Stony Brook University

PhiC31 Integrase as a More Efficient Method of Transgenesis in Zebrafish

A Thesis Presented

by

Andrew Scortea

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Master of Science

in

Biochemistry and Cell Biology

Stony Brook University

December 2016

Stony Brook University
The Graduate School

Andrew Scortea

We, the thesis committee for the above candidate for the
Master of Science degree, hereby recommend
acceptance of this thesis.

_____
Benjamin Martin – Thesis Advisor
Assistant Professor, Department of Biochemistry and Cell Biology

_____
David Matus – Second Reader
Assistant Professor, Department of Biochemistry and Cell Biology

This thesis is accepted by the Graduate School

_____
Charles Taber
Dean of the Graduate School

Abstract of the Thesis

PhiC31 integrase as a More Efficient Method of Transgenesis in Zebrafish

by

Andrew Scortea

Master of Science

in

Biochemistry and Cell Biology

Stony Brook University

2016

Abstract:

Zebrafish transgenesis is conventionally performed via the tol2 transposable element system as it is the most straightforward, providing quick and easy transgenesis. Despite being commonly used in the zebrafish community, tol2 mediated transgenesis has issues that cause workflow slowdowns and delays. The biggest of these issues are the random nonspecific integrations which can occur multiple times within a genome. In order to generate a stable transgenic line with a single insertion, one must perform several generations of outcrosses and subsequent screening. Thus, while tol2 transgenesis is effective at generating founder animals, a considerable amount of time and labor is required later to isolate a single insertion line. A promising alternative is the use of phiC31 integrase, which follows the same workflow yet provides a single specific integration, removing the need for subsequent outcrossing. Unfortunately, this does not appear to be as easy and revolutionary as would be expected. The efficiency of phiC31 integrase was known to be lower than tol2, yet out of six different plasmids injected numerous times using different aliquots of phiC31 integrase mRNA, no positive insertion has ever been found. Although insertion appeared evident at first, further experiments showed that this was not the case. This method may still be useful but future evaluations will need to be conducted.

Table of Contents

List of Figures

Abbreviations

ITR	Inverted Terminal Repeats
attP	Attachment site phage
attB	Attachment site bacterium
cmlc2	Cardiac myosin light chain 2
ubi	Ubiquitin
hsp70	Heat shock protein 70
TCFDC	TCF transcription factor deletion
SPT	Spade tail gene
EGFP	Enhanced green fluorescent protein
sfGFP	Super folder green fluorescent protein
GFP	Green fluorescent protein
HS	Heat shock
hpf	Hours post fertilization
dpf	Days post fertilization
WT	Wild-type
attL	Attachment recombination site left
attR	Attachment recombination site right
bp	Base pairs
PCR	Polymerase chain reaction
DNTP	Deoxynucleotide triphosphate

1. Introduction:

Transgenesis is the ability to incorporate a gene derived from other organisms, or even synthesized, into another organism of choice. It is a useful tool for studying developmental biology, and used in many model organisms including yeast, drosophila, mice, and zebrafish. One such experimental application is the study of stem cell differentiation and fate decisions in the zebrafish tail bud (Martin, B. L., et al. 2012). Transgenic experiments are especially useful in zebrafish as the embryo is transparent during development. This allows for easy visualization of fluorescence of the embryo, and visual access to internal organ systems for fluorescence within the embryo. Transgenesis experiments have been performed throughout the years using a variety of methods each possessing their own strengths and weaknesses. The original method of zebrafish transgenesis was meganuclease mediated transgenesis, which was followed by the more effective tol2 system in zebrafish. A promising new method of transgenesis is the phiC31 integrase system, which has been demonstrated effectively in *Drosophila*, and has been used in a limited fashion in zebrafish (Mosimann, C., et al. 2013).

1.1 Meganuclease Transgenesis:

In order to break DNA to allow for integration, researchers initially used the meganuclease transgenesis method. This method utilizes the endonuclease activity of the meganuclease, which has very large recognition sequences, to perform the transgenesis (Silva, G., et al. 2011). The huge recognition sequences allow for very specific cleavage of phosphodiester bonds thus creating a double stranded break. Initially this was performed with naturally occurring meganucleases, which would, by chance, need to have a recognition site within the target genome. This means the transgenesis is consistent, but not specific. The method works differently in fish however, as the endonuclease used in these experiments, IsceI, does not

have recognition sites within the fish genome (Thermes, V., et al 2002). Therefore, the endonuclease is used to cleave the transgene out of the plasmid, leaving a linear fragment to be integrated into random double stranded breaks. New methods allow for engineered meganucleases with nearly any desired target site (Silva, G., et al. 2011). This, in turn, allows the addition of a gene of choice to be integrated into any site via homologous recombination (Silva, G., et al. 2011). Meganuclease activity occurs at very low efficiency with only 10^{-6} to 10^{-9} of all treated cells being transformed (Silva, G., et al. 2011). Although this leads to a very low rate of integration when used for transgenic purposes, the gene can be consistently placed within the cleavage site of the meganuclease. It stands to reason that due to this rate of integration, it will take many attempts to finally end up with a usable transgenic organism. Low efficiency would generally be a huge weakness, but with meganuclease transgenesis it tends to allow for single insertions. Despite lacking this single insertion ability, tol2 transgenesis is more commonly used.

