Evaluating the biological relevance of putative enhancers using Tol2 transposon-mediated transgenesis in zebrafish

Shannon Fisher[1,2,5], Elizabeth A Grice[1,5], Ryan M Vinton[1], Seneca L Bessling[1], Akihiro Urasaki[3], Koichi Kawakami[3] & Andrew S McCallion[1,4]

[1]McKusick–Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. [2]Department of Cell Biology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. [3]Division of Molecular and Developmental Biology, National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan. [4]Department of Molecular and Comparative Pathobiology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. [5]These authors contributed equally to this work. Correspondence should be addressed to S.F. or A.S.M. (sfisher@jhmi.edu; amccalli@jhmi.edu)

Published online 12 October 2006; doi:10.1038/nprot.2006.230

Evaluating the biological relevance of the myriad putative regulatory noncoding sequences in vertebrate genomes represents a huge challenge. Functional analyses *in vivo* have typically relied on costly and labor-intensive transgenic strategies in mice. Transgenesis has also been applied in nonrodent vertebrates, such as zebrafish, but until recently these efforts have been hampered by significant mosaicism and poor rates of germline transmission. We have developed a transgenic strategy in zebrafish based on the Tol2 transposon, a mobile element that was recently identified in another teleost, Medaka. This method takes advantage of the increased efficiency of genome integration that is afforded by this intact DNA transposon, activity that is mediated by the corresponding transposase protein. The approach described in this protocol uses a universal vector system that permits rapid incorporation of DNA that is tagged with sequence targets for site-specific recombination. To evaluate the regulatory potential of a candidate sequence, the desired interval is PCR-amplified using sequence-specific primers that are flanked by the requisite target sites for cloning, and recombined into a universal expression plasmid (pGW_*cfos*EGFP). Purified recombinant DNAs are then injected into 1–2-cell zebrafish embryos and the resulting reporter expression patterns are analyzed at desired timepoints during development. This system is amenable to large-scale application, facilitating rapid functional analysis of noncoding sequences from both mammalian and teleost species.

INTRODUCTION

Functionally constrained noncoding sequences, defined as those evolving more slowly than neutral (nonfunctional) sequences[1], are frequently hypothesized to determine tissue specificity, timing and levels of gene expression[2,3], as well as other roles. The identification of putative regulatory elements has been facilitated by the analysis of multiple orthologous genomic sequence intervals and the rapid development and refinement of computational tools. However, our ability to assess and, ultimately, to predict the biological functions of conserved noncoding sequences remains extremely limited by inefficient methods for functionally testing computational predictions. Cell culture assays permit the analysis of large numbers of sequences, but overlook the complexity of developmental and tissue-specific gene regulation. Functional analyses *in vivo* typically rely on transgenesis in mice, which, although highly informative, is costly and labor-intensive, frequently precluding comprehensive analysis of even a single locus. Transgenesis has also been used in nonrodent vertebrates, such as zebrafish and *Xenopus*. However, these approaches are limited by reliance on expression from episomal DNA and visually inaccessible *Xenopus* embryos. Additionally, standard DNA transgenesis in zebrafish generates highly mosaic G0 embryos, expressing transgenes in $<10\%$ of appropriate

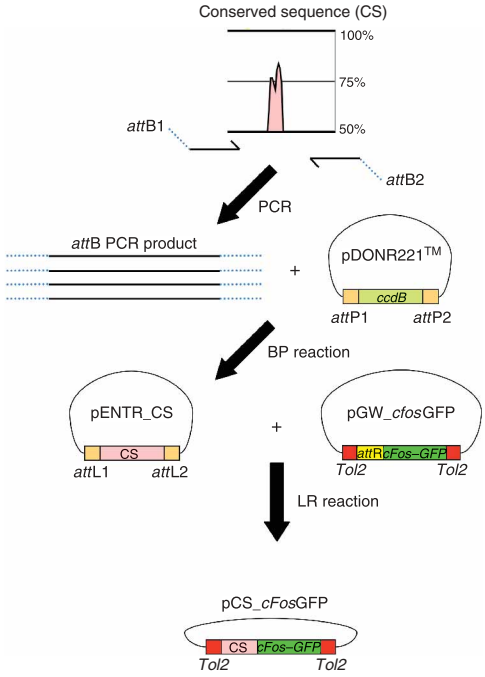

Figure 1 | Cloning a conserved noncoding sequence into the Tol2 transposon expression vector. Conserved noncoding sequences are identified by sequence alignment — in this case, using the VISTA server. Primers that contain 5′ *att*B sequences are designed to amplify the conserved noncoding sequences. The ensuing PCR product is then inserted into an entry vector (pDONR™221) via BP recombination. The resulting construct is recombined with the destination vector (pGW_*cfos*EGFP) by LR recombination, so that the conserved noncoding sequence is placed in the context of a *c-fos* minimal promoter driving EGFP expression. After purification and quantification, the construct is ready for injection into zebrafish embryos.

cells. This high degree of mosaicism has necessitated strategies such as the reconstruction of overall expression patterns from scattered positive cells in numerous G0 embryos[4].

