I*, filled in and digested with *NotI*/*ScaI*, a 2.9 kb fragment containing the above mentioned cassette was isolated.

Vector #294 was linearised with *KspI*, filled in and digested with *NotI*, the linearised vector was purified and ligated to the above cassette resulting in the vector pCGSBHSPGFP. This vector contains the above-mentioned cassette and another cassette in opposite orientation containing *GFP* driven by the cska promoter and followed by a SV40 pA. Both cassettes were isolated by digestion with *Bsp120I* (4.8

kb) and ligated to vector #765 that was linearised with *Bsp120I* resulting in the vector pzHSPSBGFPS-I (769) containing the above-mentioned cassettes flanked by two inverted *I-SceI* recognition sites.

10.7 Isolation of Flanking Genomic Sequences

Genomic regions flanking the insertions were isolated by splinkerette PCR as described (Devon et al., 1995; Henrich, 1999). In brief, genomic DNA from transgenic lines was digested with *XhoI*. Nested PCR was performed (primary PCR: spl/ left-IR/DR and spl/ right-IR/DR primers, secondary PCR spl-nest/left-IR/DR-nest and spl-nest/right-IR/DR-nest primers. 1 ml was transferred from primary to secondary PCR) 95 °C 30 sec.; 95 °C 15 sec., 71 °C 1 min –2 °C per cycle, 72 °C 2 min (5 cycles); 95 °C 15 sec., 61 °C 2 min, 72 °C 2 min + 9 sec per cycle (28 cycles)

Spl: cgaatcgtaaccgttcgtacgagaa, spl-nest: tcgtacgagaatcgctgtcctctcc, left-IR/DR: tttactcggattaaatgtcaggaattg, left-IR/DR-nest: tgagtttaaatgtatttggctaaggtg, right-IR/DR: agtgtatgtaaacttctgacccactgg, right-IR/DR-nest: cttgtgtcatgcacaaagtagatgtcc.

Accession number of genomic insertion in SV line: AJ404849

10.8 Isolation of Total RNA

The isolation of DNA free total RNA was performed as described (Chomczynski and Sacchi, 1987; Köster et al., 1997).

10.9 Transcription of mRNA *In Vitro*

SB mRNA was generated *in vitro* using the Ambion mMessage machine (SP6) according to the manufacturers protocol. The mRNA was subsequently purified using the RNeasy RNA purification kit from QIAgen according to the manufacturers protocol.

10.10 Reverse Transcription – PCR

cDNA was generated by reverse transcription using of total RNA using the Superscript II RNAse H- reverse transcriptase from GibcoBRL. Subsequent PCR was performed according to the manufacturers recommendation using 10 % of the RT reaction.

11. Supplementary Information

Two transgenic medaka lines have been analysed in more detail using confocal microscopy, time-lapse and 3D-reconstruction. QuickTime Movies of these analyses are provided on the supplementary CD. The folder “634 movies” contains three movies that show transgenic line 634, generated using the meganuclease approach (see also figure 9). Movie “634_28_head” shows transgenic line 634. Starting point is a dorsal view of the head at developmental stage 28. Anterior is to the left (in all of the movies). 3D-rendered structures include single cells of the retinal-pigmented epithelium (RPE), the lens of the left eye and the developing diencephalon. Posterior, *GFP* is expressed in the rhombomeres. Movie “634_31_diencephalon” shows a close-up surface rendering of the *GFP* positive region of the diencephalon at developmental stage 31. *GFP* domains include the left and right *habenulae* and the epiphysis that is located between the *habenulae*. The third movie “634_31_rhombomeres” is focussed on rhombomeres 1

and 3. Anterior, the ventral part of the *cerebellum* is *GFP* positive (rhombomere 1) while the second rhombomere does not show any *GFP* signal, the complete rhombomere 3 is positive for *GFP* expression. The folder 'DE movies" contains five movies that show the development of *GFP* positive tissues of transgenic line DE, generated by transposon injection. Movies "DE_23" to "DE_30" outline the development of diencephalic tissue at the respective stages 23 to 30. Starting point is again a dorsal view of the head region, anterior is to the left. In this line the *GFP* positive tissue clearly differs from that in line 634, although both lines show *GFP* in diencephalic areas. For a detailed description of the structures that are visible in these movies, please refer to figures 15 and 16. Movie "DE_30_lateral" shows the same embryo as in "DE_30" but in lateral rotation, starting from a frontal view. In addition, the folder "Thesis" contains the complete thesis in PDF format.

References

Aizawa, K., Shimada, A., Naruse, K., Mitani, H. and Shima, A. (2003) The medaka midblastula transition as revealed by the expression of the paternal genome. Gene Expr. Patterns 3, 43-7.

Allen, N.D., Cran, D.G., Barton, S.C., Hettle, S., Reik, W. and Surani, M.A. (1988) Transgenes as probes for active chromosomal domains in mouse development. Nature. 333, 852-5.

Allende, M.L., Amsterdam, A., Becker, T., Kawakami, K., Gaiano, N. and Hopkins, N. (1996) Insertional mutagenesis in zebrafish identifies two novel genes, pescadillo and dead eye, essential for embryonic development. Genes & Development. 10, 3141-55.

Amsterdam, A., Yoon, C., Allende, M., Becker, T., Kawakami, K., Burgess, S., Gaiano, N. and Hopkins, N. (1997) Retrovirus-mediated insertional mutagenesis in zebrafish and identification of a molecular marker for embryonic germ cells. Cold Spring Harbor Symposia on Quantitative Biology. 62, 437-50.

Andeol, Y. (1994) Early transcription in different animal species: implication for transition from maternal to zygotic control in development. Roux's Arch. Dev. Biol. 171, 3-10.