1.2 Tol2 Transposase Transgenesis:

Tol2 is the most widely used system for transgenesis in zebrafish, due to its many advantages over previous methods. Tol2 is a transposon based system, as opposed to the endonuclease based meganuclease system. Tol2 requires the use of a specially prepared vector containing the gene, or desired sequence, flanked by tol2 inverted terminal repeats (ITR) (Clark, K. J., et al. 2011). The ITR are the targeting mechanism for this system, normally acted upon by the transposase protein contained between the ITRs (Clark, K. J., et al. 2011). The second requirement for this system is the tol2 mRNA of which specific synthetic mRNA exists, which is synthesized to be more stable than other forms of transposon mRNA (Clark, K. J., et al. 2011). The tol2 vector and tol2 mRNA are mixed and co-injected into the zebrafish at the single cell

stage. It is possible to inject at later stages, but it is not ideal and greatly reduces integration rates (Clark, K. J., et al. 2011). Once injected, the F0 generation is observed via microscopy for a fluorescent reporter with mosaic expression (Clark, K. J., et al. 2011). This F0 generation is grown to sexual maturity and crossed with wild-type zebrafish to produce germline transmission screenable offspring in the F1 generation (Clark, K. J., et al. 2011). Due to mosaic integration of the transgene, the offspring will not be generated with expected Mendelian ratios (Clark, K. J., et al. 2011). Furthermore, the amount of expression from a transgene can vary based on numerous integration events in the F0 generation (Clark, K. J., et al. 2011). From the F1 generation onward, gene inheritance will be Mendelian in nature. At this point the offspring can be grown and crossed with wild-type animals to produce later generations with the desired amount of expression (Clark, K. J., et al. 2011). Depending on the desired use, this animal would most likely be crossed until a single insertion remains. This process, though, is time consuming considering the period of growth to sexual maturity of a zebrafish is approximately 3 months. This system requires less work on the frontend to create the transgenic zebrafish but significantly more work afterwards to cross successive generations of fish until only a single insertion is present. Thus, there is less work to do upfront with tol2, but significantly more work after insertion has occurred. Tol2 provides the advantage of high efficiency, whereas the meganucleases provide low efficiency, which generally results in a single insertion. There is a newer, promising third option which takes the strengths of both methods without the weaknesses of either, which utilizes phiC31 integrase as an insertion method.

1.3 PhiC31 Integrase Transgenesis:

PhiC31 phage genome encodes for an integrase which has been gaining popularity in zebrafish transgenesis experiments (Lister, J. A. 2011). The phiC31 integrase works on a similar

principle to the tol2 system and has a similar workflow. It requires slightly more work in the beginning as one needs to generate fish with integration sites, but this work has been performed already by Mosimann et al. The tol2 system was utilized to create these initial pDestTol2CG2_attp integration sites (Mosimann, C., et al. 2013). They contain the attP (attachment site phage) site at which the recombination will occur, and an EGFP gene driven by the cmlc2 (cardiac myosin light chain 2) promoter which acts as a positive identification for zebrafish containing the integration site (Mosimann, C., et al. 2013). These integration site positive fish are co-injected with the pDestattB transgene plasmid and the phiC31 integrase mRNA at the single cell stage. The plasmid contains the attB (attachment site bacterium) site, which recombines with the attP site in the genome, and the transgene cassette of choice (Mosimann, C., et al. 2013). The phiC31 system solves issues pertaining to the megnuclease and tol2 systems. It has high specificity, and single insertion events which are of vital importance to promoter activity experiments. The system should remove this variable easily and produce transgenic fish lines with similar gene expression levels. This removes the need for crossing transgenic offspring multiple times to remove redundant insertions, and therefore saves considerable time considering how long zebrafish take to grow to sexual maturity. It also means that many of the same line can be produced concurrently and each crossed with other fish lines for various experiments saving even more time. After the point of injection, the experiment workflow is identical to other transgenic systems in that one grows the fish and screens for germline genomic integration.

1.4 Evaluating PhiC31 Integrase Transgenesis:

PhiC31 transgenesis efficiency will be evaluated via the use of several plasmids with two different promoters, ubiquitin B (Ubi) and heat shock protein 70 (hsp70). The heat-shock (HS)

promoter plasmids used were Hsp70:NLSkikume, Hsp70:TCFDC-p2a- NLSkikume:, and Hsp70:SPT-p2a- NLSkikme with the latter two providing observable developmental phenotypes. The constitutive ubiquitin promoter plasmids used were Ubi:mem-mcherry, Ubi:EGFP, and Ubi:SfGFP-lifeact which contain constructs designed for imaging cell behavior. Positive insertion was never found for any of these plasmids although at first insertion appeared to exist. Had these plasmids successfully been inserted into the zebrafish, many more experiments could have been performed including; comparison to tol2, CRISPR/Cas9 transgenesis, tail bud stem cell fate map, and 3D imaging via light-sheet microscopy. PhiC31 showed promise as a new efficient method for transgenesis that follows the same basic procedure as older methods yet provides numerous advantages, but as transgenesis never occurred, this future is doubtful. Future experiments could be performed to elucidate the cause of transgenesis failure.

