Transgenes in F_4 pMThGH-transgenic common carp (*Cyprinus carpio* L.) are highly polymorphic

ZENG Zhiqiang & ZHU Zuoyan

State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China
Correspondence should be addressed to Zhu Zuoyan (e-mail: zyzhu@ihb.ac.cn)

Abstract To gain information on the integration pattern of pMThGH-transgene, 50 transgenes were recovered from F_4 generation of pMThGH transgenic common carp (*Cyprinus carpio* L.) and 33 recovered genes were analyzed. The restriction maps of these recovered genes were constructed by digestion with five kinds of enzymes. These transgenes can be classified into 4 types according to their restriction maps. Only one type of transgenes maintains its original molecular form, whereas the other three types are very different from the original one and vary each other on both molecular weight and restriction maps. This implies that the sequences of most transgenes have been deleted and/or rearranged during integration and inheritance. The results of PCR amplification and Southern blot hybridization indicate that *MThGH* in Type Ⅰ transgene keeps intact but most of its sequence has been deleted in other three types. All these results suggest that transgenes in F_4 generation of transgenic carp are highly polymorphic. Two DNA fragments concerning integration site of transgenes were cloned from recovered transgenes, and found to be homologous to the 5′ UTR of *â-actin* gene of common carp and mouse mRNA for receptor tyrosine kinase (*RTK*), respectively.

human growth hormone gene sequences, and then inserted into pBR322 at the *Eco*R Ⅰ site (fig. 1). DNA of pMThGH was linearized by digestion with *Bam*H Ⅰ and microinjected into the fertilized eggs of common carp to produce the founder transgenics[7]. After being confirmed by dot blotting or PCR, transgenic fishes were naturally mated and gave birth to their offspring and then led to the F_4 generation. At the age of two months old, the average body weight for 17 F_4 transgenic individuals was 7.59 g. The largest one (26.35 g) was selected for analysis.

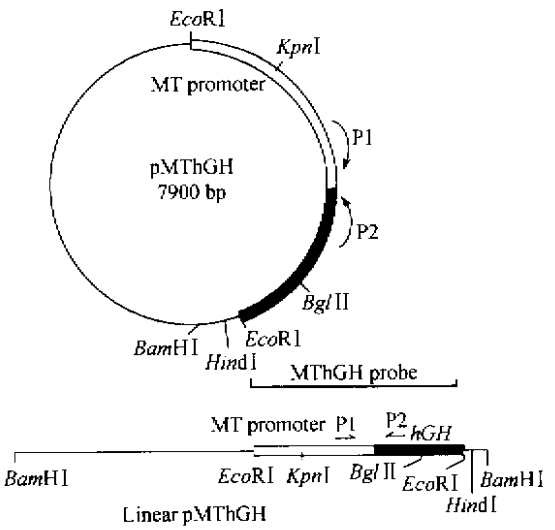

Fig. 1. Structure of transgene pMThGH and linearized DNA for microinjection. Location of PCR primer P1, P2 and *MThGH* probe used for this study are also shown in the diagram.

(ⅱ) Restriction digestion with *Bam*H Ⅰ and determination of the enriched region of pMThGH-transgene.

Genomic DNA was extracted from the liver of transgenic fish. 10 μg of DNA was completely digested with restriction endonuclease *Bam*H Ⅰ, and separated by 0.8% agarose gel electrophoresis. Southern blotting against α-^{32}P-dCTP labeled pMThGH probe was carried out as described by Maniatis et al.[9]. The molecular weight of transgene in single copy was figured out by the location of hybridization band.

(ⅲ) Recovering pMThGH-transgene

(1) Partial digestion. One of our aims in this note is to clone the host sequences flanking transgene or the sequences of integration sites. By partial digestion with *Bam*H Ⅰ, host sequence joining the linear transgene at its end with *Bam*H Ⅰ site may be cloned. 40 μg of genomic DNA were digested with *Bam*H Ⅰ for the varied period of time. The procedure can be briefly described as follows: 1/5 volume of total reaction solution were taken out at intervals of 2nd, 5th, 10th, 15th, 30th min, then the reaction was stopped by adding EDTA to a final concentration of 0.05 mol/L.