Andersson, G., Svensson, A.C., Setterblad, N. and Rask, L. (1998) Retroelements in the human MHC class II region. Trends in Genetics 14, 109-14.

Anken, R. and Bourrat, F. (1998) Brain atlas of the Medakafish *Oryzias latipes*. Institut national de la recherche agronomique, Paris.

Antoch, M.P., Song, E.J., Chang, A.M., Vitaterna, M.H., Zhao, Y., Wilsbacher, L.D., Sangoram, A.M., King, D.P., Pinto, L.H. and Takahashi, J.S. (1997) Functional identification of the mouse circadian Clock gene by transgenic BAC rescue. Cell 89, 655-67.

Balague, C., Kalla, M. and Zhang, W.W. (1997) Adeno-associated virus Rep78 protein and terminal repeats enhance integration of DNA sequences into the cellular genome. Journal of Virology. 71, 3299-306.

Bayer, T.A. and Campos-Ortega, J.A. (1992) A Transgene Containing lacZ Is Expressed in Primary Sensory Neurons in Zebrafish. Development 115, 421-426.

Belfort, M. and Roberts, R.J. (1997) Homing endonucleases: keeping the house in order. Nucleic Acids Research 25, 3379-88.

Bellen, H.J., O'Kane, C.J., Wilson, C., Grossniklaus, U., Pearson, R.K. and Gehring, W.J. (1989) P-element-mediated enhancer detection: a versatile method to study development in Drosophila. Genes Dev 3, 1288-300.

Bellen, H.J., Wilson, C., Gibson, G., Grossniklaus, U., Pearson, R.K., O'Kane, C. and Gehring, W.J. (1990) P-element-mediated enhancer detection allows rapid identification of developmentally regulated genes and cell specific markers in Drosophila. J Physiol (Paris) 84, 33-41.

Beylot, B. and Spassky, A. (2001) Chemical probing shows that the intron-encoded endonuclease I-SceI distorts DNA through binding in monomeric form to its homing site. Journal of Biological Chemistry 276, 25243-53.

Bishop, J.O. and Smith, P. (1989) Mechanism of chromosomal integration of microinjected DNA. Molecular Biology & Medicine 6, 283-98.

Boulikas, T. (1995) Chromatin domains and prediction of MAR sequences. Int Rev Cytol 162A, 279-388.

Brand, A.H. and Perrimon, N. (1993) Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118, 401-15.

Brand Andrea, H., Dormand, E.-L. and Brand, A. (1995) The GAL4 system as a tool for unravelling the mysteries of the Drosophila nervous system
GFP in Drosophila. Current Opinion in Neurobiology 5, 572-578.

Brinster, R.L., Braun, R.E., Lo, D., Avarbock, M.R., Oram, F. and Palmiter, R.D. (1989) Targeted correction of a major histocompatibility class II E alpha gene by DNA microinjected into mouse eggs. Proceedings of the National Academy of Sciences of the United States of America 86, 7087-91.

Bronchain, O.J., Hartley, K.O. and Amaya, E. (1999) A gene trap approach in Xenopus. Current Biology. 9, 1195-8.

Burns, J.C., Friedmann, T., Driever, W., Burrascano, M. and Yee, J.K. (1993) Vesicular stomatitis virus G glycoprotein pseudotyped retroviralvectors: concentration to very high titer and efficient gene transferinto mammalian and nonmammalian cells. Proc Natl Acad Sci U S, A. 90, 8033-7.

Calvi, B.R., Hong, T.J., Findley, S.D. and Gelbart, W.M. (1991) Evidence for a common evolutionary origin of inverted repeat transposons in Drosophila and plants: hobo, Activator, and Tam3. Cell 66, 465-71.

Capecchi, M.R. (1989) The new mouse genetics: altering the genome by gene targeting. Trends in Genetics 5, 70-76.

Cheng, L., Ziegelhoffer, P.R. and Yang, N.S. (1993) In vivo promoter activity and transgene expression in mammalian somatic tissues evaluated by using particle bombardment. Proceedings of the National Academy of Sciences of the United States of America 90, 4455-9.

Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Analytical Biochemistry 162, 156-9.

Chong, S.S.C. and Vielkind, J.R. (1989) Expression and fate of CAT reporter gene microinjected into fertilised medaka (Oryzias latipes) eggs in the form of plasmid DNA. Theor. Appl. Genet. 78, 369-80.

Chou, C.Y., Horng, L.S. and Tsai, H.J. (2001) Uniform GFP-expression in transgenic medaka (Oryzias latipes) at the F0 generation. Transgenic Research. 10, 303-15.

Chourrout, D. and Perrot, E. (1992) No transgenic rainbow trout produced with sperm incubated with linear DNA. Molecular Marine Biology & Biotechnology 1, 282-5.

Clark, A.J., Archibald, A.L., McClenaghan, M., Simons, J.P., Wallace, R. and Whitelaw, C.B. (1993) Enhancing the efficiency of transgene expression. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences 339, 225-32.

Collas, P. and Alestrom, P. (1997) Nuclear localization signals: a driving force for nuclear transport of plasmid DNA in zebrafish. Biochemistry & Cell Biology 75, 633-40.

Collas, P. and Alestrom, P. (1998) Nuclear localization signals enhance germline transmission of a transgene in zebrafish. Transgenic Research 7, 303-9.

Colleaux, L., D'Auriol, L., Galibert, F. and Dujon, B. (1988) Recognition and cleavage site of the intron-encoded omega transposase. Proceedings of the National Academy of Sciences of the United States of America 85, 6022-6.

