

Full Length Research Paper

# Developing a robust red fluorescent transgenic line in *tanichthys albonubes*: A breakthrough in genetic engineering

Qing Jian, Min Chen, Jun-jie Bai\*, Peng Jiang, Jia-jia Fan, Xing Ye and Shi-ling Xia

Key Laboratory of Tropical and Subtropical Fishery Resource Application and Cultivation, Ministry of Agriculture, Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou, Guangdong 510380, China.

Accepted 16 October, 2023

White cloud mountain minnow (*Tanichthys albonubes*) is a small cyprinid oviparous fresh-water fish in Southern China. For ornamental purpose, the red fluorescent protein (RFP) transgenic *T. albonubes* was generated by microinjection of a pMYLZ2-RFP gene construct containing zebrafish myosin light polypeptide 2 (*mylz2*) promoter into the fertilized eggs, and a transgenic line was established. For this line, segregation of RFP gene followed Mendelian single-gene inheritance, confirming transgene integration into the *T. albonubes* genome at a single locus. Results of reverse transcriptase-polymerase chain reaction (RT-PCR) analyses showed that RFP transgene was expressed in almost all body tissues, revealing an ectopic transgene expression. Based on ordinary and quantitative PCR analyses, integration pattern of RFP gene was determined to be three copies and head-to-tail multimers. An approximately 1.0 kb fragment amplified from 5'-flanking sequence of transgenes was identified in wild-type *T. albonubes* genome, and bears no homologous sequence in the GenBank database. In the 3'-flanking region, an approximately 1.2 kb fragment with unidentified source was amplified which has 99% homology to enterobacteria phage P1. The above results well demonstrated the pMYLZ2-RFP constructs have being inserted as a stable DNA fragment in the *T. albonubes* genome and could be stably inherited and expressed in transgenic progenies.

**Key words:** Transgenic fish, *Tanichthys albonubes*, RFP, expression, flanking sequence, copy number.

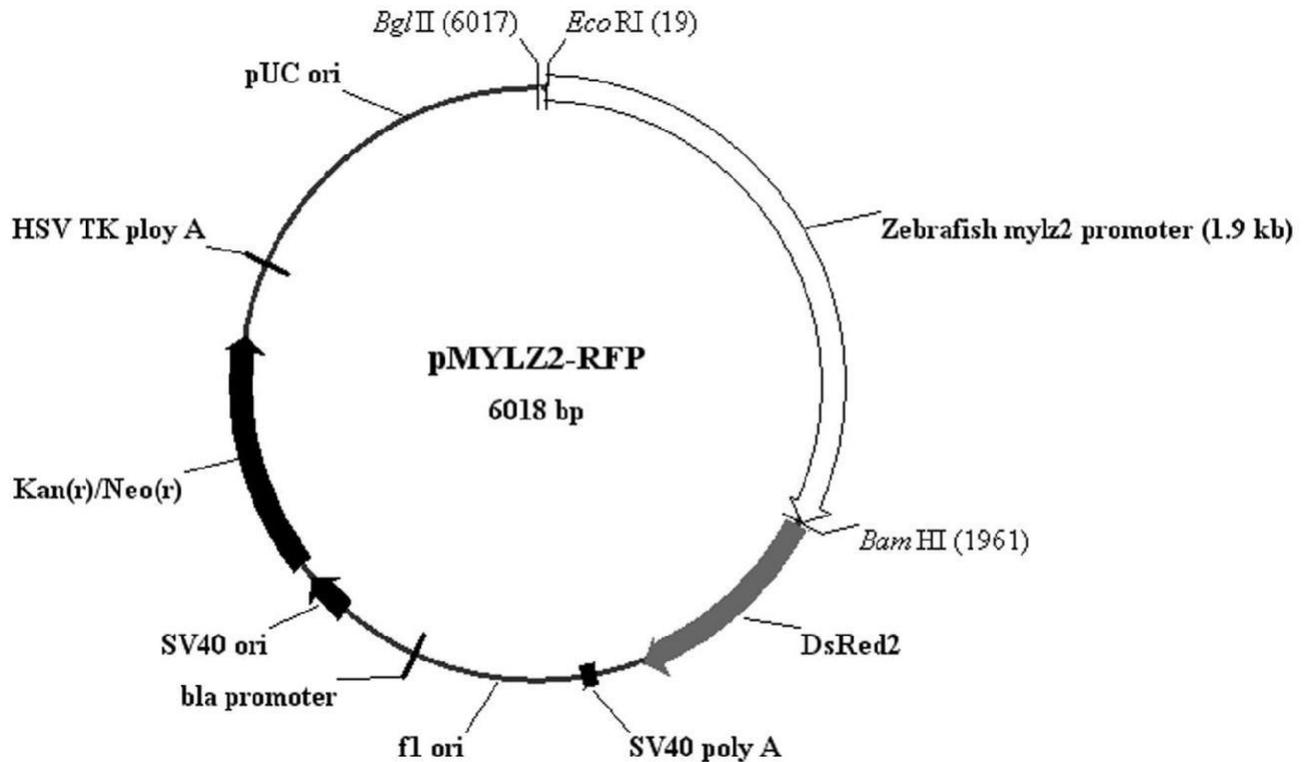
## INTRODUCTION

Fluorescent proteins originally isolated from marine invertebrates are widely used as noninvasive probes for cell and molecular biology (Matz et al., 1999; Stewart, 2006). In recent years, researchers opened up their novel application for ornamental fish by transgenic technology, such as the fluorescent transgenic zebrafish (*Danio rerio*) and Japanese medaka (*Oryzias latipes*) (Gong et al., 2003; Kinoshita, 2004). This novel application benefits from the expression of fluorescent proteins in a wide range of cell types without apparent toxic effects and the vivid fluorescent colors that are readily visible to unaided eyes (Liu et al., 1999; Shaner et al., 2004). *Tanichthys albonubes*, "white cloud mountain minnow", is a small cyprinid oviparous fresh-water fish in Southern China

(Liang et al., 2008), which have been cultured around the world as a popular aquarium fish (Yi et al., 2004). To develop new varieties of ornamental fish, our laboratory generated the red fluorescent protein (RFP) transgenic *T. albonubes* by microinjecting fertilized eggs with RFP gene under the control of zebrafish myosin light polypeptide 2 (*mylz2*) promoter. Germ-line transmitted transgenic *T. albonubes* was obtained, which displays a visible red fluorescent color instead of the original body color of wild-type fish.