(2) Recovery of DNA fragments. The result of Southern blot hybridization shows that the molecular weight of single pMThGH-transgene released from concatemer is 7.5—8.0 kb (data not shown). Considering part of transgenes may have changed in sequence length after several generations, DNA fragments ranging from 5 to 10 kb were recovered and purified from low-melting agarose gel. Recovered DNA was diluted in 100 μL TE buffer (10 mmol/L Tris·HCl, 1 mmol/L EDTA, pH 8.0) at a concentration of 40 μg/mL.

(3) Circularization of recovered DNA fragments. A modified method[10] for DNA circularization was employed here. 200 μL DNA fragments at a final concentration of 2 μg/mL were self-ligated with 1U T_4 ligase at 16 ℃ for 8 h. The circularized DNA was extracted with phenol-chloroform, precipitated by ethanol and re-suspended in 10 μL TE buffer (10 mmol/L Tris·HCl, 1 mmol/L EDTA, pH 8.0).

(4) Transformation into *E. coli*. 10 μL of circularized DNA was used to transform 100 μL of DH5α competent cell using $CaCl_2$ treatment procedure. Transformed cells were spread on LB plates with 50 μg/mL ampicillin. pMThGH DNA was transformed as the positive control.

(ⅳ) Classification of recovered transgenes. 50 clones were obtained and 33 clones were randomly picked up for further analysis. To classify these clones, plasmid DNA were mini-prepared, and doubly digested with *Eco*R Ⅰ and *Bam*H Ⅰ. Classification was performed according to the electrophoresis patterns of restriction fragments resulting from double digestion.

(ⅴ) Mapping. Recovered plasmid DNA were firstly digested with *Bam*H Ⅰ and *Hin*d Ⅲ, respectively. 1/3 of each linear DNA continued to be digested with *Eco*R Ⅰ, *Bgl* Ⅱ or *Kpn* Ⅰ. Digestion DNA fragments were separated by 1.0% agarose gel electrophoresis, stained with ethidium bromide and visualized under UVP GDS8000 system (UVP Ltd., UK). Restriction maps for five kinds of restriction endonucleases were constructed on the basis of digestion fragment sizes.

(ⅵ) Detection for *MThGH* fragment in recovered transgenes by PCR amplification and Southern blot hybridization. The detection of *MThGH* fragment by PCR was performed using the method described by Zhao et al.[8]. Sense primer P_1 and anti-sense primer P_2 were located at the *MT-1* promoter and *hGH* sequence, respectively (fig. 1). The expected size of PCR product was 450 bp. For Southern blotting, various recovered transgenes digested with *Bam*H Ⅰ were separated by 0.8% agarose gel electrophoresis, and transferred onto nylon membrane. DNA hybridization against DIG-labeled *MThGH* and de-

tection were carried out according to the user's manual (DIG DNA Labeling and Detection Kit, Boehringer Mannheim).

(vii) DNA sequencing of the regions flanking the recovered transgenes. To confirm that recovered transgenes had trapped the host DNA sequences at their ends, two DNA fragments, B/H (1 kb) on the left side of Type II transgene and B/B (0.5 kb) on the right side of Type III transgene, were subcloned into the appropriate sites on pUC18. Recombinants were named B5(B-H) and B6(B-B′), respectively. DNA sequences of these two subfragments were determined using the dideoxy sequencing method. The data collection was automatically performed on the ABI 310 Genetic Analyzer (PE Applied Biosystems). Searching for homologous sequences to the determined sequences in nucleotide sequence database (GenBank+ EMBL+ DDBJ) were carried out using database search program "BLAST 2.0".