To address the issue of extreme G0 mosaicism in the visually accessible zebrafish embryo, we developed a reporter vector to functionally examine putative enhancers in transgenic zebrafish. We based our vector on the Tol2 transposon, which was originally identified from the medaka *Orzyas latipes*[5]. Other methods have been developed to increase the efficiency of zebrafish transgenesis, based on the Sleeping Beauty transposon[6,7] or relying on I-SceI meganuclease digestion of injected DNA[8]. However, the reported rates of germline transmission for Tol2 vectors are higher[9] than for these alternative methods. We also observed substantially greater expression of an ubiquitous control construct in G0 embryos with a Tol2 vector than with one based on Sleeping Beauty (S.F. and A.S.M., unpublished observations). We adapted an existing plasmid[9], constructing a smaller Tol2 vector containing essential *cis* sequences for transposition and placed the Gateway® *ccdB* recombination cassette and mouse *cFos* minimal promoter[10] upstream of the *EGFP* gene. Without the addition of further sequences, the *cFos* minimal promoter fails to drive reporter-gene expression in transgenic zebrafish. Inserting a regulatory element with positive activity — for example, an enhancer sequence — into the Gateway® cassette will result in EGFP expression reflecting the normal regulatory activity of the enhancer, whereas insertion of a sequence with negative or no regulatory activity will not lead to detectable EGFP. We have previously reported the use of this system in a comprehensive analysis of noncoding conserved sequences around the zebrafish and human *RET* genes[11]. We further demonstrated the function of multiple human sequences as specific enhancer elements in zebrafish embryos in the absence of detectable sequence conservation across the same evolutionary span. Thus, the utility of our approach can extend to mammalian loci for which the corresponding zebrafish gene has not been characterized, or in cases in which sequence conservation is not detected beyond coding exons.

The ability to rapidly examine the regulatory potential of all putatively functional noncoding sequences in a cost-effective manner is essential for a full understanding of their biological role and to further refine the computational tools used in their prediction. We present here an approach using a high-efficiency vector in visually accessible zebrafish embryos, which will facilitate large-scale functional analysis of sequences from vertebrate genomes. At this time, the assay is designed only to identify positive regulatory elements — for example, enhancers; new strategies are necessary to facilitate the evaluation of negative regulatory elements. In addition, the potential effect of context on the detectable activities of such regulatory elements has not yet been fully evaluated, so all resulting data should be interpreted in light of this caveat. However, successful use of this technology will yield new *in vivo* substrates for lineage analysis during development and disease processes; it will facilitate the elucidation of complex regulatory networks, and it may also be used to support ongoing activities to permit functional annotation of vertebrate genomes.

MATERIALS

REAGENTS

- Zebrafish: injections are performed in embryos of the strain AB[12], raised in our facility; AB fish can be obtained from the Zebrafish International Resource Center (http://zfin.org)
- Takara LA Taq™ polymerase kit (Takara Mirus Bio; cat. no. TAK_RR002M)
- Agarose gels
- PCR Cloning System with Gateway® Technology (with pDONR™221) (Invitrogen; cat. no. 12535-019)
- Library efficiency competent cells DH5α™, SOC media included (Invitrogen; cat. no. 18263-012)
- Library efficiency competent cells DB3.1™, SOC media included (Invitrogen; cat. no. 11782-018)
- QIAprep® Spin Miniprep Kit (Qiagen; cat. no. 27104)
- Gel purification kit (Qiagen; cat. no. 28704)
- QIAquick® PCR Purification Kit (Qiagen; cat. no. 28104)
- pGW_*cfos*EGFP (universal acceptor Tol2 transposon vector available from authors)
- HiSpeed Plasmid Midi Kit (Qiagen; cat. no. 12643)
- mMessage mMachine Kit Sp6 (Ambion; cat. no. 1340)
- pCS_Tp (contains transposase gene, available from authors)
- LB broth
- Ampicillin stock (50 mg ml^{-1})
- Kanamycin stock (50 mg ml^{-1})
- Ampicillin (50 μg ml^{-1}) LB agar plates
- Kanamycin (50 μg ml^{-1}) LB agar plates
- *Eco*RV (New England BioLabs; cat. no. R0195S)
- Embryo medium: 1× salt stock and 1× bicarbonate stock (see REAGENT SETUP)
- Phenol red (2% stock in dH_2O)
- Kwik-Fil thin-wall borosilicate glass capillaries, 1.2 mm OD (WPI)
- Tricaine (3-amino benzoic acid ethyl ester, also called ethyl m-aminoboenzoate) (Sigma; cat. no. A-5040), 0.4% (v/v) stock in dH_2O
- Gene-specific oligonucleotide primers (20 μM working stock)

EQUIPMENT

- Thermocycler
- Shaking incubator
- Gel imaging equipment
- Spectrophotometer
- Pneumatic Pico Pump PV 820 injector (World Precision Instruments)
- Stemi 2000 stereomicroscope (for injection; Carl Zeiss)
- Lumar V12 Stereo stereomicroscope (for analysis and photography; Carl Zeiss)
- AxioVision Rel. 4.4 (Carl Zeiss)
- AxioCam HRm (Carl Zeiss)
- MicroPipette Puller Model P87 (Sutter Instrument Co.)
- Borosilicate glass capillaries (World Precision Instruments; cat. no. TW120F-3)
- Breeding tanks (e.g., small breeding tank; Aquatic Habitats; cat. no. SBTANK2)

REAGENT SETUP

20× salt stock Add the following components in order to 800 ml of dH_2O, allowing each salt to dissolve before adding the next one: 17.5 g NaCl, 0.75 g KCl, 2.9 g $CaCl_2$, 2.39 g $MgSO_4$, 0.41 g KH_2PO_4 and 0.13 g Na_2HPO_4. Add dH_2O to a final volume of 1 L. Filter sterilize solution and store at 4 °C.

500× bicarbonate stock Dissolve 1.5 g of $NaHCO_3$ in 50 ml of dH_2O. Store at 4 °C.