To date, microinjection is still the most common method used for gene transfer in fish (Zbikowska, 2003). However, this classical transgenic approach is linked to the random and uncontrolled integration of transgenes into the host genome (Sin, 1997). The integration behavior would inevitably disrupt the neighboring sequences and may be accompanied by some mutations (Tzfira et al., 2004; Wilson et al., 2006). In addition, the introduced foreign DNA can be modified in many ways

\*Corresponding author. E-mail: [baijj2005@21cn.com](mailto:baijj2005@21cn.com). Tel: +86 20 81616129. Fax: +86 20 81616162.



**Figure 1.** Plasmid map of pMYLZ2-RFP for generating the red fluorescent transgenic *T. albonubes*.

before, during, and even after the integration process (Würtele et al., 2003). As a result, transgenic organisms often have unexpected and unstable traits. For example, transgenic expression may vary across different sites of transgenic integration, a phenomenon called “positional effects” (Lin et al., 1994; Iyengar et al., 1996). Therefore, in establishing a transgenic line, integration of transgenes into host genome and their stable transmission and expression in offspring are the most important issues (Kinoshita et al., 1996). Nevertheless, to our knowledge there is little published information on the detailed characterization of integration site and expression pattern in the fluorescent transgenic ornamental fish.

In the present study, we describe the generation of a germline RFP transgenic *T. albonubes* and attempt to investigate its characterization in terms of transgene inheritance, expression and integration pattern. Both upstream and downstream flanking regions of the insertion site were also examined. These obtained data provide a basis for analyzing whether the inserted foreign genes can be stably inherited and expressed in transgenic progenies.

## MATERIALS AND METHODS

### Fish maintenance and transgene constructs

*T. albonubes*, purchased from a local ornamental fish market (Guangzhou, China), were maintained in well-aerated glass aquaria

containing dechlorinated tap water under fluorescent lighting on a 12L: 12D cycle. Fish were fed flake food to satiation daily, supplemented with live brine shrimp nauplius (*Artemia salina*). Fertilized eggs of *T. albonubes* were collected from pair spawning in tanks and subjected to gene transfer. An approximate of 6.0 kb gene construct pMYLZ2-RFP used in this study (Figure 1) was made by insertion of a 1.9 kb zebrafish *mylz2* promoter into the multiple cloning site of pDsRed2-1 vector (BD Biosciences Clontech, USA). The zebrafish *mylz2* promoter was isolated by PCR (Jian et al., 2004) and digested with *EcoRI* and *BamHI*. For microinjection, the recombinant vector was linearized by *BglII* and then purified by PCR Clean-Up Kit (Roche, USA).

### Generation and screening of transgenic *T. albonubes*

The linear pMYLZ2-RFP expression vector was diluted in ddH<sub>2</sub>O to a final concentration of 100 ng/μl for microinjection. The detailed procedure was carried out as previously described (Jian et al., 2004). Injected eggs were incubated at 25-26°C in pure water under laboratory conditions. Living embryos were screened under a fluorescence microscope (with 2,072 filter combination) and all the selected RFP-positive embryos were raised to sexual maturity.

Transgenic founders (F<sub>0</sub>) were mated individually with wild-type fish to produce the F<sub>1</sub> generation. The F<sub>1</sub> transgenic individuals with red body color were again crossed to wild-type fish to produce the F<sub>2</sub> generation. F<sub>3</sub> were produced by crossing two F<sub>2</sub> red transgenic siblings. In the present study, we chose transgenic F<sub>3</sub> generation and its progenies (self- or test-cross) from a germline transgenic *T. albonubes* as research objects. To confirm the inheritance pattern of RFP gene in this line, the randomly chosen five F<sub>3</sub> individuals (3 female and 2 male) were crossed with wild-type mates. The frequencies of pMYLZ2-RFP transgene transmission to F<sub>4</sub> progeny were examined under the fluorescence microscope. In addition, dot

**Table 1.** Primers used in this study.

Primer name	Sequence (5'-3')	Purpose	Expected size (bp)
Red-F Red-R	CAGGACGGCTGCTTCATCTA GGTGTAGTCCTCGTTGTGGG	RT-PCR	288
$\beta$ -actin-F $\beta$ -actin-R	GCTTTGTCACTCCAGCACCG GGGCATCATCTCCAGCGAAT	RT-PCR	220
AP1 AP2	GTAATACGACTCACTATAGGGC ACTATAGGGCACGCGTGGT	Adaptor primers	—
szq1 szq2	AAATAAAGGGTGTAAACTGTCACTGGG CATTGTTTTGCTTCTCTTGATTTTTCCC	Upstream junction cloning	—
xzzt2 xzzt2	CCACCACTTCAAGAACTCTGTAGCACC CGACCTACACCGAACTGAGATACCTAC	Downstream junction cloning	—
Dsredf Dsredr	TCAGATTCATCCCATTCCAAGAC GACCGAGAGAGAAAGAGGAAAAACT	Quantitative PCR	127
Upstrf Upstrr	CACACACCCACACATAAAGATTTCAG CATCCGCTGCATAAAAAATACACT	Quantitative PCR	143
SZYS-SY SZYS-XY	CTGGGATACATTGAGCCT GTATCCGCCAGAGTTAG	Sequence detection	722
xzxl-sy xzxl-xy	CGCCTGGCACACTGGATACG CAGTTGGCAGAATCAGTATC	Sequence detection	911

blot method was employed to determine whether there exists silent integration in the non-transgenic offspring.

#### Detection of transgene expression by RT-PCR

Total RNA was extracted from muscle, liver, gonad, spleen, gill, intestine, swim bladder and heart of adult F<sub>3</sub> transgenic *T. albonubes* and non-transgenic siblings using Trizol Reagent (Invitrogen, USA). Reverse transcription was performed according to the protocol of PrimeScript™ 1st strand cDNA Synthesis Kit (TaKaRa, Japan). The synthesized cDNA (2  $\mu$ l) from each tissue was used as template in a 20  $\mu$ l PCR mixture. The PCR primers for RFP mRNA were Red-F and Red-R shown in Table 1. A region of the endogenous  $\beta$ -actin mRNA as an internal control was amplified using primers  $\beta$ -actin-F and  $\beta$ -actin-R (Table 1). PCR was performed with an initial denaturation of 3 min at 94°C and then 30 s of denaturation at 94°C, 30 s of annealing at 64°C (55°C for  $\beta$ -actin amplified), and 1 min of extension at 72°C for 34 (RFP gene) or 30 cycles ( $\beta$ -actin gene), followed by a final extension of 7 min at 72°C. The reacted products were electrophoresed on a 1.5% agarose gel.