2. Results:

2.1 Cloning Plasmids:

Evaluation of phiC31 began by cloning various promoters and genes into a vector containing the attB recombination site. An example workflow can be seen in figure 1. The promoters and genes chosen were Hsp70:NLSkikume Hsp70:TCFDC-p2a- NLSkikume, Hsp70:SPT-p2a- NLSkikme, Ubi:mem-mcherry, Ubi:EGFP, and ubi:sfGFP-lifeact. The Hsp70:NLSkikume cassette was created initially and followed through with until injection. Afterwards it was used as the vector for cloning the TCFDC and SPT genes into. These were chosen as they produced well characterized developmental phenotypes which would affect tail morphology, and which have non uniform expression patterns throughout the embryo. TCFDC generally shows expression throughout the somites and less in the tailbud and head. SPT generally shows expression in the tail bud. A colony screen was performed to clone the plasmid and the correct plasmid contents were checked via restriction digest. Initially the Hsp70:TCFDC-p2a- NLSkikume, and Hsp70:SPT-p2a- NLSkikme plasmids were not being properly cloned. After numerous digestions the expected bands were never seen. This was corrected by creating a new starting Hsp70:NLSkikume plasmid with the gene being sourced from a different plasmid. It was discovered via sequencing that the Pac1 site within was malformed hence why it is not being cleaved, and linearized for future cloning. Cloning the TCFDC and SPT genes into this new hsp70 plasmid resulted in the expected bands being observed. The hsp70 promoter regulated genes were sequenced to ensure they were indeed present in the plasmid. The ubiquitin promoter regulated plasmids were created and analyzed in the same way to ensure the plasmid contents were correct.

2.2 Plasmid Injections:

The plasmids were injected into the embryo at the single cell, or in the yolk just under the cells during the two cell stage. A subset of the hsp70 promoted plasmid injected fish, usually 10-15 fish, were heat shocked at 6 hours post fertilization (hpf) to observe fluorescence in order to indicate the plasmid was working, in the case of TCFDC and SPT genes the developmental phenotypes, and gene specific localization were also observed. The phenotypes, and localization can be seen in figure 2. The SPT injected fish showed weaker, and overall fewer fish with expression than the hsp70, or TCFDC injected fish. These fish were observed and imaged 24h later and fluorescence and developmental phenotypes were observed consistently for each plasmid. A second heat shock using a larger subset of fish, usually 20-40 fish, was performed at 2 days post fertilization (dpf) to indicate integration of the plasmid into the zebrafish genome, and were observed 24hrs after. This suggests integration because by this point the plasmid should have been destroyed by the embryo's cells, so any visible fluorescence should not be expressing off the plasmid. Fluorescence and developmental phenotypes were consistently observed in these fish as well as seen in figure 3. Heat shocks occurring 3-5dpf were also performed and still the phenotypic expression was observed. Injection rounds that appeared to have strong mosaic expression were grown from the remaining fish that had not been heat shocked after incubating for a week. These fish were allowed to grow for 2-4 months, and then crossed with WT to look for germline transmission. Positive offspring were never discovered despite several rounds of injection of each plasmid. Offspring with the cmlc2:GFP heart marker were found in expected Mendelian ratios, but none were ever found expressing the transgene phenotypes.

Some of the hsp70 promoted plasmids that were injected appeared to become integrated into zebrafish without the cmlc2:GFP heart marker. These were very rarely seen but did occur at during the 2dpf heat shock of TCFDC, as can be seen in figure 4. This appeared to occur with each plasmid, both hsp70 and ubi regulated transgenes.

While the hsp70 transgene fish were growing to maturity the ubi transgene plasmids were generated and injected. These were observed similarly at 6hpf, and at 2dpf, although without heat shocks being necessary. Since nothing needed to be done to induce expression, they were observed randomly multiple days after the 2dpf check to see if they were still expressing, which they generally were, even at 5-6dpf. The ubi transgene fish were also grown to maturity whenever strong mosaic expression was observed. The sfGFP plasmids showed consistently weaker expression than the other two, especially the mem mcherry plasmid which had the strongest expression. Ubi transgene expression at 2dpf can be seen in figure 3. These fish were also grown 2-4 months to maturity and then crossed with WT to observe for offspring displaying expression. These too did not provide any offspring with expression despite numerous crosses. Not all fish were able to be crossed, as some were just not laying. One clutch of sfGFP offspring also displayed odd development. The tails were shortened and slightly curled, the tailfins were small or non-existent, and this occurred to each fish in the clutch. This did not seem to impact their ability to survive, as they grew to maturity, nor did it affect their ability to lay, although they were still negative for expressive offspring.

2.3 Checking for integration via PCR:

Since none of the offspring appeared to be displaying expression after months of growing and crossing, newly injected fish were verified for integration. A polymerase chain reaction (PCR) was performed for the left attachment site recombination (attL), and right attachment site

recombination (attR) sites that are formed after the attp site recombines with the attb site. A separate PCR was performed for each recombination site on every injected fish, and amplified bands were expected at 360 base pairs (bp) for the attL site, and 185bp for the attR site. Despite several fish being analyzed this way, no bands were observed on the gel as can be seen in figure 5.

3. Discussion:

Use of phiC31 integrase as the new standard for transgenesis in zebrafish seemed promising at first but unfortunately the experiments were derailed due lack of integration into the zebrafish genome. This was due to the lengthy growth time of zebrafish, as they need to be grown to maturity to search for positive integrations. The plasmids themselves were functional but either the phiC31 integrase mRNA, or the recombination sites themselves were non-functional. It is also possible that the integration rate was much lower than expected. This could be tested by continuing injections, and screening of offspring until integration is found. Time was an issue in this regard as this is a lengthy process and still would not guarantee results. It is also discouraging that not a single fish had been discovered as positive for integration or germline transmission, leading one to believe something must be fixed during the procedure.

