https://doi.org/10.47853/FAS.2022.e1



CgGH insertion functional domain analysis in transgenic G₁ and G₂ and G₃ mutiara catfish (*Clarias gariepinus*) broodstock

Ibnu Dwi Buwono^{*}, Roffi Grandiosa, Yuniar Mulyani

Aquaculture Department, Faculty of Fisheries and Marine Sciences, Universitas Padjadjaran, Jatinangor 45363, Indonesia

Abstract

Check for updates

Catfish is one of the most important freshwater fish farming commodities in Indonesia. Higher catfish production can be achieved by cultivating transgenic catfish carrying the growth hormone (*GH*) gene of African catfish (*Clarias gariepinus* GH, *CgGH*). This research focuses on analysis of the presence of the *CgGH* gene in transgenic G₁, G₂, and G₃ mutiara catfish broodstock, as an indication of stable *CgGH* inheritance. *CgGH* gene was isolated using the RNeasy mini kit and RT-PCR. RT-PCR revealed amplicons measuring approximately 600 bp in transgenic G₀, G₁, G₂, and G₃ mutiara catfish. The *CgGH* consensus sequence similarities ranged from 93.76% to 97.06%, with four functional domain sites (somatotropin-1, somatotropin-2, four α-helix, N-glycosylation, four cysteine residues) of fish GH proteins. The functional domains of fish GH proteins are conserved in G₁, G₂, and G₃ and indicate stable exogenous GH inheritance to produce transgenic catfish strains in each generation.

Keywords: CgGH, Transgenic line, Mutiara catfish, Conserved sequence, Consensus sequence

Introduction

The application of growth hormone (*GH*) transgenesis technology in fish has led to significant growth improvement as an effect of overexpression of the inserted *GH gene* (Hinits & Moav, 1999; Mori & Devlin, 1999; Nam et al., 2001). Transgenesis of *GH* causes excessive fish growth that is several times higher than that of non-transgenic fish, having the potential to increase fish culture yields. The growth of transgenic mutiara catfish (inserted *CgGH* sequence, GenBank accession no.MN249238.1) reported by Buwono et al. (2016; 2019b; 2021) was high (2–3 times that of non-transgenic fish) because of exogenous *GH* (*CgGH*) insertions at G_0 , G_1 , G_2 , and G_3 . This inserted transgene was successfully inherited in three generations through the reproduction of transgenic broodstock with *CgGH* transmission rates of 42.85% in G_1 , 50% in G_2 and 70% in G_3 (Buwono et al., 2021). Transgene inserts integrated into the fish genome have been shown to be inheritance on to offspring through broodstock reproduction. Some examples include studies on trout showing the inheritance of pSV518 in G_2 by 49%–75% (Tewari

Received: Jun 22, 2021 Revised: Sep 7, 2021 Accepted: Oct 26, 2021 *Corresponding author: Ibnu Dwi Buwono

Aquaculture Department, Faculty of Fisheries and Marine Sciences, Universitas Padjadjaran, Jatinangor 45363, Indonesia Tel: +62-22-87751918, E-mail: ibnu.dwi.buwono@unpad.ac.id

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. Copyright © 2022 The Korean Society of Fisheries and Aquatic Science

et al., 1992). Transmission of the antifreeze protein gene (AFP) that was inherited in G_2 by the transgenic G_1 Atlantic salmon broodstock pair following Mendel's law by 50% showed stable inheritance in the transgenic salmon generation (Hew et al., 1992). The percentage of transgene transmission from crosses of G_1 transgenic and non-transgenic Nile tilapia fish (*Oreochromis niloticus*) ranged 49%–52% which is consistent with Mendelian inheritance (Rahman & Maclean, 1999). Transgene inheritance of 50% from the broodstock to the offspring of this transgenic fish possible it to be transmitted to the offspring through the reproduction of the broodstock fish.

Based on the alignment of CgGH sequences from transgenic mutiara catfish, G₀, G₁, G₂, and G₃ can exhibit sequence similarities among all four generations, reflecting the consistency in CgGH inheritance in each generation (Degani et al., 2006; Pinheiro et al., 2008; Rajesh & Majumdar, 2007). Confirmation of the presence of CgGH in G₁, G₂, and G₃ transgenic mutiara catfish broodstock to ensure transgene transmission in each reproductive offspring. It is important to evaluate the similarity of functional domains of CgGH sequences in the offspring of this transgenic catfish to achieve mass production of transgenic catfish lines. The similarity of functional domains in the CgGHsequence may indicate the similarity of the GH protein molecules formed between the four generations of transgenic mutiara catfish.

Materials and Methods

Fish used in this study

Fish used this study were kept at the Aquaculture Laboratory Universitas Padjadjaran in circular tanks containing 1,000 liters of freshwater. Fish were adapted to 12 hours' daylight photoperiod conditions. Rearing conditions were kept at optimal water quality which were 27 ± 1 °C, pH 6–6.5 and optimal dissolved oxygen (continuous aeration). As much as 10% of water was replaced with freshwater while in the same time faeces and feed left over were siphoned. All fish were daily fed with commercial feed (Hiprovite 781) with the dose of 3% total biomass. Fish used for the research were the G₀ \Leftrightarrow broodstock (weighing 1,600 g, total length of 58 cm and age of 12 months), G₁ \Leftrightarrow broodstock (weighing 1,200 g, total length 55 cm and age of 12 months), G₂ \Leftrightarrow (weighing 950 g, total length of 50 cm and age of 11 months), G₃ \Leftrightarrow (weighing 910 g, total length of 46 cm and age of 11 months).