Condie, B.G., Brivanlou, A.H. and Harland, R.M. (1990) Most of the homeobox-containing Xhox 36 transcripts in early Xenopus embryos cannot encode a homeodomain protein. Molecular & Cellular Biology 10, 3376-3385.

Cooley, L., Berg, C. and Spradling, A. (1988) Controlling P element insertional mutagenesis. Trends in Genetics 4, 254-258.

Culp, P., Nusslein-Volhard, C. and Hopkins, N. (1991) High-frequency germ-line transmission of plasmid DNA sequences injectedinto fertilized zebrafish eggs. Proceedings of the National Academy of Sciences of the United States of America 88, 7953-7.

Devon, R.S., Porteous, D.J. and Brookes, A.J. (1995) Splinkerettes-improved vectorettes for greater efficiency in PCR walking. Nucleic Acids Res 23, 1644-5.

Dorer, D.R. and Henikoff, S. (1002) Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophila. Cell 77, 993-1002.

Dower, W.J., Miller, J.F. and Ragsdale, C.W. (1988) High efficiency transformation of E. coli by high voltage electroporation. Nucleic Acids Research 16, 6127-6145.

Dupuy, A.J., Clark, K., Carlson, C.M., Fritz, S., Davidson, A.E., Markley, K.M., Finley, K., Fletcher, C.F., Ekker, S.C., Hackett, P.B., Horn, S. and Largaespada, D.A. (2002) Mammalian germ-line transgenesis by transposition. Proceedings of the National Academy of Sciences of the United States of America 99, 4495-9.

Dupuy, A.J., Fritz, S. and Largaespada, D.A. (2001) Transposition and gene disruption in the male germline of the mouse. Genesis: the Journal of Genetics & Development. 30, 82-8.

Eden, S. and Cedar, H. (1994) Role of DNA methylation in the regulation of transcription. Current Opinion in Genetics & Development 4, 255-9.

Elgin, S.C. (1990) Chromatin structure and gene activity. Current Opinion in Cell Biology 2, 437-45.

Fischer, S.E., Wienholds, E. and Plasterk, R.H. (2001) Regulated transposition of a fish transposon in the mouse germ line. Proceedings of the National Academy of Sciences of the United States of America. 98, 6759-64.

Friedberg, E.C., Walker, G.C. and Siede, W. (1995) DNA repair and mutagenesis. American Society of Microbiology Press, Washington, DC.

Fritz, J.D., Herweijer, H., Zhang, G. and Wolff, J.A. (1996) Gene transfer into mammalian cells using histone-condensed plasmid DNA. Human Gene Therapy 7, 1395-404.

Fu, Y., Wang, Y. and Evans, S.M. (1998) Viral sequences enable efficient and tissue-specific expression of transgenes in Xenopus. Nature Biotechnology 16, 253-257.

Fukamachi, S., Shimada, A. and Shima, A. (2001) Mutations in the gene encoding B, a novel transporter protein, reduce melanin content in medaka. Nature Genetics 28, 381-385.

Gaiano, N., Allende, M., Amsterdam, A., Kawakami, K. and Hopkins, N. (1996a) Highly efficient germ-line transmission of proviral insertions in zebrafish. Proceedings of the National Academy of Sciences of the United States of America. 93, 7777-82.

Gaiano, N., Amsterdam, A., Kawakami, K., Allende, M., Becker, T. and Hopkins, N. (1996b) Insertional mutagenesis and rapid cloning of essential genes in zebrafish. Nature. 383, 829-32.

Garrick, D., Fiering, S., Martin, D.I.K. and Whitelaw, E. (1998) Repeat-induced gene silencing in mammals. Nature Genetics 18, 56-59.

Gibbs, P.D., Gray, A. and Thorgaard, G. (1994a) Inheritance of P element and reporter gene sequences in zebrafish. Molecular Marine Biology & Biotechnology 3, 317-26.

Gibbs, P.D., Peek, A. and Thorgaard, G. (1994b) An in vivo screen for the luciferase transgene in zebrafish. Molecular Marine Biology & Biotechnology 3, 307-16.

Golling, G., Amsterdam, A., Sun, Z., Antonelli, M., Maldonado, E., Chen, W., Burgess, S., Haldi, M., Artzt, K., Farrington, S., Lin, S.Y., Nissen, R.M. and Hopkins, N. (2002) Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. Nature Genetics 31, 135-40.

Gong, W.J. and Golic, K.G. (2003) Ends-out, or replacement, gene targeting in Drosophila. Proceedings of the National Academy of Sciences of the United States of America in press.

Gordon, J.W., Scangos, G.A., Plotkin, D.J., Barbosa, J.A. and Ruddle, F.H. (1980) Genetic transformation of mouse embryos by microinjection of purified DNA. Proceedings of the National Academy of Sciences of the United States of America 77, 7380-4.

Gossler, A., Joyner, A.L., Rossant, J. and Skarnes, W.C. (1989) Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. Science. 244, 463-5.

Grabher, C., Henrich, T., Sasado, T., Arenz, A., Furutani-Seiki, M. and Wittbrodt, J. (2002) High Frequency Generation of Specific GFP Patterns in Medaka. Nature Biotechnology submitted.

Gueiros-Filho, F.J. and Beverley, S.M. (1997) Trans-kingdom transposition of the Drosophila element mariner within the protozoan Leishmania. Science 276, 1716-1719.

Hackett, P.B. (1993) The molecular biology of transgenic fish. In Mommsen, H.a. (ed.), Biochemistry and molecular biology of fishes, Elseviers Science Publishers, Vol. 207-240.

Halloran, M.C., Sato-Maeda, M., Warren, J.T., Su, F., Lele, Z., Krone, P.H., Kuwada, J.Y. and Shoji, W. (2000) Laser-induced gene expression in specific cells of transgenic zebrafish. Development 127, 1953-1960.