Determining integration was challenging at first because since the plasmids were working, and producing fluorescence, and in the case of TCFDC and SPT producing developmental phenotypes, and the proper localization, which seems to mean the integration is occurring. The 6hpf HS was meant to see if the plasmid was working, which for each plasmid was true. The 2dpf HS however was meant to see if the phiC31 integrase itself was causing integration. The rational was that during the 6hpf HS it was possible for the plasmid to still be around thus being the cause of expression during imaging, and that by the HS at 2dpf the plasmid should be degraded already. This means that the HS at 2dpf would only be causing expression via the integrated hsp70 promoter, and all integration discovered would be just due to the phiC31 integrase activity and random insertions. This was also true for the ubi transgene plasmids which did not need to be heat shocked. Since these didn't need to be heat shocked they were able to be observed every day after injection, and each day the expression remained. Thus, leading to the

conclusion that integration had occurred, but unfortunately this did not seem to be the case in the end. Although this is an odd result since the plasmid should be degraded, as there is no evidence of plasmid DNA remaining after two days in zebrafish. It is unclear what is occurring in this case. There is evidence of plasmid DNA remaining in other animals, such as mice, for up to two years after injection so there may be some possibility here (Armengol, G., et al. 2004).

An integration event validation was performed by the group, Mossiman et al, which had produced the fish used in the experiments. The same PCR procedure was performed on injected embryos using the same primer sequences, yet the expected bands for each recombination site were not discovered. This was performed after months of attempted injections and screening looking for positive integrations. In hindsight this probably should have been performed alongside the HS and microscopic inspection of the 2dpf embryo. Had zebrafish not required such a long time to grow and cross, this would not have been as big a derail as it turned out to be.

Ultimately the delays in this experiment, and short research time available in a master's program, caused a lack of desirable results. By the time something very wrong with the experiment was found, there was not much time to figure it out, fix it, then grow up more fish. Had more time been available the error in the experiment could easily be discovered. There are some future experiments that would have been performed, both in continuing phiC31 integrase analysis, and in testing other methods such as CRISPR/Cas9, and even going back to tol2. The next step would be to see if the phiC31 integrase mRNA is actually working in the first place. This is suspect because the plasmids themselves are definitely working, otherwise no expression would have been seen. It's possible the error is occurring sometime between the injection and microscopic observation. It could also be that, as with the pac1 site sequence not being quite right, either the attp or attb sites are mutated leading to inability to recombine. If this were the

case it would most likely be the attb site causing issue, as since the fish were obtained from a group who had published a paper with these same fish, they should be fine. A sequencing of the plasmids, and fish, recombination sites used should be able to tell if either of these sequences are incorrect.

Had this still not produced results, the experiment would probably have shifted focus to CRISPR/Cas9. This system is still a very good and useful system for transgenesis, but it would still not be as ideal as phiC31 integrase transgenesis. CRISPR/Cas9 works by utilizing a guide RNA to target a specific location in the genome and induce a double strand break (Cong, L., et al. 2013). This can be used alone to disrupt a gene, or by providing additional DNA sequences with homology adjacent to the cleaved site, transgenes can be incorporated via the homology dependent repair mechanism. Due to the specific nature of CRISPR/Cas9 it would be a viable alternative to phiC31integrase in zebrafish transgenics. It would also provide the ability to add multiple transgenes at the same time in multiple locations throughout the genome. PhiC31 integrase is not able to do this, and although more complicated a process, CRISPR/Cas9 should definitely be considered wherever phiC31 integrase would be considered.

Despite both methods being very good at creating single, specific, insertions of transgenes phiC31 integrase should be the first choice unless more complicated transgenesis is required. Since CRISPR/Cas9 requires guide RNA to be made each time, more work would be required to get multiple different transgenic zebrafish. PhiC31integrase would require just a simple cloning into an already prepared vector, ready for injection. Since the fish already contain the attp site, the procedure should be easy and quick. The researcher could easily inject numerous embryos, and many with different plasmids as well, to get several transgenic zebrafish lines. Although this is slightly more complicated than just cloning a vector for phiC31 integrase,

it would still be significantly less work than the tol2 system of transgenesis. It would also produce fish with insertions just as precise as phiC31 integrase. There would be a single, specific insertion, which is one of the most important aspects of moving from tol2. A disadvantage with phiC31 integrase mediated transgenesis however is that only one plasmid insertion can be performed per fish. Due to the fish having a single attp site, and the attp-attb recombination changing the site, there can be no further transgenesis via phiC31 integrase. No future genes can be added, CRISPR/Cas9 does not have this issue as a new guide RNA can be made and a new procedure performed on the already transgenic embryo. Although this could presumably also just be performed on an embryo which was made transgenic via any transgenesis method. Another option would be just to make two separate transgenic lines and cross them together.

Once a transgenic zebrafish had been found the next step would be to utilize it in an experiment to see how viable the phiC31 integrase method truly is. These experiments would be studying the tail bud stem cell migration patterns in various ways. The first would be to simply image them and track cell movements. This could be accomplished by photoconverting the kikume fluorescent protein from green to red via a laser followed by imaging on a confocal. This allows tracking of individual cells as the cells would stand out amongst the green cells. The images could then be analyzed to determine where each cell starts, and where it travels during development. This allows for the creation of a stem cell fate map of the tail bud which would have been a quite useful reference for other researchers. A second use for the created transgenic lines would have been 3D imaging via the tiling light-sheet selective plane illumination microscopy. This would allow for a similar experiment as previously but this time in 3D. The light sheet microscopy allows for visualization and tracking of every cell in a video which would

have provided much insight into the cellular movements (Fu. Q, et al. 2016). It is unfortunate this was not able to be studied as this would have provided very interesting results.

4. Materials and Methods:

4.1 Fish Stocks:

Fish utilized in experiments were Attp2B and Attp2A transgenic zebrafish described by Mosimann et al 2013.