#### PCR analysis of transgene arrangement

Single-primer xzzt2, szq1 and their combination (Table 1) were

used to amplify the pMYLZ2-RFP construct-to-construct junctions in F<sub>3</sub> transgenic and non-transgenic fish, respectively. An amount of 20  $\mu$ l PCR reaction contained 0.2  $\mu$ l template DNA, 0.4  $\mu$ l each of forward and reverse primers, 0.3  $\mu$ l dNTP (10  $\mu$ mol/l), 2.0  $\mu$ l 10  $\times$  PCR buffer (Mg<sup>2+</sup> 25 mmol/l), 1.0 U Taq polymerase. The reaction was denatured for 3 min at 94°C and then cycled 30 times at 94°C for 30 s, 55°C for 30 s and 72°C for 1 min, followed by a final extension of 7 min at 72°C.

#### Cloning of flanking regions of the transgene by genomic walking

A total of 100  $\mu$ l blood sample was collected from the heart ventricle of five F<sub>3</sub> transgenic siblings with a capillary tubule. Genomic DNA was isolated from the whole blood sample using Blood and Cell Culture DNA Kit (QIAGEN, Germany). Briefly, *EcoR* V and *Dra* I genome walker libraries were constructed including digestion, purification of genomic DNA, and ligation of genomic DNA to GenomeWalker Adaptors. To obtain the 5' flanking region of the transgene insertion locus, primers szq1 and szq2 (Table 1) were designed according to the known pMYLZ2-RFP sequence. The first-round PCR was performed using *EcoR* V library as template with Adaptor primer AP1 and primer szq2 (Table 1). An amount of 50  $\mu$ l PCR reaction contained 1  $\mu$ l template DNA, 1  $\mu$ l each primer (20  $\mu$ mol/l), 8  $\mu$ l dNTP (10  $\mu$ mol/l), 5.0  $\mu$ l 10  $\times$  PCR buffer (Mg<sup>2+</sup> 25 mmol/l), 0.5  $\mu$ l LA Taq polymerase and 33.5  $\mu$ l ddH<sub>2</sub>O. Amplification

conditions for the primary PCR were 25 s at 94°C; 3 min at 72°C for 7 cycles, 25 s at 94°C, 3 min at 67°C for 32 cycles, and 7 min at 67°C for the final extension. The second-round PCR with Adaptor primer AP2 and primer sqz1 was performed using a 1 in 50 dilution of the primary amplification as template. Secondary PCR had the same reaction composition and parameters as described above except amplification was allowed to proceed for 5 and 20 cycles. Also, the 3'-flanking region of the transgene was isolated employing the genome walking techniques as described above. The two rounds of PCR were performed using *Dra* I library as template with Adaptor primers AP1/AP2 and specific primers xzst1/xzst2 (Table 1). The resulting PCR products were electrophoresed on a 0.8% agarose gel and then purified using a gel extraction kit (Omega, USA). The purified PCR products were subcloned into the pMD18-T Vector and sequenced.

#### Determination of transgene copy number by quantitative PCR (qPCR)

Genomic DNA templates were isolated from fin samples of one F<sub>5</sub> and one F<sub>6</sub> transgenic hemizygotes, respectively. For the transgene amplification, primers Dsredf and Dsredr (Table 1) were used to amplify a 127-bp fragment from RFP gene. Based on the obtained 5'-integration junction sequence, primers Upstrf (locating at the 5'-flanking genomic sequence) and Upstrr (locating at the upstream transgene sequence) (Table 1) were designed to amplify a 143-bp single copy fragment as a reference. Both primer pairs were tested to ensure amplification of single discrete bands with no primer-dimers. PCR products were subcloned into the pMD18-T vector for sequencing. The purified plasmids were quantified by spectrophotometry using an Eppendorf BioPhotometer (Germany) and copy numbers were calculated by the formula: plasmid DNA (copies/ml) = [Concentration of plasmid DNA (g/ml) × 6.02 × 10<sup>23</sup> (copies/mol)] / Molecular weight of plasmid DNA (g/mol); Molecular weight of plasmid DNA (g/mol) = Length (bp) × 660 (g/mol/bp), where 6.02 × 10<sup>23</sup> is the Avogadro constant, 660 is the average molecular weight of one base-pair (double strand DNA). Serial dilutions were performed to give final concentrations ranging from 10<sup>4</sup> to 10<sup>9</sup> copies. Standard curves were generated for both target and internal control based on six 10-fold dilutions.

Quantitative PCR analysis was performed using Platinum<sup>®</sup> SYBR<sup>®</sup> Green qPCR SuperMix-UDG kit (Invitrogen, USA) and monitored on an ABI PRISM<sup>®</sup> 7300 Sequence Detection System (Applied Biosystems). The qPCR reactions were carried out in 96-well plates using a 25- $\mu$ l volume per well that included 12.5  $\mu$ l SYBR PCR Master Mix, 0.5  $\mu$ l ROX Reference Dye, 0.25  $\mu$ l each primer (20  $\mu$ mol/l), 0.4  $\mu$ l of plasmid or genomic DNA template and 11.1  $\mu$ l ddH<sub>2</sub>O. The qPCR reactions were performed under the following conditions: 2 min at 50°C, 2 min at 95°C and 42 cycles of 15 s at 95°C, 30 s at 60°C and 30 s at 72°C. Each reaction was repeated four times to ensure the reproducibility of the results. To prevent all possible false-positive results, including primer-dimer-induced positive results, all qPCR procedures were followed by melting curve analysis.

According to each standard curve, the C<sub>T</sub> values were converted to the initial copy number in the samples. Transgene copy number was calculated using the double standard curves method, following the formula: Transgene copy number = initial copies of target RFP gene / initial copies of internal reference sequence. Data were expressed as means  $\pm$  standard deviation (S.D.).

## RESULTS

### Generation of transgenic *T. albonubes*

In the spring-summer of 2003, the linear pMYLZ2-RFP

expression vector was microinjected into wild-type *T. albonubes* fertilized eggs at 1-2 cell stage. Out of a total of 719 eggs microinjected, 321 embryos (45%) survived after injected, 265 embryos hatched out normally and 83 fry had RFP expression under the fluorescence microscope (31% RFP-positive rate).