2 Results

(i) Recovery of transgene. 50 ampicillin-resistant colonies were recovered from genomic DNA of a F_4 transgenics. 33 colonies were randomly selected for analysis. The 33 transgenes can be classified into four types according to the results of double digestion (fig. 2). Detailed classifying results are shown as follows. Type I: Nos. 2, 15, 16, 19, 21, 26 (6 in total, 18.0%), their electrophoresis patterns were the same as that of pMThGH; Type II: Nos. 1, 3–5, 8, 10, 12–14, 17, 18, 20, 23–25, 27, 28, 30–33, (21 in total, 63.6%); Type III: Nos. 6, 7; Type IV: Nos. 9, 11, 22, 29.

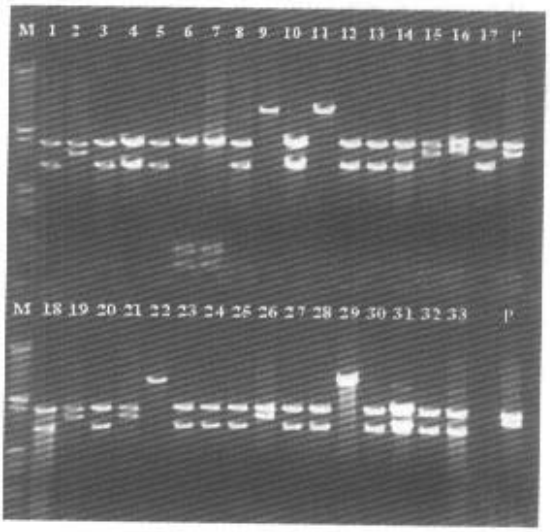

Fig. 2. The electrophoresis pattern of plasmid DNA of recovered colonies after double digestion with *Eco*R I and *Bam*H I. P, pMThGH; 1—33, recovered colonies.

(ii) The restriction maps of recovered transgenes. The restriction maps for four types of recovered transgenes were constructed by digestion with *Bam*H I, *Hind* III, *Eco*R I, *Bgl* II and *Kpn* I (fig. 3). The map for Type I is the same as that for pMThGH, maps for the other three are very different from that for Type I.

(iii) Results of PCR amplification and Southern blot analysis for *MThGH* fragment. PCR products with 450 bp in length did not appear in all types of transgene except for Type I as revealed on electrophoresis gel (fig. 4), while all recovered transgenes could hybridize against *MThGH* probe (fig. 5). These results indicated that the sequences of most transgenes were quite different from their original form. The fact is that the analyzed 450 bp

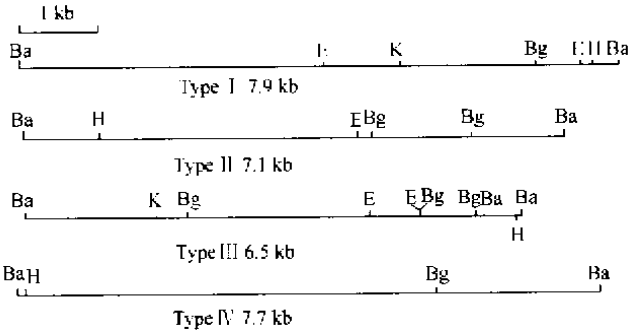

Fig. 3. Restriction maps of recovered transgenes. Type I, Original form; Types II—IV, deficient form. Ba, *Bam*H I; E, *Eco*R I; K, *Kpn* I; Bg, *Bgl* II; H, *Hind* III.

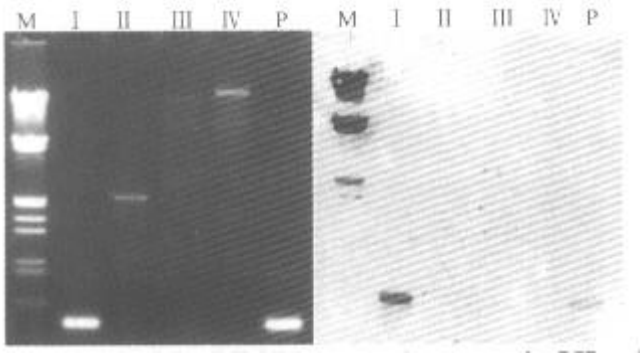

Fig. 4. Detection for *MThGH* in recovered transgenes by PCR and confirmation by Southern blotting. DIG-labeled *MThGH* and λDNA were used as probes. M, λDNA (*Hind* III/*Eco*R I); P, plasmid of pMThGH; Type I, original form; Types II—IV, deficient forms.

fragment changed significantly while those regions responsible for plasmid replication and *E. coli* ampicillin resistance still kept intact.