4.2 Ligation Independent Cloning (LIC) into the PhiC31 Vector:

All the inserts were PCR amplified using the following 50uL reaction mixture: 5X phusion buffer 10uL, DNTPS 1uL, forward primer 1.25uL, reverse primer 1.25uL, template 1uL, ddH2O 35.5uL, and phusion enzyme 1uL with the (LIC150 cycle). Insert PCR amplicons were run on a 1% agarose gel then extracted with the EZNA gel extraction kit. LIC is performed via T4 reaction on the insert (15uL gel purified PCR fragment, 2uL T4 buffer, 2uL 25mM dCTP, 1uL 100mM DTT, 0.2uL T4 pol) and the vector (15uL gel purified PCR fragment, 2uL T4 buffer, 2uL 25mM dGTP, 1uL 100mM DTT, 0.2uL T4 pol) and both reactions were incubated for 40min at room temperature then heated for 20min at 75C. Annealing was performed with 2uL of vector and 6uL of insert at room temp for 30min. DH5a cells were transformed with 2uL of the annealed plasmid and incubated on ice for 5min. The cells were incubated with 250mL of LB broth for 1hr, then plated and left overnight at 37C.

4.3 Plasmid purification:

A colony PCR was performed on the transformed DH5a colonies using the following 20uL reaction mixture: 10X taq buffer 2uL, DNTPs 0.4uL, forward primer 0.5uL, reverse primer 0.5uL, template 0ul (colonies picked), ddH2O 16.2uL, taq enzyme 0.4uL with the (col240 cycle). PCR amplicons were ran on 1% gel then colony lanes were chosen for plasmid mini prep. Chosen colonies were grown in 10mL LB broth with 10uL amp and incubated overnight at 37C.

The mini prep was performed with the EZNA Plasmid DNA mini kit I. Concentration of purified plasmid was measured on a spectrophotometer. A restriction digest, varying based on insert, was then performed on 1ug of plasmid, to verify insertion into the plasmid. A phenol chloroform extraction was performed to precipitate and purify the plasmid. Concentration was measured again. Plasmids were sequenced with an 8uL reaction mixture: 1uL 4pmol primer, 200-300ug of plasmid based on measured concentration, and then brought up to 8uL with H2O.

4.4 Injections:

Injection was performed by injecting approximately 1nl of the injection mixture: 25ug/ml plasmid 5ul, 25ug/ml phiC31 integrase mRNA 5ul into each embryo. Reagents were thawed on ice to preserve mRNA. The mixture was injected as close to directly into the embryo at the single cell stage, or just under the embryo in the yolk at the 2 cell stage. Embryos were stored at 28.5C until HS or microscopic observation.

4.5 Heat Shock:

Heat shocks were performed for 30 min at 40C for all HS promoter plasmids. They were performed mostly at 2 different times, shield stage HS at 6 hours post fertilization, or at 2dpf. Heat shocks were occasionally performed between 3-5 dpf as well. The 6hpf HS was to ensure the plasmid worked, and the 2dpf HS was meant to determine if the insertion worked. The 3-5dpf heat shocks served the same purpose as the 2dpf HS.

4.6 Insertion verification PCR:

The embryo was extracted from the corion and lysed with 50uL lysis buffer for 20min at 94C, and 5ul of 10mg/ml Proteinase K for 1hr at 55C followed by 20 min at 94C. The lysed embryo was PCR amplified with the following primers:

attL: 5’-CACCTCTCGAGGTACCTGCAGTACTGACG-3’ Forward

5’-GCTGACATGCCCGCCGTGACCG-3’ Reverse

attR: 5’-CACCGTCGACGATGTAGGTCACGGTCTCG-3’ Forward

5’-GTCGACGACTCTAGATCG-3’ Reverse

PCR was performed for both attR and attL individually in the 20ul reaction mixture: 10X taq buffer 2ul, DNTPs 0.4ul, forward primer 0.5ul, reverse primer 0.5ul, template 2ul, ddH2O 14.2ul, taq enzyme 0.4ul with the (col240 cycle). The amplicons were gel analyzed for bands at 185bp and 360bp indicating presence of attR and attL respectively.

4.7 Microscopy:

Injected fish were observed 24hrs after each HS for the hsp70 transgene plasmids, or 24hrs and 2dpf post fertilization for ubiquitin transgene plasmids, on a Leica M165 FC fluorescence microscope, for cmlc2:GFP heart promoter, and expression of fluorescence or developmental phenotype. Embryo images were captured in embryo media or knocked out with tricane then fixed in glycerol.

4.8 Fish Crossing:

Fish were set-up for mating between 3pm-8pm the day before they lay. Embryos were collected between 9am-11am and incubated at 28.5C until need for use. The attp2A, and attp2B fish were outcrossed with WT fish to produce a heterozygous offspring for injection.

The F1 generation fish were screened by crossing with either WT or negative insertion fish from the same clutch. They were observed microscopically for the cmlc2 heart promoter driven GFP indicating presence of the attP site, thus ability to insert injected plasmid. Fish that

expressed the trans gene and cmlc2:GFP would be designated positive for insertion and grown to maturity. Insertion negatives were used for mating or euthanized.

References:

Armengol, G., Ruiz L. M., Orduz, S. (2004) "The injection of plasmid DNA in mouse muscle results in lifelong persistence of DNA, gene expression, and humoral response" Molecular Biotechnology **27**(2): 109-118.

Clark, K. J., Urban, M. D., et al. (2011) "Transgenic Zebrafish Using Transposable Elements" Methods Cell Biology **104**: 137-149.