At 25 to 26°C, the survival *T. albonubes* embryos hatched out by 36 h post-fertilization (hpf), and no fluorescence signal was detected under the fluorescence microscope. RFP expression was initially observed at around 48 hpf. Subsequently, RFP expression gradually expanded to developing somite and skeletal muscles (Figure 2A, B). After 7 days post-hatching (dph), red fluorescent color was visible to unaided eyes under normal daylight (Figure 2C). Up to 30 dph, RFP expression increased continuously, and transgenic founders displayed the patched RFP expression in body muscles (Figure 2D, E).

### Germline transmission of pMYLZ2-RFP gene constructs

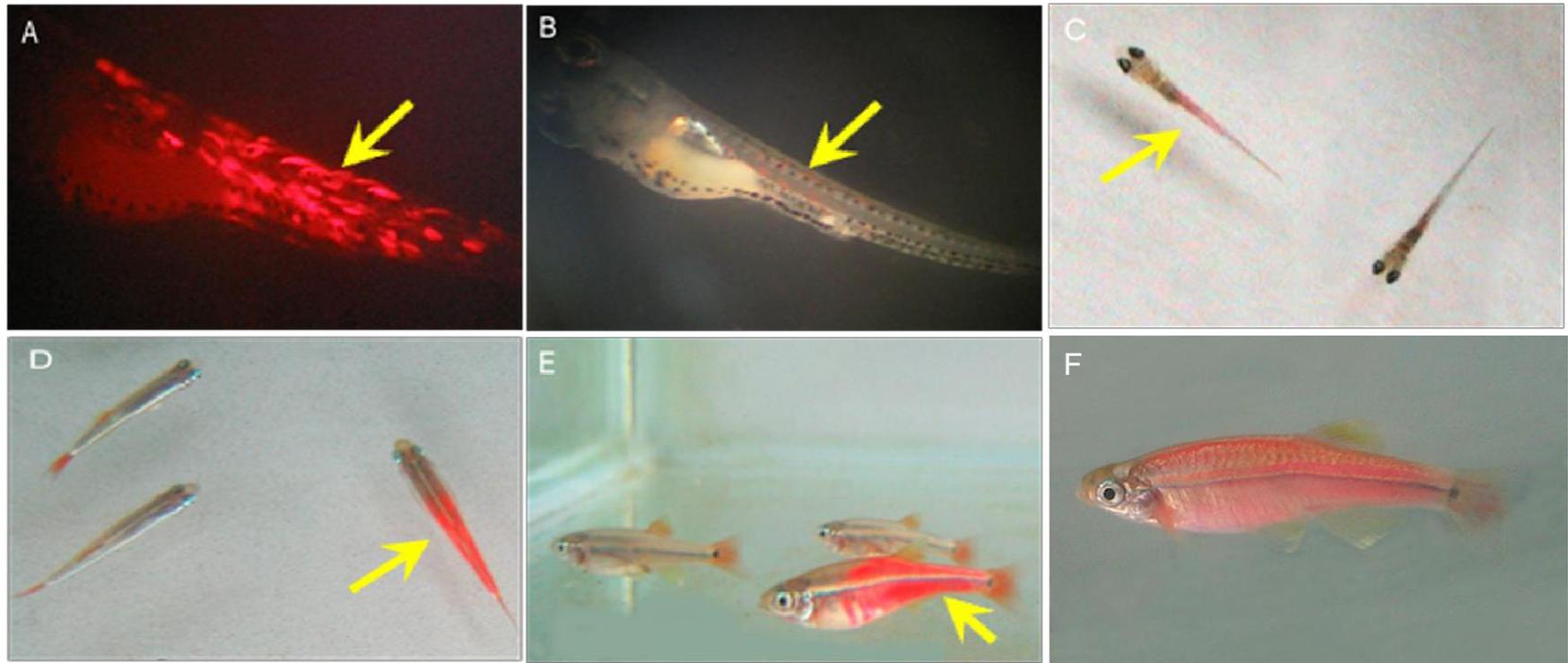
The randomly chosen five F<sub>3</sub> individuals from the germline transgenic *T. albonubes* were used to analyze the genetic character of RFP gene. When three F<sub>3</sub> female and two F<sub>3</sub> male were crossed with wild-type fish, 100, 45.2, 51.5, 51.1 and 52% of their F<sub>4</sub> transgenic progeny showed uniform RFP expression (Figure 2F), respectively. Segregation studies indicated that the transgene follows Mendelian single-gene inheritance in this line (Table 2). Furthermore, dot blot analysis showed that no hybridization signal was observed in all the RFP-negative offspring (data not shown). These results reveal that the pMYLZ2-RFP transgene DNA is integrated into the *T. albonubes* genome and exists at a single locus.

### Expression levels of RFP in different tissues

To examine the characteristic of transgene expression in different tissues of adult F<sub>3</sub> transgenic individuals, the presence of RFP mRNA was determined by RT-PCR. An expected 288 bp band of RFP was specifically amplified from transgenic fish, while no band was observed in the non-transgenic siblings. As shown in Figure 3, relatively higher RFP expression levels were detected in muscle, liver and gonad, while a lower expression was detected in spleen, gill, intestine, swim bladder and heart. Furthermore, similar expression patterns were observed between male and female transgenic fish.

### The arrangement of pMYLZ2-RFP gene constructs in transgenic fish

The quantitative PCR analysis revealed that multiple copies of the pMYLZ2-RFP construct were inserted into the host genome. To further characterize the arrangement



**Figure 2.** Expression of RFP in *T. albonubes* injected with the linearized pMYLZ2-RFP vector. Red fluorescence in transgenic larva was observed under a fluorescence microscope (A) and an anatomical lens (B). RFP expression in 7 dph (C) and adult (D, E) transgenic founder fish was observed under normal daylight. The arrows indicate muscle-specific expression of RFP in transgenic fish. Wild-type fish as controls are also shown in (C), (D) and (E). (F) Adult transgenic F<sub>4</sub> progeny with uniform RFP expression photographed under normal daylight.

of the expression vector, single primer szq1 binds in the promoter region and extends in the upstream direction whereas single primer xzst2 binds in the terminator region and extends in the downstream direction. If the constructs were arranged in a head-to-head fashion, single primer szq1 would be anticipated to amplify an 892 bp band. Likewise, single primer xzst2 would be anticipated to amplify a 748 bp band if the constructed were arranged in a tail-to-tail fashion. However, as shown in Figure 4, no product was

amplified by either single primer while an approximate 800 bp band was amplified using their combination. These results indicated that the constructs were not organized in either a head-to-head or a tail-to-tail arrangement fashion but in a head-to-tail orientation (Figure 5).