(iv) Sequencing results of DNA fragments at the flanks of two transgenes. 470 bp of DNA sequence from the *Bam*H I to *Hind* III sites on Type II transgene was determined. The homology between its 438 bp and the

promoter or 5′ untranslation region (5′ UTR) of common carp *b-actin* gene on the piscine *GFP* expression vector FRMwg (accession number in GeneBank: AF1709151) was pronounced (98% identity, fig. 6). The full length of B/B fragment in Type Ⅲ transgene was sequenced and 509 bp was determined. The blast research result showed that the homology between DNA sequence of B/B fragment and mouse mRNA for receptor tyrosine kinase (accession number in DBJ: D13738.1) was significant. 288 bp of the former sequence (base positions 33 —320) matched the latter with base identity of 99% (fig. 7).

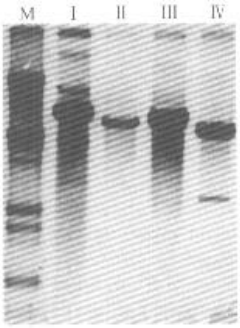

Fig. 5. Southern hybridization analysis of recovered transgene. DIG-labeled *MThGH* and λDNA were used as probes. M, λDNA (*Hind* Ⅲ/*Eco*R Ⅰ); P, plasmid of pMThGH; Type Ⅰ, Original form; Types Ⅱ — Ⅳ, deficient forms.

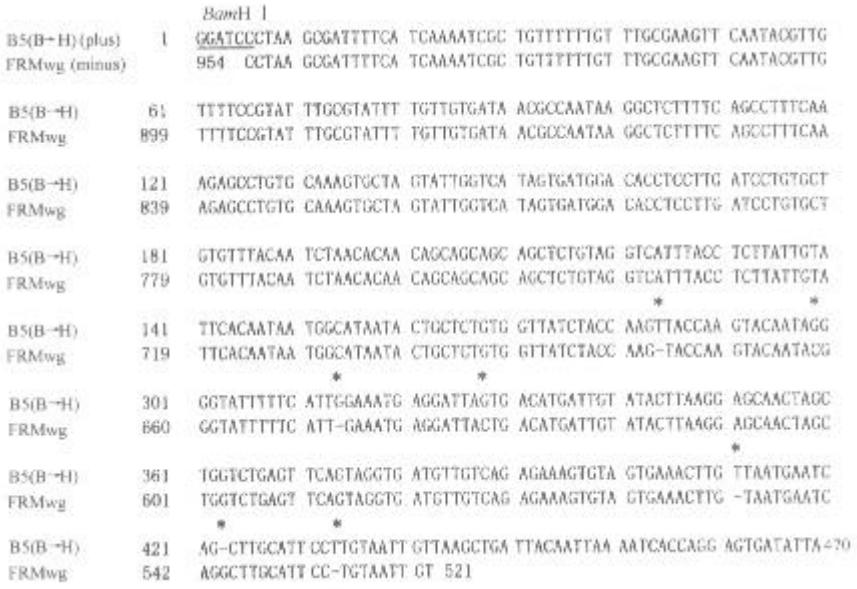

Fig. 6. Nucleotide sequence comparison between B/H fragment in the flank of Type Ⅱ transgene and the promoter or 5′ UTR of common carp *b-actin* gene on the piscine *GFP* expression vector FRMwg. Restriction site at the end of fragment is underlined. Base difference and gaps are indicated by asterisks.