Embryo medium (8 L) Mix 400 ml of 20× salt stock, 16 ml of bicarbonate stock, and dH_2O to a final volume of 8 L. Optional: To minimize fungal growth in embryo dishes, methylene blue ($C_{16}H_{18}ClN_3S$) can be added to the embryo medium. Prepare a 0.1% solution of methlyene blue in embryo medium; add 8 ml of methylene blue stock along with other stocks to an 8 L batch of embryo medium.

Identification of conserved noncoding sequences in the interval of interest Identify orthologous sequence intervals for multiple organisms encompassing the interval by creating a multiple sequence alignment using appropriate software. This may now be readily performed using the publicly available UCSC genome browser (http://genome.ucsc.edu), which permits an investigator to align and evaluate sequences *in situ* with sophisticated tools such as phastCons[13]. There are many freely available stand-alone alignment algorithms, each aiming to predict functional sequences predicated on overlapping but subtly different parameters; the one selected is a matter of personal preference. Some of the more

commonly used algorithms include VISTA[14], MultiPipmaker[15], Multi-species Conserved Sequences[16], Regulatory Potential[17] and LAGAN[18].

Primer design for PCR cloning Design primers to amplify the sequence of interest, typically including ≥30 bp flanking DNA on either side of the conserved sequence, as the boundaries of functional elements may not be readily predicted. Clusters of noncoding conserved sequences can be amplified in a single PCR product and their individual roles can be dissected subsequently if necessary. For primer design, use Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) or similar primer design software. To enable Gateway® cloning (see below), add 4 guanine (G) nucleotides to the 5' end of the forward primer, followed by the 25 bp *att*B1 site, followed by 18–25 bp of template-specific sequence (5′-GGGGACAAGTTTGTACAAAAAAGCAGGCT-template-specific sequence-3′). For the reverse primer, add 4 guanine (G) nucleotides followed by the 25 bp *att*B2 site, followed by 18–25 bp of template-specific sequence (5′-GGGGACCACTTTGTACAAGAAAGCTGGGT-template-specific sequence-3′). Once primers are obtained for the sequence of interest, they should be diluted to 20 μM concentration. For more information on primer design, consult the Gateway® website (http://www.invitrogen.com/content.cfm?pageid=4072). The addition of *att*B sites to PCR primers can decrease the efficiency of primer annealing, Therefore, standard restriction enzyme-based cloning strategies or gene-specific primers incorporating selected restriction sites may also be used to clone amplicons into an alternative entry vector (pENTR™2B, Invitrogen).

Gateway cloning For cloning purposes, we use the Gateway® Technology, a rapid, highly efficient universal cloning system based on the site-specific recombination properties of the bacteriophage lambda. We have found that sequences of less than 6 kb may be readily managed by both the Gateway® system and Tol2 transposition capabilities. Once primers are designed and the desired sequence is amplified with flanking *att*B sites, a recombination reaction transfers the PCR product to a donor vector pDONR™221, containing *att*P sites (**Fig. 1**). This is the BP reaction, and the resulting construct, referred to as an entry clone, contains the sequence of interest flanked by *att*L sites. Note that the term BP is not an acronym; it refers to the recombination event that occurs between the *att*B and *att*P sites (BP) on the PCR product and the donor vector (pDONR), respectively. From the entry clone, the noncoding conserved sequence can be shuttled by LR recombination to any Gateway® ready destination vector — for example, pGW_*cfos*EGFP, which contains a *ccdB* gene and chloramphenicol gene flanked by *att*R recombination sites (**Fig. 1**). As above, the term LR is not an acronym; it refers to the recombination event that occurs between the *att*L and *att*R sites (LR) (see **Fig. 1**). The *ccdB* gene serves as a negative selection gene for the destination vector. *ccdB* encodes a protein that interferes with *Escherichia coli* DNA gyrase and is therefore lethal except in certain bacterial strains, such as DB3.1™ (Invitrogen).

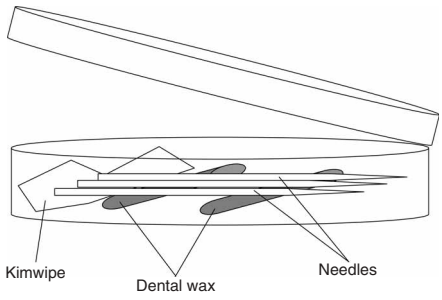

Figure 2 | Filled needles are stored in a humidified holding dish. In a 150-mm Petri dish, place two parallel rolls of dental wax or poster mounting material. Lay needles on top with gentle pressure to fix in place. Insert a Kimwipe soaked with water in the dish, to prevent evaporation of the injection solution from filled needles. Pulled needles can be stored in the closed dish before filling; after filling, the dish can be sealed with a strip of Parafilm to further prevent solution loss.

Therefore, the destination vector should only be propagated in DB3.1™ cells. When LR recombination occurs, the *ccdB* gene and chloramphenicol-resistance gene are replaced by the sequence of interest, and therefore are able to be propagated in DH5α™ strains. We recommend taking the time to familiarize oneself with the manufacturer's manual on Gateway® cloning and associated protocols (http://www.invitrogen.com/content.cfm?pageid=4072).

EQUIPMENT SETUP

Preparation of injection needle Injection needles are pulled from 1.2 mm OD filament capillary glass, with a program designed to yield a strong tip with a fairly sharp taper, to penetrate intact chorions. Break the tips by hand under a stereomicroscope to an outer diameter of approximately 15 μm, using a clean razor blade and a micrometer slide to measure the diameter. Prepared needles can be made the day before injections; store in a covered needle-holding dish (**Fig. 2**) to keep clean. ▲ CRITICAL The taper of the needles and the diameter of the tips are important factors in the ease of injections. If the needle tapers too gradually, then the tip will be too flexible to easily penetrate the chorion. Conversely, if the taper is too sharp, it will be difficult to break the tip to the correct diameter. If the tip diameters are inconsistent, then it will be necessary to recalibrate the injection volumes between needles.