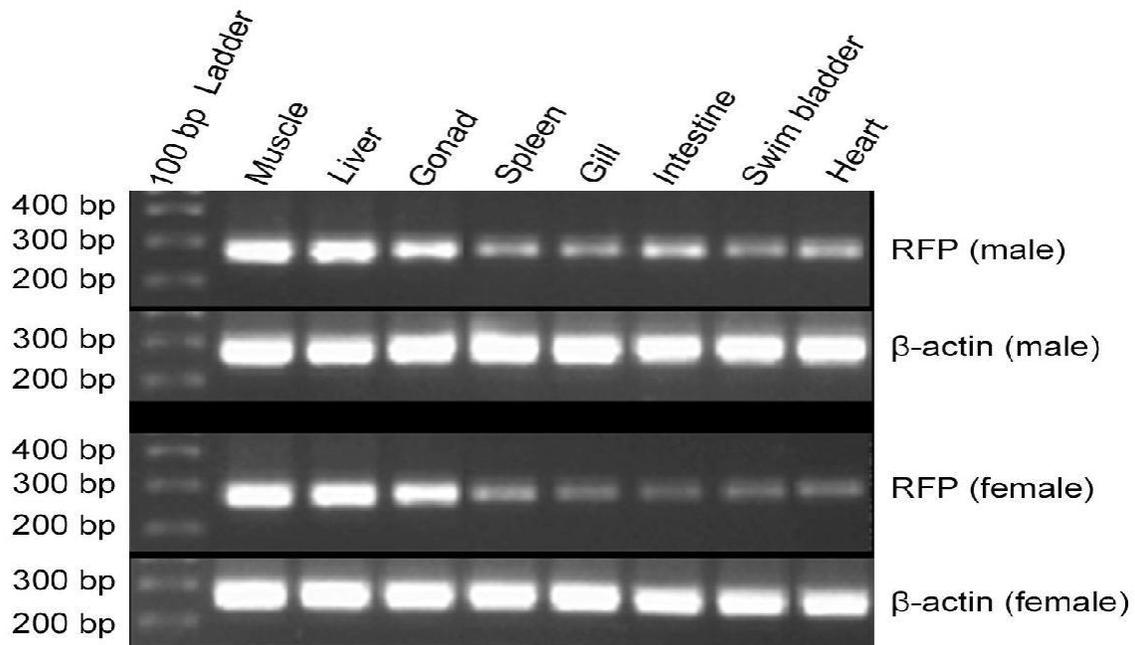
#### Flanking DNA sequence analysis

Approximately 1.5 kb fragment was generated

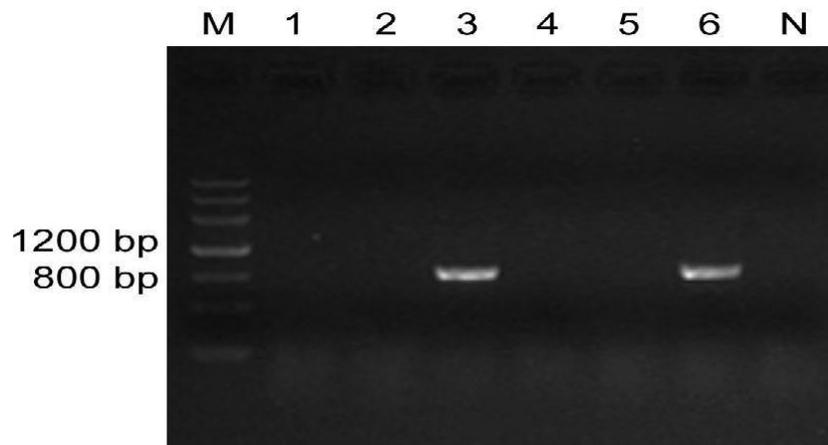
from *EcoR V* library, which contained 1075 bp of 5'-flanking region of the transgene insertion locus and 442 bp of the transgene sequence. Compared to the corresponding 5'-end sequence of pMYLZ2-RFP vector, the 442-bp transgenic sequence was no major modification but the deletion of cohesive terminus of *Bgl*II at the upstream transgene-host genomic DNA junction. The 5'-flanking sequence analysis showed that there was no significant homology to sequences present in the GenBank database. In addition,

**Table 2.** Segregation of pMYLZ2-RFP transgene from transgenic F<sub>3</sub> generation to F<sub>4</sub> progenies.

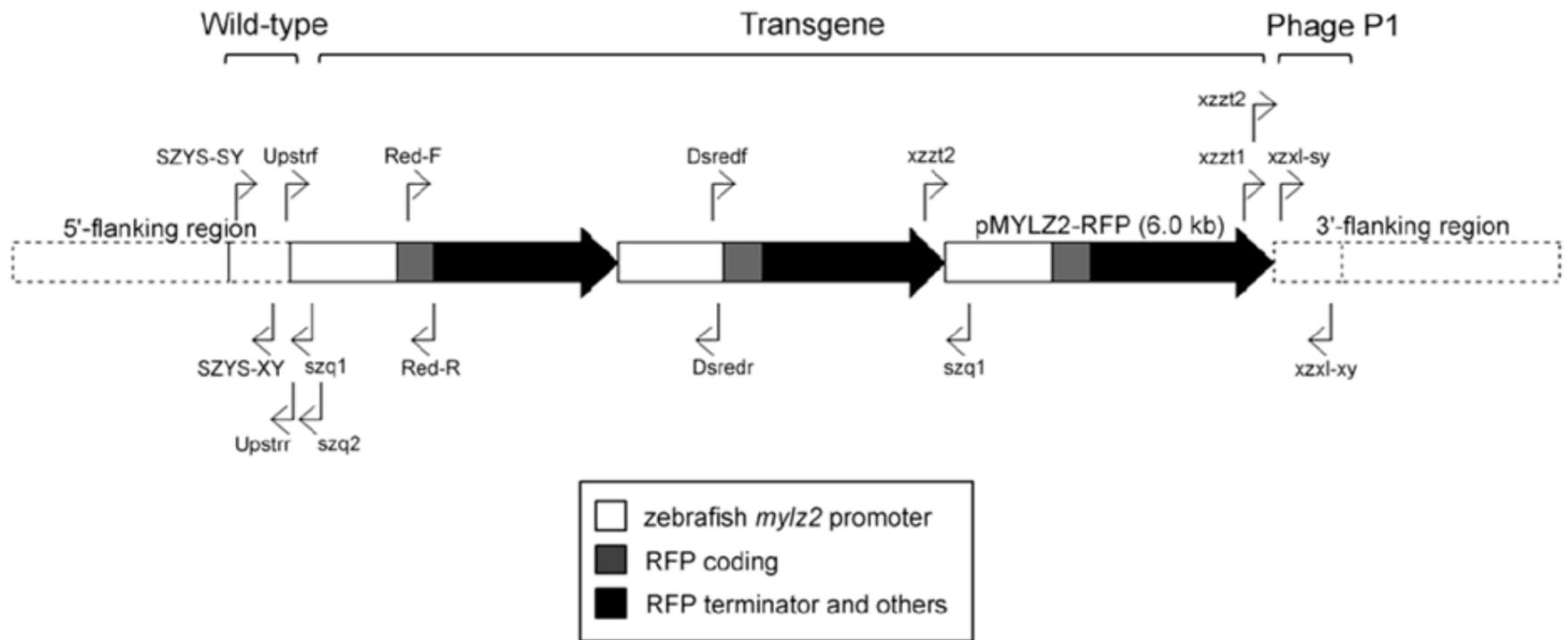
Trial	Sex	Crosses with wild-type fish			
		Positive	Negative	Total	Transgene transmission / %
1	Female	96	0	96	100.0
2	Female	28	34	62	45.2
3	Female	34	32	66	51.5
4	Male	24	23	47	51.1
5	Male	39	36	75	52.0