Handler, A.M., Gomez, S.P. and O'Brochta, D.A. (1993) A functional analysis of the P-element gene-transfer vector in insects. Archives of Insect Biochemistry & Physiology 22, 373-84.

Henrich, T. (1999) Enhancer trap in medaka using the reconstructed transposon "Sleeping Beauty". Ruprecht-Karls-Universität Heidelberg, Heidelberg, pp. 106.

Higashijima, S., Okamoto, H., Ueno, N., Hotta, Y. and Eguchi, G. (1997) High-frequency generation of transgenic zebrafish which reliably express GFP in whole muscles or the whole body by using promoters of zebrafish origin. Developmental Biology. 192, 289-99.

Hong, Y., Winkler, C. and Schartl, M. (1996) Pluripotency and differentiation of embryonic stem cell lines from the medakafish (Oryzias latipes). Mechanisms of Development 60, 33-44.

Hong, Y., Winkler, C. and Schartl, M. (1998) Production of medakafish chimeras from a stable embryonic stem cell line. Proceedings of the National Academy of Sciences USA 95, 3679-84.

Horie, K., Kuroiwa, A., Ikawa, M., Okabe, M., Kondoh, G., Matsuda, Y. and Takeda, J. (2001) Efficient chromosomal transposition of a Tc1/mariner- like transposon Sleeping Beauty in mice. Proceedings of the National Academy of Sciences of the United States of America. 98, 9191-6.

Houdebine, L.M. and Chourrout, D. (1991) Transgenesis in fish. Experientia 47, 891-7.

Hsiao, C.D., Hsieh, F.J. and Tsai, H.J. (2001) Enhanced expression and stable transmission of transgenes flanked by inverted terminal repeats from adeno-associated virus in zebrafish. Developmental Dynamics 220, 323-36.

Hyodo-Taguchi, Y. and Egami, N. (1985) Establishment of inbred strains of the medaka Oryzias latipes and the usefulness of the strains for biomedical research. Zoological Science 2, 305-316.

Inoue, K., Yamashita, S., Hata, J., Kabeno, S., Asada, S., Nagahisa, E. and Fujita, T. (1990) Electroporation as a new technique for producing transgenic fish. Cell Differ Dev 29, 123-8.

Ishikawa, Y. (1997) Embryonic development of the medaka brain. The Fish Biology Journal MEDAKA 9, 17-31.

Ishikawa, Y. (2000) Medakafish as a model system for vertebrate developmental genetics. Bioessays 22, 487-495.

Ivics, Z., Hackett, P.B., Plasterk, R.H. and Izsvak, Z. (1997) Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell 91, 501-10.

Ivics, Z., Izsvak, Z. and Hackett, P.B. (1999) Genetic applications of transposons and other repetitive elements in zebrafish. Methods in Cell Biology 60, 99-131.

Iwamatsu, T. (1994) Stages of normal development in the medaka Oryzias latipes. Zoo. Sci. 11, 825-839.

Iyengar, A. and Maclean, N. (1995) Transgene concatemerisation and expression in rainbow trout (Oncorhynchus mykiss). Mol. Mar. Biol. Biotechnol. 4, 248-54.

Iyengar, A., Muller, F. and Maclean, N. (1996) Regulation and expression of transgenes in fish -- a review. Transgenic Research 5, 147-166.

Jacquier, A. and Dujon, B. (1985) An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell 41, 383-94.

Jaenisch, R. (1988) Transgenic animals. Science. 240, 1468-74.

Johnson, R.D. and Jasin, M. (2001) Double-strand-break-induced homologous recombination in mammalian cells. Biochem Soc Trans 29, 196-201.

Jowett, T. (1999) Transgenic zebrafish. Methods in Molecular Biology 97, 461-86.

Joyner, A.L., Herrup, K., Auerbach, B.A., Davis, C.A. and Rossant, J. (1991) Subtle cerebellar phenotype in mice homozygous for a targeted deletion of the En-2 homeobox. Science 251, 1239-1243.

Kaneda, Y., Iwai, K. and Uchida, T. (1989) Increased expression of DNA cointroduced with nuclear protein in adult rat liver. Science 243, 375-8.

Kawakami, K., Shima, A. and Kawakami, N. (2000) Identification of a functional transposase of the Tol2 element, an Ac-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage. Proceedings of the National Academy of Sciences of the United States of America 97, 11403-11408.

Kellum, R. and Schedl, P. (1991) A position-effect assay for boundaries of higher order chromosomal domains. Cell 64, 941-50.

Kellum, R. and Schedl, P. (1992) A group of scs elements function as domain boundaries in an enhancer-blocking assay. Molecular & Cellular Biology 12, 2424-31.

Kimmel, C.B. (1993) Patterning the brain of the zebrafish embryo. Annu Rev Neurosci 16, 707-32.

Klein, T.M. and Fitzpatrick-McElligott, S. (1993) Particle bombardment: a universal approach for gene transfer to cells and tissues. Current Opinion in Biotechnology 4, 583-590.

Koga, A., Inagaki, H., Bessho, Y. and Hori, H. (1995) Insertion of a novel transposable element in the tyrosinase gene is responsible for an albino mutation in the medaka fish, Oryzias latipes. Molecular & General Genetics 249, 400-405.

Koga, A., Suzuki, M., Inagaki, H., Bessho, Y. and Hori, H. (1996) Transposable element in fish. Nature 383, 30.

Kondo, S., Kuwahara, Y., Kondo, M., Naruse, K., Mitani, H., Wakamatsu, Y., Ozato, K., Asakawa, S., Shimizu, N. and Shima, A. (2001) The medaka rs-3 locus required for scale development encodes ectodysplasin-A receptor. Current Biology 7, 1202-1206.