PROCEDURE

Cloning sequences of interest into the transposon vector, pGW-*cfos*EGFP ● TIMING approximately 1 week

1| Set up PCRs, as in the table below, to amplify the noncoding conserved sequence with specific *att*B-containing primers designed as outlined in REAGENT SETUP; use total genomic DNA or a large insert genomic clone as a template. Use the Takara *LA Taq*™ system, or similar *Taq* polymerase with proofreading capabilities.

▲ CRITICAL STEP Use of a proofreading polymerase is highly advisable for avoiding the introduction of potentially deleterious mutations in sequences that are to be functionally evaluated. We regularly use Takara™ *Taq* polymerase, which will amplify sequences up to 20 kb in length, which is significantly in excess of our present requirements (0.5–2.5 kb).

▲ CRITICAL STEP Thaw reagents and set up reactions on ice. Make sure all reagents are thoroughly mixed prior to the addition to the reaction mix. Although reaction components need not be added in the strict order that is laid out below, we find it helpful to have a printed template of components that can be checked/ticked off as reagents are added.

PCR reaction mix:

Component	Amount (per reaction)	Final amount/concentration
Sterile water	20 μl	
10× LA PCR buffer	3 μl	1×
dNTP mix (2.5 mM)	4.8 μl	1×
*att*B1 forward primer (20 μM)	0.4 μl	0.27 μM
*att*B2 reverse primer (20 μM)	0.4 μl	0.27 μM
Genomic DNA (100 ng μl^{-1})	1 μl	100 ng
Takara *Taq* polymerase (5 U μl^{-1})	0.4 μl	2 units
TOTAL volume	30 μl	

2| Transfer the reactions to a thermocycler programmed as detailed below.

PCR cycling conditions:

Cycle	Denature	Anneal/Extend	Extend
1	95 °C, 1 min		
2–30	95 °C, 30 s	68 °C, 1 min/1 kb	
31			68 °C, 10 min

3| Run the entire PCR product on an agarose gel and excise the band**.** Choose an appropriate percentage gel based on the size of the PCR product.
? TROUBLESHOOTING

4| Purify the PCR product with the QIAquick® Gel Extraction kit (Qiagen) or equivalent, eluting the DNA from the column with 20–50 µl of Buffer EB. This kit can be used for PCR products ranging in size from 70 bp to 10 kb. Each column is capable of binding up to 10 µg, and recovery is typically 70–80%. To determine recovery, it is useful to run 3–5 µl of the extracted DNA on an agarose gel to assess the efficiency of the extraction. The purified PCR product is then quantified with a spectrophotometer. We generally only accept yields in excess of 25 ng μl^{-1} for subsequent cloning steps.

5| Generate the *Entry Clone* (pENTR_CS, **Fig. 1**) by incubating the purified PCR product containing *att*B recombination sites with a donor vector (pDONR™ 221) containing *att*P recombination sites, and the BP Clonase™ recombination enzyme, as described in the Gateway® manual. The resulting construct, referred to as an *Entry Clone*, now contains the noncoding conserved sequence of interest, flanked by *att*L sites (see **Fig. 1**). Importantly, conventional methods, such as restriction-enzyme-based cloning strategies, may also be used to sub-clone PCR products or restriction fragments to create pENTR_CS (refer to the manufacturer's manual on Gateway® cloning for additional details).
? TROUBLESHOOTING

6| Transfer the amplified sequence from pENTR_CS into the pGW-*cfos*EGFP destination vector by LR recombination; detailed instructions of these steps are provided in the Gateway® manual. This vector contains Tol2 transposon sequences and Gateway® *att*R recombination sequences, upstream of a *cFos* minimal promoter[10] and the EGFP coding sequence. The manufacturer also provides a positive control for the LR reaction.
? TROUBLESHOOTING

7| To verify the product of the LR recombination, digest ~500 ng of plasmid with *Eco*RV, using the manufacturer's recommended conditions, to release the insert. Confirm the size of the insert by agarose-gel electrophoresis. However, as mutations introduced during amplification and cloning may influence the biological activity of the sequence being tested, sequencing is recommended; primers used for amplification may be used for sequencing. In our experience, the vast majority of clones obtained by LR recombination will contain the insert of interest, although sequence-specific variation in efficiency may occur. In this case, see the manufacturers guide for troubleshooting suggestions.
? TROUBLESHOOTING

8| Once an accurate clone has been identified, prepare plasmid DNA using the Qiagen HiSpeed® Plasmid Midi Kit. Pick selected colony into 1 ml of LB medium (50 µg ml^{-1} ampicillin), incubate at 37 °C with agitation (250 rpm) for 8 h, then transfer 500 µl to a flask containing 50 ml of LB medium (50 µg ml^{-1} ampicillin). Incubate flask and contents at 37 °C with agitation (250 rpm) for 16 h before extracting plasmid DNA according to the manufacturer's instructions.

9| Optional: to further purify the plasmid, we use the QIAquick® PCR Purification Kit, according to the manufacturer's protocol. We have included this as an optional step because many investigators will be satisfied to inject DNA prepared as in Step 8. However, embryos are often sensitive to contaminants that can be carried through standard DNA preparation protocols. We suggest this method as a means of circumventing any potential toxicity associated with injected DNAs. Equivalent kits may also be used, but we have found this specific kit to provide consistent high-quality results. Elute the DNA with 30 µl RNase-free water. Please note that although RNase-free water may be purchased or prepared, we have observed that embryo toxicity may occasionally be associated with the organic chemicals that are used to prepare RNase-free water. However, we have never observed such toxicity issues with Ultrapure™ Millipore filtered water; neither have we observed RNA degradation with this water. Please substitute your preparation of choice where RNase-free water is mentioned. Quantify the eluted samples by spectrophotometry and dilute to a concentration of 125 ng μl^{-1}.
■ **PAUSE POINT** The plasmid stocks may be stored for extended periods at 4 °C.