**Figure 3.** RT-PCR analysis of RFP expression in various tissues of male and female F<sub>3</sub> transgenic *T. albonubes*. The endogenous  $\beta$ -actin mRNA was used as an internal control.



**Figure 4.** The arrangement of pMYLZ2-RFP constructs in transgenic fish. Lane M: Marker III; lanes 1, 4: single-primer PCR with primer szq1 in transgenic individuals; lanes 2, 5: single-primer PCR with primer xzst2 in transgenic individuals; lanes 3, 6: PCR with both szq1 and xzst2 primers in transgenic individuals; lane N: PCR with both szq1 and xzst2 primers in non-transgenic fish.



**Figure 5.** Genetic structure of transgenic *T. albonubes*. The boxes made up of dotted lines represent flanking regions of transgene insertion locus. Three complete copies of the 6.0 kb pMYLZ2-RFP construct were integrated into *T. albonubes* genome, and arranged in a head-to-tail fashion. The arrows represent the amplified direction and location of primers.

primers SZYS-SY and SZYS-XY (Table 1) were designed according to the obtained 5'-flanking sequence, and an expected size fragment was specifically amplified from wild-type genome (data not shown). This further demonstrated that the amplified 1075-bp fragment existed in the host genome DNA.

Approximately 1.6 kb fragment was generated from *Dra* I library, which contained 374 bp of the transgene sequence and 1233 bp of 3'-flanking region of the transgene insertion locus. Compared to the corresponding 3'-end sequence of pMYLZ2-

RFP vector, the 374-bp transgenic sequence was not found the mutations including base substitution and insertion, and *Bgl*II restriction site was well preserved at the downstream junction. Sequence analysis showed that the 3'-flanking sequence was 99% identical to that of phage P1 partial genome. Furthermore, the same sequence was also amplified using *Pvu*II genome walker library. It indicated that the amplified 1.6-kb fragment belonged to the 3'-flanking region of the insertion point. Nevertheless, PCR analysis revealed that no product was amplified from wild-

type *T. albonubes* genome DNA with primers xzxl-sy and xzxl-xy (Table 1), which were designed according to the obtained 3'-flanking sequence.

#### Transgene copy numbers estimated by qPCR

In the qPCR assay, standard curves for target gene and internal reference sequence were produced by using ABI 7300 System SDS software. The correlation coefficient ( $R^2$ ) between  $C_T$  value and log DNA concentration was 0.9992 for target gene and 0.9982 for internal reference sequence.

**Table 3.** Estimation of transgene copy number in F<sub>5</sub> and F<sub>6</sub> transgenic *T. albonubes*.

Trial	F <sub>5</sub> generation			F <sub>6</sub> generation		
	Initial copy of internal reference	Initial copy of target gene	Transgene copy number	Initial copy of internal reference	Initial copy of target gene	Transgene copy number
1	3.40E+5	8.96E+5	2.64	9.14E+4	2.71E+5	2.96
2	3.34E+5	9.74E+5	2.92	8.80E+4	2.48E+5	2.82
3	2.99E+5	9.26E+5	3.10	1.00	2.97E+5	2.97
4	3.44E+5	9.75E+5	2.83	9.32E+4	2.73E+5	2.93
Mean ± SD	(3.29±0.21)E+5	(9.43±0.39)E+5	2.87±0.19	(9.31±0.51)E+4	(2.72±0.20)E+5	2.92±0.07

The PCR reaction efficiencies were 95.29 and 100.39% for target gene and internal reference sequence PCR assay, respectively. Good linearity between DNA quantities and C<sub>T</sub> values indicated that the established assays were well suited for quantitative measurements.

Copy number value for RFP gene is shown in Table 3. The initial copies for RFP gene and internal reference sequence were  $(9.43 \pm 0.39) \times 10^5$  and  $(3.29 \pm 0.21) \times 10^5$  in the F<sub>5</sub> sample, and  $(2.72 \pm 0.20) \times 10^5$  and  $(9.31 \pm 0.51) \times 10^4$  in the F<sub>6</sub> sample, respectively. From these data, transgene copy number value was determined to be three per haploid genome.

## DISCUSSION

In this study, 45% of *T. albonubes* eggs survived after microinjection and 31% of the hatched fry tested RFP-positive which represents a relative high efficiency of transgenic fish production. Nonetheless, most transgenic founders displayed the patched RFP expression in body muscles. This mosaic phenomenon has been widely reported (Chou et al., 2001; Pan et al., 2008) and thought to result from delayed transgenic integration after the pre-cleavage stage of embryonic development (Gross et al., 1992; Figueiredo et al., 2007). In fact, the stable transgenic *T. albonubes* line described in this paper, although mosaic, was able to transmit RFP gene to the offspring because the transgene was integrated into its germ cells. In addition, Mendelian transmission pattern suggested the transgene exists at a single locus in this line.

Results of RT-PCR analysis revealed that RFP gene driven by zebrafish *mylz2* promoter was expressed in most *T. albonubes* tissues, but previous studies reported that this muscle-specific promoter can drive reporter gene expression solely in the skeletal muscle (Gong et al., 2003; Ju et al., 2003; Zeng et al., 2005). Such ectopic expression may be due to lack of putative tissue-specific or repressor elements in our isolated zebrafish *mylz2* promoter (Jian et al., 2004; Hobbs and Fletcher, 2008). On the other hand, the influence of integration site context also may result in ectopic expression (Clark et al., 1994). For example, expression patterns often differ

among transgenic lines carrying the same gene construct (Nam et al., 1999). Yet, the possible mechanism for ectopic transgene expression needs further study.