Korn, R., Schoor, M., Neuhaus, H., Henseling, U., Soininen, R., Zachgo, J. and Gossler, A. (1992) Enhancer trap integrations in mouse embryonic stem cells give rise to staining patterns in chimaeric embryos with a high frequency and detect endogenous genes. Mechanisms of Development 39, 95-109.

Köster, R., Stick, R., Loosli, F. and Wittbrodt, J. (1997) Medaka spalt acts as a target gene of hedgehog signaling. Development 124, 3147-3156.

Kroll, K.L. and Amaya, E. (1996) Transgenic Xenopus embryos from sperm nuclear transplantations reveal FGF signaling requirements during gastrulation. Development 122, 3173-83.

Kroll, K.L. and Gerhart, J.C. (1994) Transgenic X. laevis embryos from eggs transplanted with nuclei of transfected cultured cells. Science 266, 650-653.

Kuspa, A. and Loomis, W.F. (1992) Tagging developmental genes in Dictyostelium by restriction enzyme-mediated integration of plasmid DNA. Proceedings of the National Academy of Sciences of the United States of America 89, 8803-7.

Lai, L., Kolber-Simonds, D., Park, K.W., Cheong, H.T., Greenstein, J.L., Im, G.S., Samuel, M., Bonk, A., Rieke, A., Day, B.N., Murphy, C.N., Carter, D.B., Hawley, R.J. and Prather, R.S. (2002) Production of alpha-1,3-

galactosyltransferase knockout pigs by nuclear transfer cloning. Science 295, 1089-92.

Lampe, D.J., Churchill, M.E. and Robertson, H.M. (1996) A purified mariner transposase is sufficient to mediate transposition in vitro. EMBO Journal 15, 5470-5479.

Lankenau, S., Barnickel, T., Marhold, J., Lyko, F., Mechler, B.M. and Lankenau, D.-H. (2003) Knockout Trageting of the Drosophila Nap1 Gene and Examination of DNA Repair Tracts in the Recombination Products. Genetics in press.

Lavitrano, M., Camaioni, A., Fazio, V.M., Dolci, S., Farace, M.G. and Spadafora, C. (1989) Sperm cells as vectors for introducing foreign DNA into eggs: genetic transformation of mice. Cell 57, 717-723.

Lee, K.-Y., Huang, H., Ju, B., Yang, Z. and Lin, S. (2002) Cloned zebrafish by nuclear transfer from long-term-cultured cells. Nature Biotechnology 20, 795-99.

Lee, K.S., Yasumasu, S., Nomura, K. and Iuchi, I. (1994) Hce, a Constituent of the Hatching Enzymes of Oryzias Latipes Embryos, Releases Unique Proline-Rich Polypeptides from Its Natural Substrate, the Hardened Chorion. FEBS Letters 339, 281-284.

Lin, S. (2000) Transgenic zebrafish. Methods in Molecular Biology 136, 375-83.

Lin, S., Gaiano, N., Culp, P., Burns, J.C., Friedmann, T., Yee, J.K. and Hopkins, N. (1994a) Integration and germ-line transmission of a pseudotyped retroviral vector in zebrafish. Science 265, 666-669.

Lin, S., Yang, S. and Hopkins, N. (1994b) lacZ Expression in Germline Transgenic Zebrafish Can Be Detected in Living Embryos. Developmental Biology 161, 77-83.

Linney, E., Hardison, N.L., Lonze, B.E., Lyons, S. and DiNapoli, L. (1999) Transgene expression in zebrafish: A comparison of retroviral-vector and DNA-injection approaches. Developmental Biology 213, 207-216.

Loosli, F., Köster, R.W., Carl, M., P., K.R., Henrich, T., M., M., Krone, A. and Wittbrodt, J. (2000) A genetic screen for mutations affecting embryonic development in medaka fish (*Oryzias latipes*). Mechanisms of Development 97, 133-139.

Loosli, F., Winkler, S., Burgtorf, C., Wurmbach, E., Ansorge, W., Henrich, T., Grabher, C., Arendt, D., Carl, M., Krone, A., Grzebisz, E. and Wittbrodt, J. (2001) Medaka *eyeless* is the key factor linking retinal dertermination and eye growth. Development 128, 4035-4044.

Loukeris, T.G., Arca, B., Livadaras, I., Dialektaki, G. and Savakis, C. (1995) Introduction of the transposable element Minos into the germ line of Drosophila melanogaster. Proceedings of the National Academy of Sciences of the United States of America 92, 9485-9.

Lurquin, P.F. (1997) Gene transfer by electroporation. Molecular Biotechnology 7, 5-35.

Ma, C.G., Fan, L.C., Ganassin, R., Bols, N. and Collodi, P. (2001) Production of zebrafish germ-line chimeras from embryo cell cultures. Proceedings of the National Academy of Sciences USA 98, 2461-2466.

McCreath, K.J., Howcroft, J., Campbell, K.H., Colman, A., Schnieke, A.E. and Kind, A.J. (2000) Production of gene-targeted sheep by nuclear transfer from cultured somatic cells. Nature 405, 1066-9.

McKnight, R.A., Shamay, A., Sankaran, L., Wall, R.J. and Hennighausen, L. (1992) Matrix-attachment regions can impart position-independent regulation of a tissue-specific gene in transgenic mice. Proceedings of the National Academy of Sciences of the United States of America 89, 6943-7.

Mehtali, M., LeMeur, M. and Lathe, R. (1990) The methylation-free status of a housekeeping transgene is lost at high copy number. Gene 91, 179-84.