▲ CRITICAL STEP Use RNase-free water to preserve the integrity of the transposase RNA at the injection stage. Early embryos are sensitive to amounts of injected plasmid DNA or impurities in plasmid preparations. The cleanliness of the plasmid DNA is critical for good survival and normal development of injected embryos, and the quantification must be accurate. The optical density ratio 260 nm:280 nm ($OD_{260:280}$) should be between 1.7 and 1.9; outside this range, we suggest repurification. However, this ratio is not an absolute indicator of DNA purity and experiments should incorporate appropriate controls (discussed later) to uncover DNA that is suspended in a solution that is toxic to the embryos. Thus, embryo death can similarly indicate a need to re-purify DNA.

***In vitro* transcription of transposase RNA** ● TIMING **approximately 24 h**

10| RNA encoding functional Tol2 transposase enzyme is transcribed *in vitro* from the pCS-Tp vector[9]. Purify the pCS-Tp plasmid using the Qiagen Midi-Prep kit. Bacterial cultures should be established from a single colony picked from freshly streaked (≤4 weeks old) plates and prepared as described in Step 8 above. Linearize 10–20 μg plasmid with *Not*I using the manufacturer's recommended conditions. Set up the digest in a total volume of 100 μl, in a 1.5 ml micro-centrifuge tube.

11| Add Proteinase K to the entire linearized template from above to a final concentration of 100–200 μg ml^{-1} and incubate for an additional 15 min at 37 °C, to ensure destruction of restriction enzyme or other proteins, particularly contaminating RNases.

12| Add an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) to the sample in microcentrifuge tubes. Mix the contents until an emulsion form, then centrifuge at maximum speed for 1 min at room temperature.
! CAUTION Phenol and chloroform are extremely toxic and should be used in an externally vented fume hood. Protective clothing (gloves, eye wear and lab coat) is required.

13| Transfer the aqueous (upper) phase to a fresh microcentrifuge tube and discard the interface and organic phase. Add an equal volume of chloroform and repeat the centrifugation and recovery of the aqueous phase.
! CAUTION Chlorofom is extremely toxic and should be used in an externally vented fume hood. Protective clothing (gloves, eye wear and lab coat) is required.

14| Add sodium acetate to a final concentration of 0.3 M and 1 volume of isopropanol and incubate at –20 °C for 2–16 h. Yield varies little between these time points but it is often a convenient place to pause for the day.

15| Spin chilled solution at maximum speed for 15 min at 4 °C. Discard the supernatant, being careful not to pour off the pellet.

16| Wash the pellet with 70% ice-cold ethanol and recentrifuge at maximum speed for 5 min at 4 °C. Air-dry the pellet for 5 min in a fume hood, and resuspend in RNase-free water to yield a final concentration of 200 ng μl^{-1} to 2 μg μl^{-1}.

17| Set up a transcription reaction with the mMessage mMachine® Sp6 kit according to the manufacturer's instructions. From a single reaction starting with 1 μg of template, a typical yield is 20 μg of RNA.

18| Purify and precipitate the RNA according to kit instructions. Resuspend RNA to a final concentration of ~1 μg μl^{-1} — that is, 20 μl for a single reaction, in RNase-free water, and quantify by UV spectrophotometry. In addition, analyze ~1 μg of RNA by agarose-gel electrophoresis to verify full-length transcription. Although a standard TAE or TBE gel is adequate for this analysis, the denaturing sample buffer included with the transcription kit should be used according to kit instructions.
▲ CRITICAL STEP The purity, integrity and quantity of transposase RNA are critical to the success of the injections. RNA should provide an $OD_{260:280}$ between 1.8 and 2.0; outside this range, we suggest repurification. We suggest the Qiagen RNeasy® mini kit. In addition, separate batches of RNA may have different activities. It may be useful to test each new batch of RNA with a control plasmid to verify good activity.
■ PAUSE POINT Aliquots of transposase RNA (175 ng μl^{-1}) can be stored at –80 °C (≤6 months).

Fish husbandry and matings ● TIMING **approximately 2 d**

19| Maintain the zebrafish on a regular light–dark cycle, with 14 h of light. The day prior to performing microinjections, set up fish for timed matings in small breeding tanks, each consisting of a base tank, a slotted insert and a plastic lid. Create parallel rows of single sex tanks of fish; each row should comprise tanks with either three females or two males per tank. Placement of a small plastic tree in each tank prevents males from fighting overnight. For detailed guidance regarding zebrafish husbandry and associated techniques, please refer to *The Zebrafish Book*[19].

20| The morning of the microinjections, shortly after the light cycle begins, set up 2 tanks containing 2 males and 3 females in clean system-treated water. Egg production should initiate shortly thereafter, permitting the production of ≥200 eggs within 15 min. Timed production of good-quality eggs can typically be continued over a 2-h period after the normal 'lights on' time, by mixing tanks of males and females just prior to use. The yield of eggs depends critically on the light–dark cycle; females are

most likely to lay shortly after the lights come on, and over the next several hours the quality and quantity of eggs laid decreases. Clutches of >200 eggs are preferable for injections, as they allow several experimental groups of 50 embryos to be injected, and an uninjected dish to also be set aside as a control for egg quality. Although smaller batches of eggs may be of good quality, they are less convenient for injections. Poor-quality eggs or unfertilized eggs will generally fail to progress to the 2-cell stage. These eggs should not be used for injections. However, some clutches will undergo early cell divisions and will be used for injection only to fail to develop normally, hence the benefit of a control plate of uninjected embryos to discern whether embryo death is a consequence of injection conditions or embryo health.