Traditionally, Southern blot hybridization is the conventional method for transgene copy number determination (Rahman et al., 2000). But this method is time consuming, costly, and requires a large amount of DNA sample for each assay (Ballester et al., 2004; Yuan et al., 2007). To avoid these drawbacks, we adopted quantitative PCR to determine RFP gene copy number in transgenic *T. albonubes*, characterized by speediness, sensitivity and accurateness. For qualitative PCR assays, an endogenous reference gene of known copy number is always considered as necessity, but it is relatively difficult to screen a suitable gene in *T. albonubes*. Specifically, this study developed a 143-bp fragment from the 5' transgene-host junction region used as an internal control, which was validated to be present as a single copy per haploid transgenic *T. albonubes* genome. This novel design has also been applied to the detection of genetically modified foods (Hernández et al., 2003; Zhu et al., 2008). Meanwhile, the same result obtained for the different generations demonstrated the feasibility of qPCR method to estimate transgene copy number.

Both flanking sequences were obtained from the transgene insertion locus using genome walking technology. BLAST analysis showed that the 5'-flanking sequence bore no significant homology to sequences present in the GenBank database, while the 3'-flanking sequence was 99% identical to that of phage P1 partial genome. A similar event was reported by Uh et al. (2006) that the downstream flanking sequences of an insertion site in transgenic coho salmon (*Oncorhynchus kisutch*) has high homology to the cercaria antigen (CA) membrane protein gene from *Schistosoma japonicum*. One explanation is that the P1 phage sequences had existed in the genome DNA of the wild-type *T. albonubes* receptor prior to injecting the pMYLZ2-RFP gene constructs. This phenomenon indicates the possible presence of "hot spot" for DNA integration. Indeed, a few studies reveal some interesting trends that integration of foreign DNA preferentially occurs near the repetitive regions, AT-rich regions, ends of chromosomes, etc (Hamada et al., 1993; Rijkers et al., 1994). Nevertheless,

most transgene integration events are often viewed as a random process (Merrihew et al., 1996; Smith, 2001). Certainly, it is impossible to ignore an alternative possibility that the 3'-flanking sequence may be from the original sample for injection, which was inserted together with the pMYLZ2-RFP construct. The presumptive evidences are listed as follows:

(1) No product was amplified from some other wild-type *T. albonubes* individuals with primers xzxl-sy and xzxl-xy, which were designed according to the obtained 3'-flanking sequence.

(2) P1 is a temperate phage which infects and lysogenizes *Escherichia coli* and other enteric bacteria (Łobocka et al., 2004). This provides the probability that phage P1 genomic DNA was blended into the purified recombinant vector because of the employment of *E. coli* DH5α competence cells in the experiment.

(3) When the exogenous DNA sequences are integrated into the host genome, the deletion or modification of nucleotides always take place at the transgene-host DNA junction (Gheysen et al., 1991; Mayerhofer et al., 1991).

However, sequence analysis showed that the downstream transgene region adjacent to the P1 phage homologous sequences was not found any kind of alteration at nucleotide level, and *Bgl*II restriction site was well preserved at the downstream junction. If this assumption is true, it not only puts a higher safety requirement for transgenic manipulation, but also shows the unpredictable nature of transgenesis.

In summary, this germline transgenic *T. albonubes* is of some unique features such as ectopic expression and an unusual 3'-flanking sequence, however, Mendelian transmission, consistent expression, complete pMYLZ2-RFP constructs and an appropriate genomic integration context well demonstrated the transgenes have been inserted as a stable DNA fragment in the *T. albonubes* genome and can be stably inherited and expressed in transgenic progenies.

## ACKNOWLEDGMENTS

Research results reported in this paper were supported by grants from the National Facilities and Information Infrastructure for Science and Technology Program of China (Grant No. 2005DKA21103 and 2010B20308002), Science and Technology Plan Projects of Guangdong Province, China (Grant No. 2005B20301018), Fish Germplasm Improvement and Evaluation Research (Grant No. A200899F01) and National High-Tech Research and Development Program 863 of China (Grant No.2009AA10Z105).

## REFERENCES

Ballester M, Castelló A, Ibáñez E, Sánchez A, Folch JM (2004). Real-time quantitative PCR-based system for determining transgene copy