Monteilhet, C., Perrin, A., Thierry, A., Colleaux, L. and Dujon, B. (1990) Purification and characterization of the in vitro activity of I-Sce I, a novel and highly specific endonuclease encoded by a group I intron. Nucleic Acids Research 18, 1407-13.

Mueller, J.E., Bryk, M., Loizos, N. and Belfort, M. (1993) Linn, S.M. et al. (eds.), Nucleases, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 111-43.

Müller, F., Blader, P. and Strähle, U. (2002) Search for enhancers: teleost models in comparative genomic and transgenic analysis of cis regulatory elements. Bioessays 24, 564-72.

Muller, F., Ivics, Z., Erdelyi, F., Papp, T., Varadi, L., Horvath, L. and Maclean, N. (1992) Introducing foreign genes into fish eggs with electroporated sperm as a carrier. Molecular Marine Biology & Biotechnology 1, 276-81.

Muller, U. (1999) Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. Mechanisms of Development 82, 3-21.

Noma, K., Allis, C.D. and Grewal, S.I. (2001) Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. Science 293, 1150-55.

O'Kane, C.J. and Gehring, W.J. (1987) Detection in situ of genomic regulatory elements in Drosophila. Proceedings of the National Academy of Sciences USA 84, 9123-9127.

Ono, H., Hirose, E., Miyazaki, K., Yamamoto, H. and Matsumoto, J. (1997) Transgenic medaka fish bearing the mouse tyrosinase gene: expression and transmission of the transgene following electroporation of the orange-colored variant. Pigment Cell Research 10, 168-75.

Ozato, K., Kondoh, H., Inohara, H., Iwamatsu, T., Wakamatsu, Y. and Okada, T.S. (1986) Production of transgenic fish: introduction and expression of chicken delta-crystallin gene in medaka embryos. Cell Differ 19, 237-44.

Palmiter, R.D., Sandgren, E.P., Avarbock, M.R., Allen, D.D. and Brinster, R.L. (1991) Heterologous introns can enhance expression of transgenes in mice. Proceedings of the National Academy of Sciences of the United States of America 88, 478-82.

Perrin, A., Buckle, M. and Dujon, B. (1993) Asymmetrical recognition and activity of the I-SceI endonuclease on its site and on intron-exon junctions. EMBO Journal 12, 2939-47.

Philip, R., Brunette, E., Kilinski, L., Murugesh, D., McNally, M.A., Ucar, K., Rosenblatt, J., Okarma, T.B. and Lebkowski, J.S. (1994) Efficient and sustained gene expression in primary T lymphocytes and primary and cultured tumor cells mediated by adeno-associated virus plasmid DNA complexed to cationic liposomes. Molecular & Cellular Biology 14, 2411-8.

Plasterk, R.H. (1996) The Tc1/mariner transposon family. Current Topics in Microbiology & Immunology 204, 125-143.

Plasterk, R.H.A., Izsvak, Z. and Ivics, Z. (1999) Resident aliens - the Tc1/mariner superfamily of transposable elements. Trends in Genetics 15, 326-332.

Prioleau, M.N., Huet, J., Sentenac, A. and Mechali, M. (1994) Competition Between Chromatin and Transcription Complex Assembly Regulates Gene Expression During Early Development. Cell 77, 439-449.

Raz, E., van Luenen, H.G., Schaerringer, B., Plasterk, R.H.A. and Driever, W. (1998) Transposition of the nematode Caenorhabditis elegans Tc3 element in the zebrafish Danio rerio. Current Biology 8, 82-8.

Rio, D.C., Barnes, G., Laski, F.A., Rine, J. and Rubin, G.M. (1988) Evidence for Drosophila P element transposase activity in mammalian cells and yeast. Journal of Molecular Biology 200, 411-5.

Robertson, H.M., Preston, C.R., Phillis, R.W., Johnson-Schlitz, D.M., Benz, W.K. and Engels, W.R. (1988) A stable genomic source of P element transposase in Drosophila melanogaster. Genetics 118, 461-70.

Rong, Y.S. and Golic, K.G. (2000) Gene targeting by homologous recombination in Drosophila. Science 288, 2013-8.

Rong, Y.S. and Golic, K.G. (2001) A targeted gene knockout in Drosophila. Genetics 157, 1307-12.

Rong, Y.S., Titen, S.W., Xie, H.B., Golic, M.M., Bastiani, M., Bandyopadhyay, P., Olivera, B.M., Brodsky, M., Rubin, G.M. and Golic, K.G. (2002) Targeted Mutagenesis by homologous recombination in D. melanogaster. Genes & Development 16, 1568-1681.

Rouet, P., Smih, F. and Jasin, M. (1994) Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Molecular & Cellular Biology 14, 8096-106.

Rubin, G.M. and Spradling, A.C. (1982) Genetic transformation of Drosophila with transposable element vectors. Science 218, 348-353.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning, A Laboratory Manual. Cold Spring Habor Laboratory Press.

Sandmeyer, S.B., Hansen, L.J. and Chalker, D.L. (1990) Integration specificity of retrotransposons and retroviruses. Annu Rev Genet 24, 491-518.

Scheer, N. and Campos-Ortega, J.A. (1999) Use of the Gal4-UAS technique for targeted gene expression in the zebrafish. Mechanisms of Development 80, 153-158.

Sebestyen, M.G., Ludtke, J.J., Bassik, M.C., Zhang, G., Budker, V., Lukhtanov, E.A., Hagstrom, J.E. and Wolff, J.A. (1998) DNA vector chemistry: the covalent attachment of signal peptides to plasmid DNA. Nature Biotechnology 16, 80-5.

Sentry, J.W. and Kaiser, K. (1992) P element transposition and targeted manipulation of the Drosophila genome. Trends in Genetics 8, 329-331.