Cong, L., Ann Ran, F., et al. (2013) "Multiplex Genome Engineering Using CRISPR/Cas Systems" Science **339**(6121): 819-823.

Fu, Q., Martin, B. L., Matus, D. Q., Gao, L. (2016) "Imaging multicellular specimens with real-time optimized tiling light-sheet selective plane illumination microscopy" Nature Communications **7**: 11088.

Lister, J. A. (2011) "Use of phage PhiC31 integrase as a tool for zebrafish genome manipulation" Methods Cell Biology **104**: 195-208.

Martin, B. L., Kimelman, D. (2012). "Canonical Wnt Signaling Dynamically Controls Multiple Stem Cell Fate Decisions" Developmental Cell **22**: 223-232.

Mosimann, C., Puller, A., et al. (2013) "Site-Directed Zebrafish Transgenesis Into Single Landing Sites With the phiC31 Integrase System" Developmental Dynamics **242**: 949-963.

Munoz-Lopez, M. Garcia-Perez, J. L. (2010) "DNA Transposons: Nature and Applications in Genomics" Current Genomics **11**: 115-128.

Roberts, J. A., Miguel-Escalada, I., et al. (2014) "Targeted transgene integration overcomes variability of position effects in zebrafish" Development **141**:715-724.

Sander J. D., Joung, J. K. (2014) "CRISPR-Cas systems for editing, regulating and targeting genomes" Nature Biotechnology **31**:347-355.

Silva, G., Poirot L., et al. (2011) "Meganucleases and Other Tools for Targeted Genome Engineering: Perspectives and Challenges for Gene Therapy" Current Gene Therapy **11**: 11-27.

Thermes, V., Grabher, C., et al. (2002) "I-SceI meganuclease mediates highly efficient transgenesis in fish" Mechanics of Development **118**: 91-98

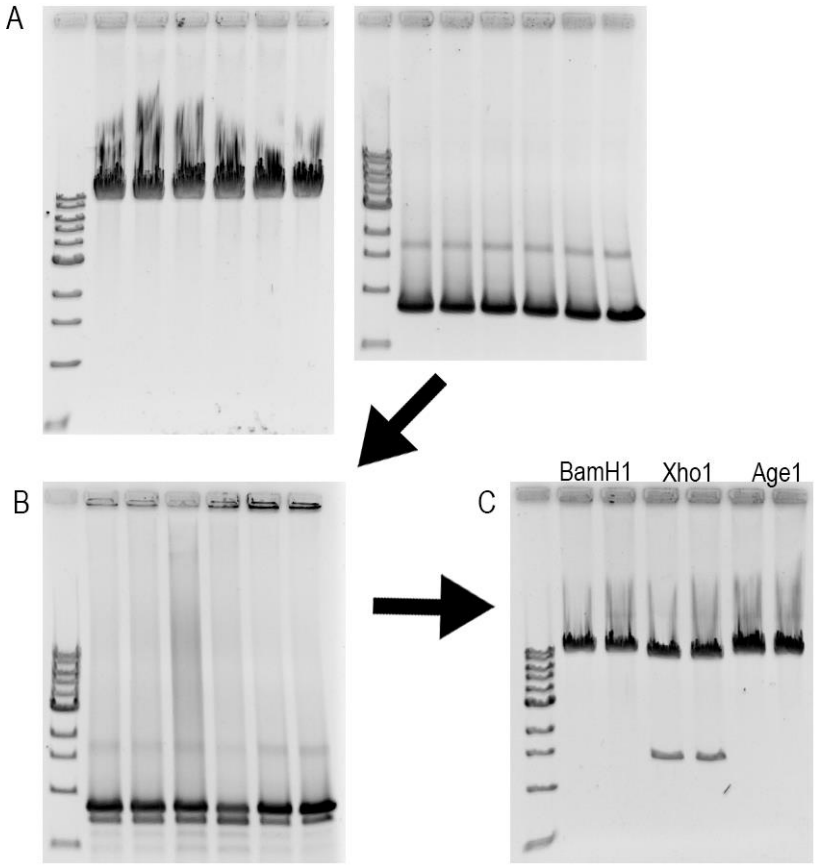

Figure 1. Example workflow displaying each step of the cloning process. Figure 1 displays the workflow used for cloning the mem-mcherry gene into the ubi promoter regulated plasmid. After step C the plasmid would be sequenced to ensure proper cloning, then finally phenol-chloroform extracted and precipitated. (a) Display of isolating the ub1 vector (left) via pac1 digestion, and isolating the mem-mcherry gene via PCR amplification. Both are then gel extracted and ligated together. (b) A colony PCR is performed to find ligated vector-gene plasmids which have transformed DH5a E. coli cells. Two bands are chosen and gel extracted, then their concentration is measured. (c) A restriction enzyme digestion is performed with a few enzymes and expected band drop outs are observed for consistency with expected bands.