3 Discussion

(ⅰ) The polymorphism of transgene in transgenic fish. The restriction maps for these four types of transgenes showed that only a few transgenes maintained their original construction for the founder transgenics, most transgenes are totally different from their original form in product sizes and endonucleases recognition sites in sort, number and location. This may imply that part of the sequences in most transgenes have been deleted and/or rearranged during the course of integration and inheritance, and appear to be very polymorphic.

Sequencing results confirm that some sequences in recovered transgenes Type Ⅱ and Type Ⅲ not only were deleted, but also were interrupted by host sequences. Similar evidence was also found in the early reported case. In the case of early postimplantation embryo lethality due to DNA rearrangements in a transgenic mouse strain (*HUGH*/3), approximately five copies of transgenes were arrayed in tandem but interrupted at least twice by mouse cellular sequences [2]. These results lead us to propose that multiple copies of transgenes may not be simply arranged in tandem and directly inserted into the chromosome at integration site, the course of integration may involve homologous and/or illegitimate recombination between transgene and transgene, and between transgene and genome. Some transgenes sequences will be deleted, arranged and/or interrupted by host sequences. Whether matters involving the polymorphism of transgenes take place in founder transgenics, or during the inheritance to the following generations, or in both of case, needs further studies.

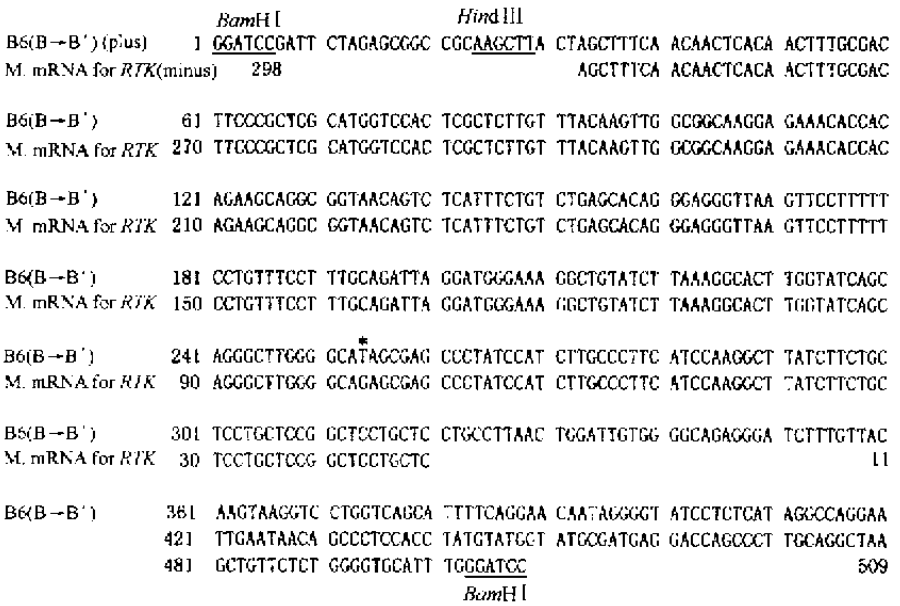

Fig. 7. Nucleotide sequence of B/B fragment on the flank of Type III transgene and alignment of its 288 bp with the sequence of mouse mRNA for receptor tyrosine kinase. Restriction sites are underlined. The single difference base is indicated by an asterisk.

Due to the complex composition of fragments in digested genomic DNA, it is very inefficient to self-ligate these fragments and then transform circularized transgenes into *E. coli* cell despite target fragments being enriched. The recovered transgenes are only a fraction of transgenes with molecular weight ranging from 5 to 10 kb, and those regions responsible for ampicillin selection and plasmid replication are intact. These results strongly suggest that transgenes in F_4 transgenics are highly polymorphic.

(ii) Function of transgenes in F_4 transgenics.
PCR results showed that only in Type I transgene can the structure of *MThGH* fragment remain intact. Part of *MThGH* sequence in other three types of transgene is confirmed to be lost by Southern blot. Of course, transgene will be expressed and results in proper biological consequences if only there are a few intact transgenes integrated where are suitable for expression. In the cases of transgenic mouse[11], drosophila[12]and fish[7], it has been demonstrated that the expression level of transgene is significantly affected by the location of transgene. Body weight of the F_4 transgenic fish selected for this project was 3.5 times as large as the average of 17 individuals. This fact argues that the dramatic fast growing effect of this transgenics was due to a few intact transgenes and their proper location on chromosomes. We believe that some transgene integrations in this individual are genetically functional as described in our previous paper[7].