21| To collect embryos, lift the slotted insert out of the base tank and place the fish into a new base filled with system-treated water. Allow the embryos to settle to the bottom of the tank. Pour off most of the water and then pour the embryos into a 60 × 15 mm Petri dish.

22| With a wide-bore, 5 1/4″ glass pasteur pipette fitted with a latex bulb, sort the collected embryos into 60 × 15 mm Petri dishes, partially filled with embryo medium (see REAGENT SETUP), in groups of 50 embryos. Mark the time collected and number of embryos on the lid of each dish. It is convenient to inject embryos in groups of 50; this normally provides enough embryos expressing the construct extensively to allow characterization of the expression pattern, and a 60 mm dish has sufficient volume of water to keep 50 embryos for 5–6 d.

▲ **CRITICAL STEP** The timing of injections, at the late one-cell to early two-cell stage, is important for extensive transgene expression and normal development. For ease in injecting large clutches of eggs, it is helpful to carefully monitor the fish and collect eggs within a few minutes of laying. Otherwise, the fish may continue to lay over an extended period, and the clutch will not be well synchronized.

Injection of embryos with transposons ● TIMING approximately 3 h

This timing refers to the likely productive period within which multiple clutches of eggs may be collected (see above) plus the time taken to inject them.

23| Prepare fresh injection solution by mixing the following in a microcentrifuge tube on ice:

Component	
Transposon plasmid stock (125 ng μl^{-1})	1 μl
Transposase RNA stock (175 ng μl^{-1})	1 μl
Phenol red stock (2% in H_20)	0.5 μl
RNase-free water	2.5 μl

24| Lay injection needles in the holding dish (see EQUIPMENT SETUP and **Fig. 2**), and fill by pipetting 500 nl drops of injection solution onto the wide end of each needle. After the liquid is drawn to the tip through capillary action, additional injection solution can be added to a total of 1.5–2 μl; letting the liquid draw to the tip before adding more liquid helps to prevent air bubbles in the needle. At least two needles should be prepared for each injection solution, depending on the number of different constructs and the total number of embryos to be injected. This provides a backup in case a needle becomes blocked or breaks. In general, plan on about one needle per 100 embryos to be injected, with at least one extra per construct in case of breakage or blockage. Keep the needle dish covered as much as possible, and place a Kimwipe soaked in water in the dish to minimize evaporation of injection solution (see **Fig. 2**). We have never examined the maximum time that solution is stable in the needle, but do not observe a drop in efficacy over a 3-h period of injection.

25| Load a filled needle into the hand-held needle holder of a Pneumatic Pico-Pump or similar pressure injector, configured and connected to a N_2 tank per manufacturer's instructions. Calibrate injection volumes by measuring the diameter of droplets expelled into mineral oil on a micrometer slide (**Fig. 3a**). Typically, an injection time of 120 ms with a pressure of 20 p.s.i. will yield a droplet of approximately 1 nl, but slight variations in needle diameter will affect these parameters and recalibration may be required between needles. Once the parameters are adjusted to give the desired injection volume, place the tip into the liquid in an injection dish and adjust the back pressure until injection solution is extruded slowly from the tip between injections. The back pressure will prevent dilution or contamination of the injection solution in the needle.

26| Perform injections with the aid of a stereomicroscope at 6–10× magnification. Some investigators will prefer to line up embryos in an agarose injection tray to stabilize them for injection[19]. We find this an unnecessary time delay, and prefer, instead, to use a pair of fine forceps to hold the embryo in place. However, take care not to put any pressure on the embryo after the needle penetrates the chorion, to avoid pushing the embryo out through the small hole. Push the injection needle with steady pressure through the chorion and into the yolk of an embryo at the late one-cell or early two-cell stage. Ideally,

position the needle tip in the yolk just below the blastomeres. Expel approximately 1 nl of injection solution and withdraw the needle. The expelled volume should be visible as a phenol red stained drop below the blastomeres (**Fig. 3b**). Although some investigators will prefer to use a micromanipulator to perform injections, we find that we can move swiftly injecting by hand. Experienced personnel should be able to inject at least 600 embryos in a 2-h period, by collecting embryos from several successive lays. We prefer to inject 150–200 embryos per construct. Thus, 3–4 petri dishes of approximately 50 embryos per dish will be completed for each construct. Injection of larger numbers of embryos — for example, 600 as discussed above — will require multiple egg collections to ensure that injected embryos are synchronized. Embryos may take up to 30 min to progress beyond the 2-cell stage. Embryo collection should be repeated until sufficient embryos have been collected to complete desired injections ($\leq$200 embryos per construct) or until embryo production ceases.

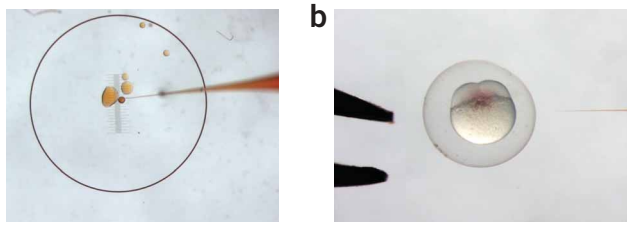

Figure 3 | Calibration and injection of fixed volumes into zebrafish embryos. (**a**) Injection volume is calibrated by measuring the diameter of a droplet expelled into mineral oil. (**b**) A two-cell embryo immediately after injection; phenol red is visible in the yolk below the blastomeres.