- number in transgenic animals. *Biotechniques*, 37: 610-613.
- Chou CY, Horng LS, Tsai HJ (2001). Uniform GFP-expression in transgenic medaka (*Oryzias latipes*) at the F0 generation. *Transgenic Res.* 10: 303-315.
- Clark AJ, Bissinger P, Bullock DW, Damak S, Wallace R, Whitelaw CB, Yull F (1994). Chromosomal position effects and the modulation of transgene expression. *Reprod. Fertil. Dev.* 6: 589-598.
- Figueiredo MA, Lanes CFC, Almeida DV, Marins LF (2007). Improving the production of transgenic fish germlines: in vivo evaluation of mosaicism in zebrafish (*Danio rerio*) using a green fluorescent protein (GFP) and growth hormone cDNA transgene co-injection strategy. *Genet. Mol. Biol.* 30: 31-36.
- Gheysen G, Villarroel R, Van Montagu M (1991). Illegitimate recombination in plants: a model for T-DNA integration. *Genes Dev.* 5: 287-297.
- Gong ZY, Wan HY, Tay TL, Wang H, Chen M, Yan T (2003). Development of transgenic fish for ornamental and bioreactor by strong expression of fluorescent proteins in the skeletal muscle. *Biochem. Biophys. Res. Commun.* 308: 58-63.
- Gross ML, Schneider JF, Moav N, Moav B, Alvarez C, Myster SH, Liu Z, Hallerman EM, Hackett PB, Guise KS, Faras AJ, Kapuscinski AR (1992). Molecular analysis and growth evaluation of northern pike (*Esox lucius*) microinjected with growth hormone genes. *Aquaculture*, 103: 253-273.
- Hamada T, Sasak H, Seki R, Sakak Y (1993). Mechanism of chromosomal integration of transgenes in microinjected mouse eggs: sequence analysis of genome-transgene and transgene-transgene junctions at two loci. *Gene*, 128: 197-202.
- Hernández M, Pla M, Esteve T, Prat S, Puigdomènech P, Ferrando A (2003). A specific real-time quantitative PCR detection system for event MON810 in maize Yield Gard based on the 3'-transgene integration sequence. *Transgenic Res.* 12: 179-189.
- Hobbs RS, Fletcher GL (2008). Tissue specific expression of antifreeze protein and growth hormone transgenes driven by the ocean pout (*Macrozoarces americanus*) antifreeze protein OP5a gene promoter in Atlantic salmon (*Salmo salar*). *Transgenic Res.* 17: 33-45.
- Iyengar A, Müller F, Maclean N (1996). Regulation and expression of transgenes in fish—a review. *Transgenic Res.* 5: 147-166.
- Jian Q, Bai JJ, Ye X, Xia SL, Liang XF, Luo JR (2004). Cloning of Mylz2 promoter and generation of green fluorescence transgenic zebrafish. *J. Fish. Sci. China*, 11: 391-395.
- Ju B, Chong SW, He J, Wang X, Xu Y, Wan H, Tong Y, Yan T, Korzh V, Gong Z (2003). Recapitulation of fast skeletal muscle development in zebrafish by transgenic expression of GFP under the mylz2 promoter. *Dev. Dyn.* 227: 14-26.
- Kinoshita M, Toyohara H, Sakaguchi M, Inoue K, Yamashita S, Satake M, Wakamatsu Y, Ozato K (1996). A stable line of transgenic medaka (*Oryzias latipes*) carrying the CAT gene. *Aquaculture*, 143: 267-276.
- Kinoshita M (2004). Transgenic medaka with brilliant fluorescence in skeletal muscle under normal light. *Fish. Sci.* 70: 645-649.
- Liang XF, Chen GZ, Chen XL, Yue PQ (2008). Threatened fishes of the world: *Tanichthys albonubes* Lin 1932 (Cyprinidae). *Environ. Biol. Fish.* 82: 177-178.
- Lin S, Yang S, Hopkins N (1994). *LacZ* expression in germline transgenic zebrafish can be detected in living embryos. *Dev. Biol.* 161: 77-83.
- Liu H-S, Jan M-S, Chou C-K, Chen P-H, Ke H-J (1999). Is green fluorescent protein toxic to the living cells? *Biochem. Biophys. Res. Commun.* 260: 712-717.
- Łobocka MB, Rose DJ, Plunkett G, Rusin M, Samoedny A, Lehnher H, Yarmolinsky MB, Blattner FR (2004). Genome of bacteriophage P1. *J. Bacteriol.* 186: 7032-7068.
- Matz MV, Fradkov AF, Labas YA, Savitsky AP, Zaraisky AG, Markelov ML, Lukyanov SA (1999). Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat. Biotechnol.* 17: 969-973.
- Mayerhofer R, Koncz-Kalman Z, Nawrath C, Bakkeren G, Cramer A, Angelis K, Redei GP, Schell J, Hohn B, Koncz C (1991). T-DNA integration: a mode of illegitimate recombination in plants. *EMBO J.* 10: 697-704.
- Merrihew RV, Marburger K, Pennington SL, Roth DB, Wilson JH (1996). High-frequency illegitimate integration of transfected DNA at preintegrated target sites in a mammalian genome. *Mol. Cell. Biol.*

- 16: 10-18.
- Nam YK, Noh CH, Kim DS (1999). Transmission and expression of an integrated reporter construct in three generations of transgenic mud loach (*Misgurnus mizolepis*). *Aquaculture*, 172: 229-245.
- Pan X, Zhan H, Gong Z (2008). Ornamental expression of red fluorescent protein in transgenic founders of white skirt tetra (*Gymnocorymbus ternetzi*). *Mar. Biotechnol.* 10: 497-501.
- Rahman MA, Hwang G-L, Razak SA, Sohm F, Maclean N (2000). Copy number related transgene expression and mosaic somatic expression in hemizygous and homozygous transgenic tilapia (*Oreochromis niloticus*). *Transgenic Res.* 9: 417-427.
- Rijkers T, Peetz A, R  ther U (1994). Insertional mutagenesis in transgenic mice. *Transgenic Res.* 3: 203-215.
- Shaner NC, Campbell RE, Steinbach PA, Giepmans BN, Palmer AE, Tsien RY (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* 22: 1567-1572.
- Sin FYT (1997). Transgenic Fish. *Rev. Fish Biol. Fish.* 7: 417-441.
- Smith K (2001). Theoretical mechanisms in targeted and random integration of transgene DNA. *Reprod. Nutr. Dev.* 41: 465-485.
- Stewart CN Jr (2006). Go with the glow: fluorescent proteins to light transgenic organisms. *Trends Biotechnol.* 24: 155-162.
- Tzfira T, LiJ, Lacroix B, Citovsky V (2004). *Agrobacterium* T-DNA integration: molecules and models. *Trends Genet.* 20: 375-383.
- Uh M, Khattra J, Devlin RH (2006). Transgene constructs in coho salmon (*Oncorhynchus kisutch*) are repeated in a head-to-tail fashion and can be integrated adjacent to horizontally-transmitted parasite DNA. *Transgenic Res.* 15: 711-727.
- Wilson AK, LathamJR, Steinbrecher RA (2006). Transformation-induced mutations in transgenic plants: Analysis and biosafety implications. *Biotechnol. Genet. Eng. Rev.* 23: 209-237.
- W  rtele H, Little KCE, Chartrand P (2003). Illegitimate DNA integration in mammalian cells. *Gene Theor.* 10: 1791-1799.
- Yi ZS, Chen XL, Wu JX, Yu SC, Huang CE (2004). Rediscovering the wild population of white cloud mountain minnows (*Tanichthys albonubes* Lin) on Guangdong province. *Zool. Res.* 25: 551-555.
- Yuan JS, Burris J, Stewart NR, Mentewab A, Stewart CN (2007). Statistical tools for transgene copy number estimation based on real-time PCR. *BMC Bioinformatics*, 8: S6.
- Zbikowska HM (2003). Fish can be first- advances in fish transgenesis for commercial applications. *Transgenic Res.* 12: 379-389.
- Zeng Z, Liu X, Seebah S, Gong Z (2005). Faithful expression of living color reporter genes in transgenic medaka under two tissue-specific zebrafish promoters. *Dev. Dyn.* 234: 387-392.
- Zhu H, Zhao X, Jia J, Sun J, Zhao K (2008). A specific qualitative and real-time PCR detection of MON863 maize based on the 5'-transgene integration sequence. *J. Cereal Sci.* 48: 592-597.