(iii) The method of plasmid rescue useful to analyzing the structural feature of transgene integration site.
To gain information on the sequences of integration sites, probes from transgene were used to screen genomic DNA library[2, 3]. It is well known that constructing and screening genomic DNA library are complex and expensive. It sounds impracticable to construct genomic DNA libraries for every transgenic individual in case that we need to analyze the transgene integration and inheritance for several transgenics.

It is well believed that transgenes, arraying in a head-to-tail manner, are integrated into host genome at random sites. Single copy of transgene can be released from concatemer by digestion with proper restriction enzyme. Released transgene at both ends of concatemer always traps the host sequences or sequences of integration site at its ends. Based on this feature, transgenes can be directly recovered from genome of transgenic animal. This method can be applied further in gene-transfer studies because the whole vector sequence may be reserved in most cases for purposes of properties screen and functional sequence protection from being deleted. In this note, two DNA fragments concerning integration sites were cloned using this method, and we have demonstrated its usefulness.

Acknowledgements We are grateful to Wang Yaping for his constructive suggestion and discussion, and to Hu Wei, Sun Yonghua and Wang Wei for their kindly assistance throughout the whole project. This work

was supported by the National Natural Science Foundation of China (Grant No. 39730290).

References

1. Brinster, R. L., Chen, H. Y., Trumbauer, M. E., Somatic expression of Herpes Thymidine Kinase in mice following injection of a fusion gene into eggs, Cell, 1981, 27: 223.
2. Covarrubias, L., Nishida, Y., Mintz, B., Early postimplantation embryo lethality due to DNA rearrangements in a transgenic mouse strain, Proc. Natl. Acad. Sci. USA, 1986, 83: 6020.
3. Hamada, T., Sasaki, H., Seki, R. et al., Mechanism of chromosomal integration of transgenes in microinjected mouse eggs: sequence analysis of genome-transgene and transgene-transgene junctions at two loci, Gene, 1993, 128: 197.
4. He, X., Liu, G., Chen, Q. et al., Exogenous gene localization on chromosomes of transgenic pig, Acta Zoologica Sinica (in Chinese), 1999, 26: 241.
5. Miao, C., Lu, G., Wang, Y. et al., Mapping BNLF-1 transgene on transgenic mouse progeny chromosomes by fluorescence *in situ* hybridization, Acta Genetic Sinica (in Chinese), 1998, 25: 422.
6. Zhang, J., Lao, W., Zhang, X. et al., Genetics and expression stability of exogenous gene construct in transgenic mice, Acta Genetic Sinica (in Chinese), 1999, 26: 135.
7. Zhu, Z., Xu, K., Xie, Y. et al., A model of transgenic fish, Scientia Sinica, Ser. B (in Chinese), 1989(2): 147.
8. Zhao, H., Chen, S., Sun, Y. et al., Time course of foreign gene integration and expression in transgenic fish embryos, Chinese Science Bulletin, 2000, 45(8): 734.
9. Maniatis, T., Fritsch, E. F., Sambrook, J., Molecular cloning, A Laboratory Manual, 2nd ed., New York: Cold Spring Harbor Laboratory Press, 1982, 9.31–9.37.
10. Collins, F. S., Weissman, S. M., Directional cloning of DNA fragments at a large distance from an initial probe: A circularization method, Proc. Natl. Acad. Sci. USA, 1984, 81: 6812.
11. Palimiter, R. D., Brinster, R. L., Transgenic mice (Minireview), Cell, 1985, 41: 343.
12. Howard-Flander, P., West, S. C., Stsiak, A., Role of recA protein spiral filament in genic recombination, Nature, 1984, 309: 215.

(Received June 29, 2000)