Shima, A., Himmelbauer, H., Mitani, H., Furutani-Seiki, M., Wittbrodt, J. and Schartl, M. (2003) Fish genomes flying. EMBO Rep. 4, 121-5.

Shima, A. and Shimada, A. (1991) Development of a possible nonmammalian test system for radiation- induced germ-cell mutagenesis using a fish, the Japanese medaka (Oryzias latipes). Proceedings of the National Academy of Sciences USA 88, 2545-2549.

Shinomiya, A., Tanaka, M., Kobayashi, T., Nagahama, Y. and Hamaguchi, S. (2000) The vasa-like gene, olvas, identifies the migration path of primordial germ cells

during embryonic body formation stage in the medaka, Oryzias latipes. Development Growth & Differentiation 42, 317-326.

Sin, F.Y., Walker, S.P., Symonds, J.E., Mukherjee, U.K., Khoo, J.G. and Sin, I.L. (2000) Electroporation of salmon sperm for gene transfer: efficiency, reliability, and fate of transgene. Molecular Reproduction & Development 56, 285-8.

Sippel, A.E., Saueressig, H., Winter, D.M., Grewal, T., Faust, N., Hecht, A. and Bonifer, C. (1992) The regulatory domain organization of eukaryotic genomes: implications for stable gene transfer. In Grosveld, F. and Kollias, G. (eds.), Transgenic animals, Academic press, pp. 1-26.

Smith, K.R. (2002) Gene transfer in higher animals: theoretical considerations and key concepts. Journal of Biotechnology 99, 1-22.

Srivastava, A., Lusby, E.W. and Berns, K.I. (1983) Nucleotide sequence and organization of the adeno-associated virus 2 genome. Journal of Virology 45, 555-64.

Stief, A., Winter, D.M., Stratling, W.H. and Sippel, A.E. (1989) A nuclear DNA attachment element mediates elevated and position-independent gene activity. Nature 341, 343-5.

Stuart, G.W., McMurray, J.V. and Westerfield, M. (1988) Replication, integration and stable germ-line transmission of foreignsequences injected into early zebrafish embryos. Development 103, 403-12.

Stuart, G.W., Vielkind, J.R., McMurray, J.V. and Westerfield, M. (1990) Stable lines of transgenic zebrafish exhibit reproducible patterns of transgene expression. Development 109, 577-84.

Sussman, R. (2001) Direct DNA delivery into zebrafish embryos employing tissue culture techniques. Genesis: the Journal of Genetics & Development 31, 1-5.

Sutherland, H.F., Lovell-Badge, R.H. and Jackson, I.J. (1993) Characterisation of two identical independent non-homologous integration sites in mouse embryonic stem cells. Gene 131, 265-8.

Szelei, J. and Duda, E. (1989) Entrapment of high-molecular-mass DNA molecules in liposomes for the genetic transformation of animal cells. Biochemical Journal 259, 549-53.

Tanaka, M. and Kinoshita, M. (2001) Recent progress in the generation of transgenic medaka (Oryzias latipes). Zoological Science 18, 615-622.

Tawk, M., Tuil, D., Torrente, Y., Vriz, S. and Paulin, D. (2002) High-efficiency gene transfer into adult fish: a new tool to study fin regeneration. Genesis: the Journal of Genetics & Development 32, 27-31.

te Riele, H., Maandag, E.R. and Berns, A. (1992) Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. Proceedings of the National Academy of Sciences of the United States of America 89, 5128-5132.

Tewari, R., Michard-Vanhee, C., Perrot, E. and Chourrout, D. (1992) Mendelian transmission, structure and expression of transgenes following their injection into the cytoplasm of trout eggs. Transgenic Research 1, 250-60.

Thermes*, V., Grabher*, C., Ristoratore*, F., Bourrat, F., Choulika, A., Wittbrodt, J. and Joly, J.-S. (2002) I-SceI meganuclease mediates highly efficient transgenesis in fish. Mechanisms of Development 118, 91-98.

Tsai, H.J., Wang, S.H., Inoue, K., Takagi, S., Kimura, M., Wakamatsu, Y. and Ozato, K. (1995) Initiation of the transgenic lacZ gene expression in medaka (Oryzias latipes) embryos. Molecular Marine Biology & Biotechnology 4, 1-9.

Tsubota, S., Ashburner, M. and Schedl, P. (1985) P-element-induced control mutations at the r gene of Drosophila melanogaster. Mol Cell Biol 5, 2567-74.

Venkatesh, B., Gilligan, P. and Brenner, S. (2000) Fugu: a compact vertebrate reference genome. FEBS Letters 476, 3-7.

Vielkind, J.R. (1992) Medaka and zebrafish: ideal as transient and stable transgenic systems. In Hew, C.L. and Fletcher, G.L. (eds.), Transgenic fish, World Scientific, pp. 72-91.

Vigdal, T.J., Kaufman, C.D., Izsvak, Z., Voytas, D.F. and Ivics, Z. (2002) Common Physical Properties of DNA Affecting Target Site Selection of Sleeping Beauty and other Tc1/mariner Transposable Elements. J. Mol. Biol. 323, 441-452.

Vogelstein, B. and Kinsler, K.W. (1998) The genetic basis of human cancer. McGraw-Hill, New York.

Volckaert, F.A., Hellemans, B.A., Galbusera, P., Ollevier, F., Sekkali, B. and Belayew, A. (1994) Replication, expression, and fate of foreign DNA during embryonic and larval development of the African catfish (Clarias gariepinus). Molecular Marine Biology & Biotechnology 3, 57-69.

Vos, J.C., De Baere, I. and Plasterk, R.H. (1996) Transposase is the only nematode protein required for in vitro transposition of Tc1. Genes & Development 10, 755-61.