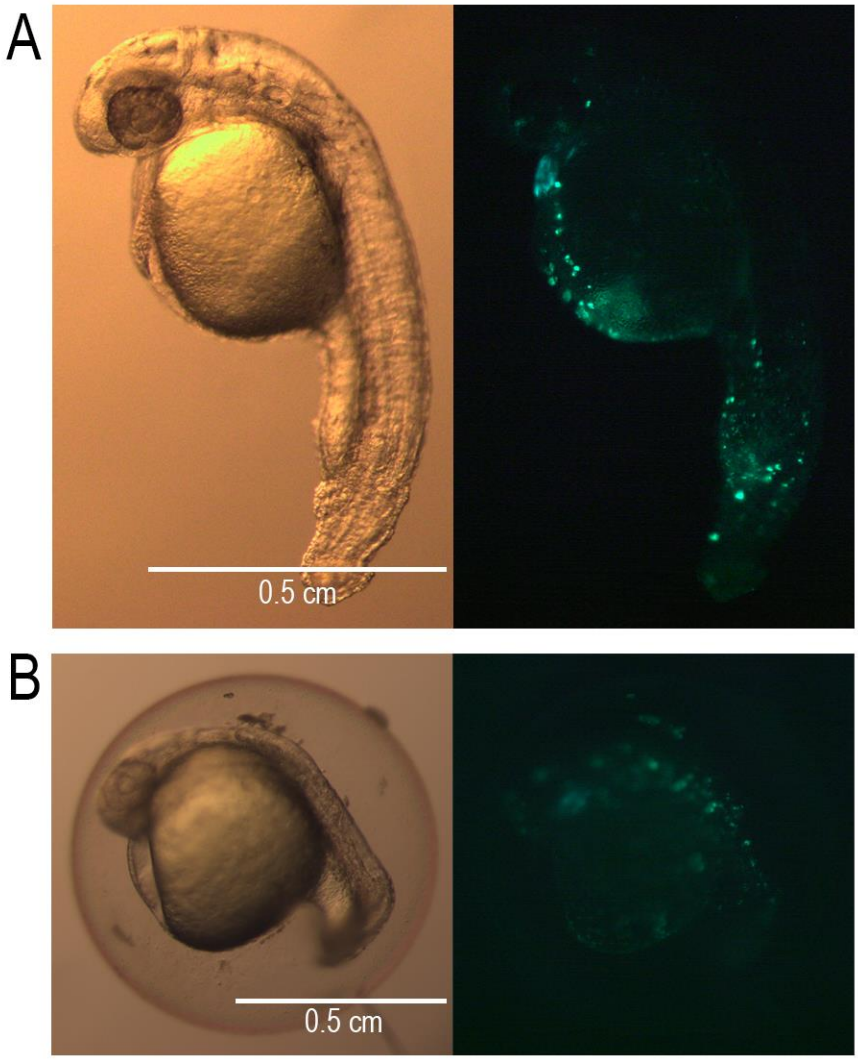

Figure 2. Developmental phenotypes, and localized expression of SPT and TCFDC. Figure 2 displays the phenotypes visible from the 6hpf HS, and the localized expression due to how the genes cause their phenotype. Scale bars are also in scale for the fluorescent images to the right of each. (a) SPT phenotype and tail bud localized expression. (b) TCFDC phenotype and somite localized expression.