27| After injections are completed, sort the embryos by removing unfertilized eggs, damaged embryos and failed injections (embryos with no phenol red in blastomeres).
▲ CRITICAL STEP Unfertilized eggs and damaged embryos must be removed promptly to ensure normal development of the remaining embryos in the dish. Otherwise, the remaining live embryos may be killed or severely delayed in development.
? TROUBLESHOOTING

Analysis of expression patterns ● TIMING **approximately 1–6 d (depending on type of analysis)**

28| After culture for the appropriate time, screen the G0 embryos for EGFP expression. At early stages, prior to 24 h post-fertilization, the embryos can be directly observed. At later stages, when the embryos are motile and have begun hatching out of their chorions, they can be anesthetized with Tricaine ($\sim$10 drops of 0.4% stock in a 50-mm dish) to facilitate observation. Large clutches of embryos are most conveniently observed on a stereomicroscope that is fitted for epifluorescence, such as a Zeiss SV11 or Lumar V12. For high-resolution photography, the Lumar V12 or a compound microscope will be necessary. If fluorescent reporters are being used, it will be necessary to obtain appropriate filters to visualize the corresponding signal; you may wish to discuss the desired specifications with your microscope vendor. Observations of the live embryos may be continued throughout the first 5–6 d. See **Figure 4** for comparisons of injections performed with and without the transposase RNA and also for examples of mosaic (G0) expression patterns.
? TROUBLESHOOTING

29| After 5–6 d, select appropriate G0 embryos, move them to tanks and raise to sexual maturity. The likelihood and rate of germline transmission seems to correlate with the extent of mosaic expression; therefore, those G0 embryos with the most expression are selected for raising.

Sexual maturation of zebrafish ● TIMING **3 months**

30| Cross sexually mature G0 adults with wild-type stocks to obtain germline transmission and establish founder G1 transgenic stocks. Although this transposon-based approach results in multiple independent insertion events per G1 individual, we recommend that multiple independent G1 lines are established from different founders to avoid the confounding influence of position effects (see ANTICIPATED RESULTS).
? TROUBLESHOOTING

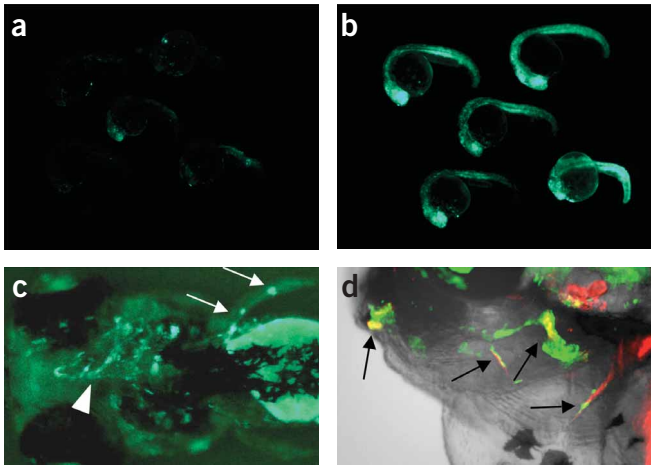

Figure 4 | Demonstration of the increased efficiency of transgenesis observed using transposon-mediated transgenesis. Comparison of the efficiencies of Tol2 transgene integration observed in the absence (**a**) and the presence (**b**) of transposase RNA. The construct used in both **a** and **b** uses a strong ubiquitous promoter (elogation factor 1-alpha) driving EGFP[9]. Examples of specific expression patterns observed in mosaic G0 embryos (**c**,**d**). (**c**) Ventral view of EGFP expression in the branchial arch vasculature (arrowhead) and in the pectoral fin (arrows) of a G0 embryo. (**d**) EGFP expression in developing bones of the craniofacial skeleton (arrows); embryo (G0) is also stained with alizarin red to detect calcification of bones.

● TIMING

Steps 1–9 (cloning sequences of interest into the transposon vector, pGW-*cfos*EGFP): approximately 1 week
Steps 10–18 (*in vitro* transcription of transposase RNA): approximately 24 h and can be performed in parallel with Steps 1–9
Steps 19–22 (fish husbandry and matings): approximately 2 d
Steps 23–27 (injection of embryos with transposons): approximately 3 h
Steps 28–30 (analysis of expression patterns): 1–6 d
Steps 28–30 (sexual maturation of zebrafish): 3 months
Steps 28–30 (fish husbandry and matings): approximately 2 d
Steps 28–30 (identification of G1 progeny): ≤1 h

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 1**.

TABLE 1 | Troubleshooting table.