Wakamatsu, Y., Ju, B.S., Pristyaznhyuk, I., Niwa, K., Ladygina, T., Kinoshita, M., Araki, K. and Ozato, K. (2001) Fertile and diploid nuclear transplants derived from embryonic cells of a small laboratory fish, medaka (Oryzias latipes). Proceedings of the National Academy of Sciences USA 98, 1071-1076.

Weitzman, M.D., Fisher, K.J. and Wilson, J.M. (1996) Recruitment of wild-type and recombinant adeno-associated virus into adenovirus replication centers. J Virol 70, 1845-54.

Westerfield, M. (1995) The zebrafish book. The University of Oregon Press, Eugene.

Westerfield, M., Wegner, J., Jegalian, B.G., DeRobertis, E.M. and Puschel, A.W. (1992) Specific activation of mammalian Hox promoters in mosaic transgenic zebrafish. Genes Dev 6, 591-8.

Williams, D.W., Muller, F., Lavender, F.L., Orban, L. and Maclean, N. (1996) High transgene activity in the yolk syncytial layer affects quantitative transient expression assays in zebrafish Danio rerio) embryos. Transgenic Research 5, 433-42.

Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J. and Campbell, K.H. (1997) Viable offspring derived from fetal and adult mammalian cells. Nature 385, 810-3.

Winkler, C., Vielkind, J.R. and Schartl, M. (1991) Transient expression of foreign DNA during embryonic and larval development of the medaka fish (Oryzias latipes). Mol Gen Genet 226, 129-40.

Wittbrodt, J., Shima, A. and Schartl, M. (2002) Medaka - A model organism from the Far East. Nature Reviews Genetics 3, 53-64.

Wu, Y., Reece Richard, J. and Ptashne, M. (1996) Quantitation of putative activator-target affinities predicts transcriptional activating potentials. Embo 15, 3951-3963.

Yamamoto, T. (1975) Medaka (Killifish), Biology and Strains. Keigaku Publishing Company, Tokyo.

Yamauchi, M., Kinoshita, M., Sasanuma, M., Tsuji, S., Terada, M., M., M. and Ishikawa, Y. (2000) Introduction of a foreign gene into medakafish using the particle gun method. J. Exp. Zool 287, 285-93.

Zelenin, A.V., Alimov, A.A., Barmintzev, V.A., Beniumov, A.O., Zelenina, I.A., Krasnov, A.M. and Kolesnikov, V.A. (1991) The delivery of foreign genes into fertilized fish eggs using high-velocity microprojectiles. FEBS Lett 287, 118-20.

Zhu, Z., Li, G., He, L. and Chen, S. (1985) Novel gene transfer into fertilized eggs of goldfish (Carassius auratus L. 1758). Z. Angew. Ichtyiol. 1, 31-34.

Zhu, Z.Y. and Sun, Y.H. (2000) Embryonic and genetic manipulation in fish. Cell Research 10, 17-27.

Acknowledgements

My appreciation goes to Jochen Wittbrodt for the opportunity to do this thesis in his lab. The enthusiasm and encouraging atmosphere he is able to spread among his fellows renders him an ideal scientific supervisor. He was always there to help and discuss despite his numerous duties. Felix Loosli, our senior postdoc always was a great help, he turned out to be a very talented and funny teacher and I am still learning. I want to thank Thorsten Henrich who introduced me to the *SB* project until he left the lab. Matthias Carl, my bench mate and second senior postdoc, became a friend who spent a lot of his time listening to my problems over a beer. Thanks to Alexander Arenz for bringing up some fresh air into our projects by joining the group. My compliments to all members of the Wittbrodt lab for the pleasant working atmosphere, especially to: Filo, Martina, Erika, Juan, Rebecca, Annette, Beate and Caroline.

I am indebted to Pam; she was on my side during all this time. In addition to her important mental support, she never hesitated to help me out with her golden cloning abilities. Thank you, Pamela.

I would like to dedicate this thesis to my parents who enabled my entire scientific career so far by giving me freedom to choose whatever I wanted and by their crucial mental and financial support. I will not forget to mention my brother Oliver and my sister Patricia.

You have all been great support!

Abbreviations

aa	Amino acids	LINE	Long interspersed element
AAV	Adeno-associated virus	LTR	Long terminal repeat
bp	Base pairs	M	Molar
BSA	Bovine serum albumine	MAR	Matrix attachment region
cDNA	Coding DNA	MBT	Mid-blastula transition
CFP	Cyan fluorescent protein	MCS	Multiple cloning site
CNS	Central nervous system	mM	Millimolar
DNA	Deoxyribonucleic acid	mRNA	Messenger RNA
dNTP	Deoxynucleic triphosphate	nls	Nuclear localisation signal
ds	Double stranded	NHEJ	Non homologous end joining
DSB	Double strand break	pA	Polyadenylation
ES	Embryonic stem	PCR	Polymerase chain reaction
F1	Filial generation 1	REMI	Restriction endonuclease mediated integration
FACS	Fluorescence activated cell sorting	RNA	Ribonucleic acid
G0	Generation 0	RT	Reverse transcriptase
GFP	Green fluorescent protein	rt	room temperature
hpf	Hours post fertilisation	SAR	Scaffold attached region
HR	Homologous recombination	SB	Sleeping Beauty
HSP	Heat-shock protein	SINE	Short interspersed element
IPTG		TcE	Tc1 like element
IR	Infrared	Tn	Transposon
IR/DR	Inverted/direct repeat	UAS	Upstream activating sequence
IS	Insertion sequence	UV	Ultraviolette
ITR	Inverted terminal repeat	YSL	Yolk syncytial layer
kb	Kilobase		