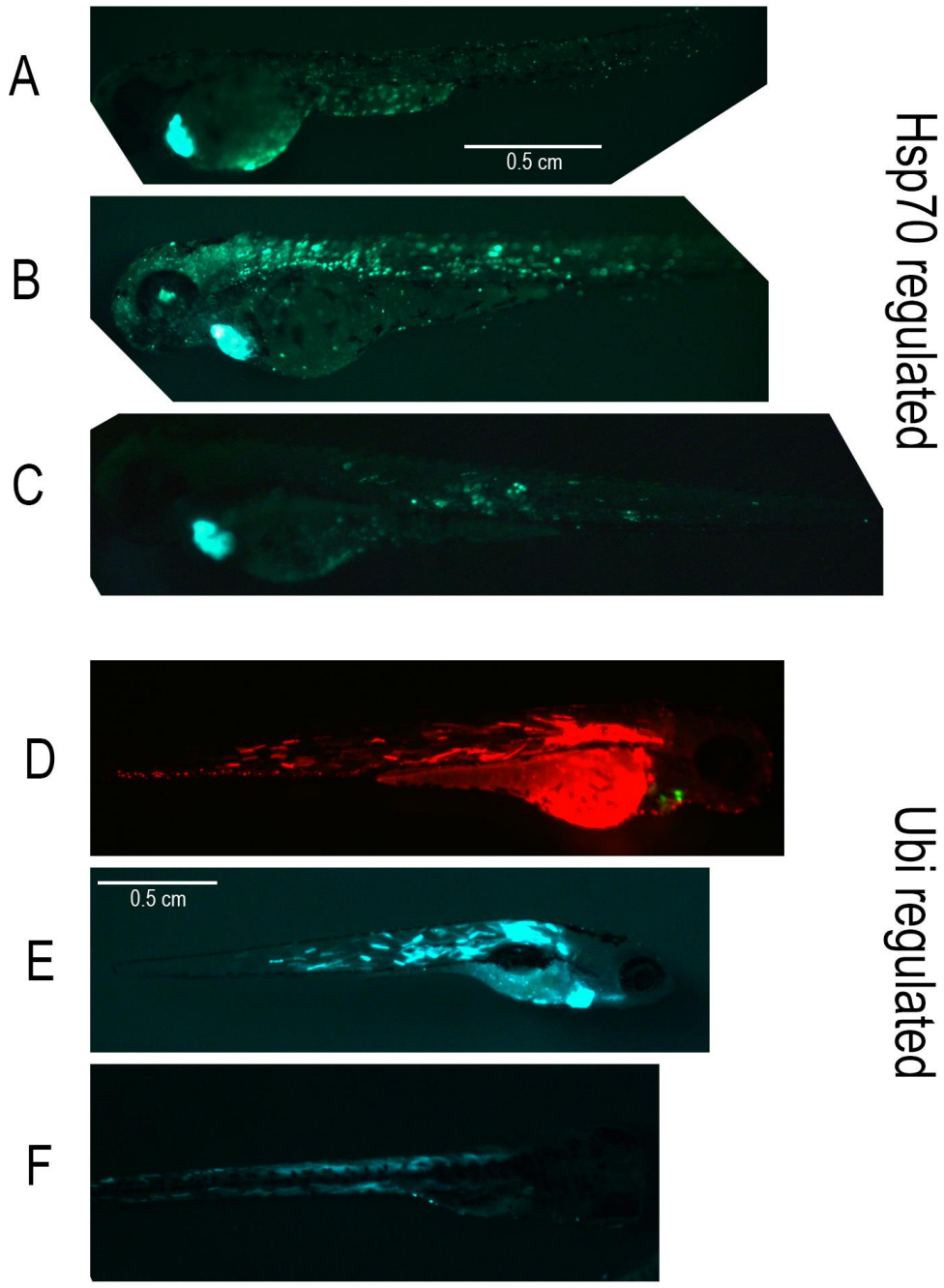

Figure 3. Representative plasmid expression at 2dpf for hsp70, and ubi regulated transgenes. Figure 3 displays samples with the most amount of expression seen for each plasmid. Some express much stronger than others. Each displays the cmlc2 heart marker. Scale bars for b, c, and d, are in scale with a. The same is true with e for f. (a) Hsp70 expression, which generally appeared very diffuse. (b) TCFDC expression was the strongest of the hsp70 transgenes. (c) SPT expression was generally weaker than the other hsp70 transgenes. (d) Expression of mem-mcherry, which in general had the strongest expression of any plasmid. The cmlc2 heart marker was imaged on a separate channel and overlayed for visibility. (e) EGFP expression was similar to mem-mcherry expression, but usually not as strong. (f) SfGFP had weak expression, like SPT usually. The cmlc2 heart marker is present, but not visible due to the angle of the fish.

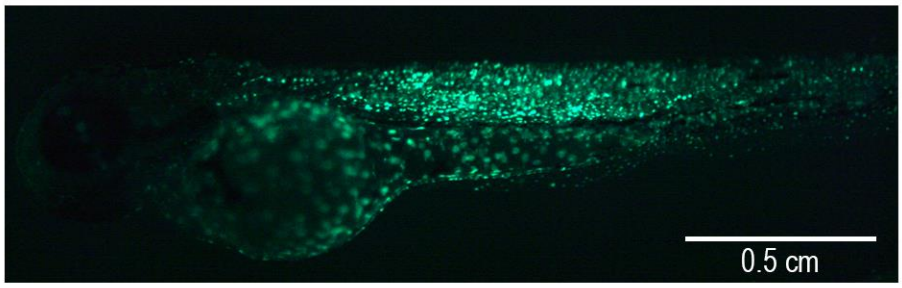

Figure 4. Expression of TCFDC at 2dpf without cmlc2 heart marker. Display of apparent integration and expression in a fish that should not be able to support integration via phiC31 integrase. Lack of cmlc2 heart marker indicates not attP site, therefore nowhere for the plasmid to recombine.

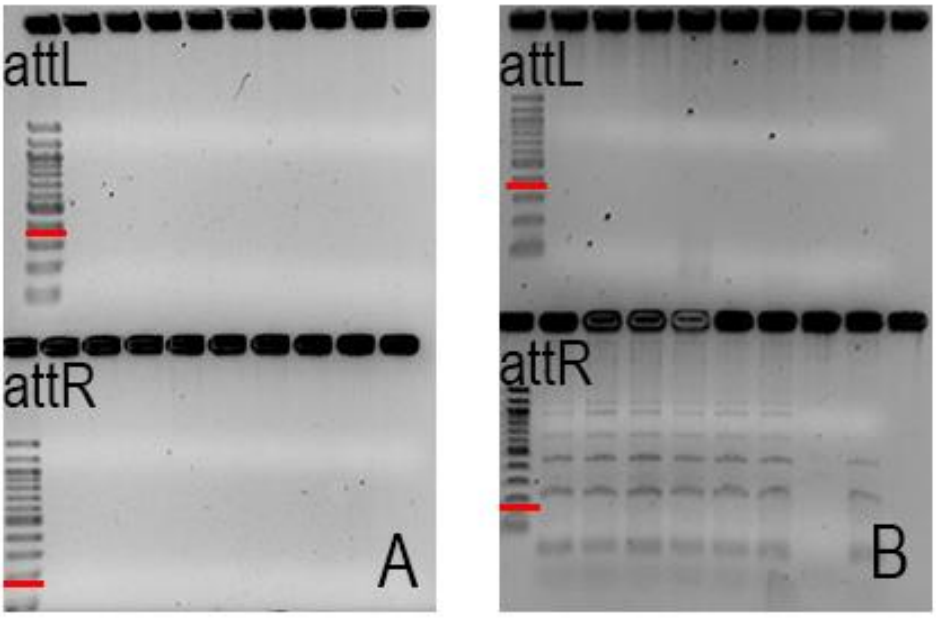

Figure 5. PCR Validation of hsp70 transgenes, and ubi Transgenes, did not result in bands expected for attL and attR recombination. Figure 3 displays the lack of bands which would indicate successful attL and attR recombination. Each lane is a different fish embryo, and red lines indicate where the expected band should be found. A band at 360bp expected for attL recombination, and a band at 180bp expected for attR recombination. (a) Representative figure of hsp70 transgenes displaying no bands. (b) Representative figure of ubi transgenes displaying only background bands.