PROBLEM	SOLUTION
Step 3: PCR does not yield expected product	The addition of *att*B sites to PCR primers can decrease the efficiency of primer annealing. If using genomic DNA as template, consider obtaining a BAC clone encompassing the region for template in the PCR reaction. If this does not yield a PCR product, design new primers for the sequence of interest.
Steps 5 and 6: Few or no colonies obtained from BP or LR recombination	Ensure that primers contain proper *att*B sites. If PCR product is especially long (>5 kb), BP and LR reactions may need to be incubated for longer periods of time. For a more complete troubleshooting guide, consult the Gateway® manual, which is available online at http://www.invitrogen.com/content.cfm?pageid=4072.
Step 7: High background with few colonies containing insert	The *ccdB* gene in the destination vector may be mutated/deleted. Determine the integrity of the destination vector, and always keep the vector under antibiotic selection. Additionally, reagents used in the LR reaction may be contaminated with a different plasmid containing the antibiotic-resistance gene.
Step 27: A large number of injected embryos die or develop abnormally	The likely problem is either a toxic impurity in a component of the injection solution or too much DNA injected per embryo. Check the quantification of the DNA and the injection volume again. To look for impurities, perform control injections of similar volumes with individual components missing; remake or replace any that correlate with abnormal embryos.
Step 28: Little or no expression in mosaic embryos	The putative regulatory element may not have enhancer activity in the embryo. However, if you think this is unlikely, or if you see expression inconsistently or at low levels, the transposase RNA may have become degraded or have poor activity. Perform injections with a control plasmid (with a ubiquitously expressed or known promoter), with and without transposase RNA, to compare expression levels.
Step 30: Only a single line is successfully passed through the germline	This is often considered insufficient to confidently report the expression pattern regulated by an element under examination. As discussed in the text, the potential of position effects influencing the expression pattern often necessitate evidence from multiple independent lines. Continue to set up matings using transgene-positive G0 fish. Additional rounds of injection should be performed to provide new stocks of G0 fish for matings to achieve germline transmission.

ANTICIPATED RESULTS

Under optimal injection conditions, the large majority (≥80%) of injected embryos will develop normally. In general, expression patterns that are consistent among at least 10–20% of embryos will be highly representative of the nonmosaic expression observed from the same constructs after germline transmission. However, detailed characterization of an expression pattern will require the establishment of transgenic lines. To ensure that position effects on individual transgene insertions are not confounding the interpretation of expression patterns, multiple independent lines should be established for each construct. The term 'position effect' refers to differences in expression that can be observed from identical transgenes because of regulatory control imposed on them by the genomic context in which they have inserted. Consequently, we suggest the generation of two or more independent lines for evaluation. Because of the high rate of integration of Tol2 vectors (see **Fig. 4**), in most cases fewer than 20 G0 adults need to be screened to identify more than one transgenic founder. From individual founders, we have observed germline transmission rates from <5% to >95%, although ~35% is more typical. Note that the above protocol, as written, is designed specifically to permit the evaluation of potential activating regulatory sequences (e.g., enhancers). Although alterations in the design of the protocol may be readily envisaged to permit the evaluation of

suppressor-type sequences, they have not yet been established. Additionally, by appending the Tol2 sequences to any desired transgene construct, this protocol may be used to create zebrafish transgenics more efficiently with any desired sequence.

ACKNOWLEDGMENTS This work was supported by a Basil O'Conner starter scholar award from the March of Dimes (A.S.M.) and grant GM071648 from the NIGMS (A.S.M.).

COMPETING INTERESTS STATEMENT The authors declare that they have no competing financial interests.

Published online at http://www.natureprotocols.com
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Kimura, M. & Ota, T. Protein polymorphism as a phase of molecular evolution. *Nature* **229**, 467–469 (1971).
2. Pennacchio, L.A. & Rubin, E.M. Genomic strategies to identify mammalian regulatory sequences. *Nature Rev. Genet.* **2**, 100–109 (2001).
3. Waterston, R.H. *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
4. Woolfe, A. *et al.* Highly conserved noncoding sequences are associated with vertebrate development. *PLoS Biol* **3**, e7 (2005).
5. Koga, A., Suzuki, M., Inagaki, H., Bessho, Y. & Hori, H. Transposable element in fish. *Nature* **383**, 30 (1996).
6. Davidson, A.E. *et al.* Efficient gene delivery and gene expression in zebrafish using the Sleeping Beauty transposon. *Dev. Biol.* **263**, 191–202 (2003).
7. Ivics, Z., Hackett, P.B., Plasterk, R.H. & Izsvak, Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* **91**, 501–510 (1997).
8. Thermes, V. *et al.* I-SceI meganuclease mediates highly efficient transgenesis in fish. *Mech. Dev.* **118**, 91–98 (2002).
9. Kawakami, K. *et al.* A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish. *Dev. Cell* **7**, 133–144 (2004).
10. Dorsky, R.I., Sheldahl, L.C. & Moon, R.T. A transgenic Lef1/β-catenin-dependent reporter is expressed in spatially restricted domains throughout zebrafish development. *Dev. Biol.* **241**, 229–237 (2002).
11. Fisher, S., Grice, E.A., Vinton, R.M., Bessling, S.L. & McCallion, A.S. Conservation of RET regulatory function from human to zebrafish without sequence similarity. *Science* **312**, 276–279 (2006).
12. Johnson, S.L. & Zon, L.I. Genetic backgrounds and some standard stocks and strains used in zebrafish developmental biology and genetics. *Methods Cell Biol* **60**, 357–359 (1999).
13. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050 (2005).
14. Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M. & Dubchak, I. VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* **32**, W273–W279 (2004).
15. Schwartz, S. *et al.* PipMaker — a web server for aligning two genomic DNA sequences. *Genome Res.* **10**, 577–586 (2000).
16. Margulies, E.H., Blanchette, M., Haussler, D. & Green, E.D. Identification and characterization of multi-species conserved sequences. *Genome Res.* **13**, 2507–2518 (2003).
17. Kolbe, D. *et al.* Regulatory potential scores from genome-wide three-way alignments of human, mouse, and rat. *Genome Res.* **14**, 700–707 (2004).
18. Brudno, M., Chapman, M., Gottgens, B., Batzoglou, S. & Morgenstern, B. Fast and sensitive multiple alignment of large genomic sequences. *BMC Bioinformatics* **4**, 66 (2003).
19. Westerfield, M. (ed.) *The Zebrafish Book* (University of Oregon Press, Eugene, OR, 1995).