Isolation of CgGH

RNA was isolated from 10 mg of fish tail fin tissue of G_0 , G_1 , G_2 , and G₃ was isolated using the RNeasy mini kit (Qiagen, Venlo, Netherlands), following the kit instructions for RNA isolation. Sampling was carried out on the tail fin of the broodstock, not taken from gonadal tissue or liver because the broodfish was used for the production of the next generation. Transgenes (including CgGH) can be inserted in fish tissues along the head to tail region (Rahman & Maclean, 1999; Uh et al., 2006). Synthesis of cDNA and RT-PCR (semi-quantitative PCR) of CgGH were performed using My Taq OneStep RT-PCR (Bioline, London, UK) with the following cycling programme: 48° C for 20 min; 40 cycles of 95 $^\circ$ C for 1 min, 95 $^\circ$ C for 10 s, 60 $^\circ$ C for 30 s and 72 $^\circ\!\mathrm{C}\,$ 30 s; and 72 $^\circ\!\mathrm{C}\,$ for 5 min. Confirmation of transgenic catfish was achieved by detecting the presence of a PCR product of approximately 600 bp using primers GH-F (5'-ATGGCTC-GAGTTTTGGTGCTGCT-3') and GH-R (5'-CTACAGAGT-GCAGTTGGAATCCAGGG-3') (Buwono et al., 2021; Zhang et al., 2009).

Sequencing of CgGH

The *CgGH* gene amplicon (PCR product of G_0 , G_1 , G_2 sample A and B, G_3 sample A, B, and C) was then sequenced using the Sanger sequencing method through service 1st BASE (Molecular Biology Company) Singapore because the amplicon size is less than 1,200 bp (Sanger et al., 1977). Nucleotide sequence similarity analysis of the *CgGH* gene in G_0 , G_1 , G_2 , and G_3 transgenic mutiara catfish was performed using the BioEdit 7.0.5.3 software (http://www.mbio.ncsu.edu/BioEdit/BioEdit.html) to identify the consensus sequences of the four *CgGH* nucleotide similarity analysis, we used the consensus sequence, to equate the complementary sequence from the forward direction, so that this consensus sequence could be used to *CgGH* sequence similarity analysis to G_0 , G_1 , G_2 , and G_3 as written in Table 1 and Fig. 2.

Table 1. Consensus similarity of the CgGH sequences between transgenic mutiara catfish¹⁾

Nucleotide (nt)	<i>CgGH</i> -G ₁ (%)	<i>CgGH</i> -G ₂ (%)	<i>CgGH</i> -G ₃ (%)
<i>CgGH</i> -G₀	93.76	93.78	95.15
CgGH-G1		97.06	96.42
CgGH-G ₂			96.29

¹⁾ analyzed with BioEdit 7.0.5.3 (pairwise alignment). GH, growth hormone.

Functional domain analysis of the CgGH sequence

The presence of *CgGH* in transgenic catfish was confirmed by aligning the gene sequence with the *Clarias gariepinus GH* coding sequence (cds) in GenBank to determine the similarity of the nucleotide base sequences encoding the fish *GH* protein (Peyush et al., 2000). *CgGH* nucleotide base sequence similarity among the generation of transgenic mutiara catfish was identified using the BLAST (Basic Local Alignment Search Tools) programme (http://www.ncbi.nlm.nih.gov/BLAST/). The functional domains of *CgGH* sequences, especially amino acid residues, were analysed using SWISS-MODEL (https://swissmodel. expasy.org/) to map the molecular structure of fish *GH* protein.

Results

CgGH amplicons and CgGH consensus sequence

ThePCR analysis results showed that CgGH was amplified using the primers GH-F and GH-R in the test samples G₀, G₁, G₂, and G₃ and had a size of approximately 600 bp (Fig. 1). The size of this amplified sequence was also not markedly different from that of the *Clarias gariepinus* GH sequence in GenBank (accession nos. EF411172 and MN249238.1), with sizes of 603 and 615 bp, respectively. This indicated that the *CgGH* gene sequence present in the four generations of transgenic mutiara catfish is the *GH* gene of *C. gariepinus*. Consensus sequence analysis results of the four generations of transgenic mutiara catfish also showed high nucleotide similarity of *CgGH* consensus sequences (Table 1 and Fig. 2).

CgGH functional domain

Considering that G_0 transgenic mutiara catfish is a germline transmitter and G_1 fish is produced from a transgenic × non-transgenic cross, functional domain analysis of *CgGH* (forward direction) was conducted between G_2 transgenic mutiara catfish resulting from crossing A (sample A coded 1st_BASE_3044995 and sample B coded 1st_BASE_3044997) (Buwono et al., 2019a) with G_3 transgenic mutiara catfish sample A (transgenic female-1 × transgenic male-1), B (transgenic female-2 × transgenic male-2) and C (transgenic female-3 × transgenic male-3) (Buwono et al., 2021) using the Sanger method for the sequence process while the alignment used the CLUSTALW BioEdit software. The results of the analysis showed that the *CgGH* gene sequences (forward direction) of thw G_3 fish (samples A, B, and C) with G_2 fish showed high similarity (96.21%, 96.38%, and 95.91%, respectively).

Differences in CgGH nucleotides between G_2 (code 1st_BASE_3044995 sample A) and G_3 fish sample A (1st_BASE_3527728_A) occurred of 16 nucleotides in the first



Fig. 1. Electropherogram of the G_0 , G_1 , G_2 , and G_3 transgenic mutiara catfish (marked by the arrow *CgGH* 600 bp). 1kb, DNA ladder 1kb; 1, sample from fin of G_0 \bigcirc transgenic broodstock; 2, sample from fin of G_0 \bigcirc transgenic broodstock; 3, sample from fin of G_0 \bigcirc transgenic broodstock; 4, sample from fin of G_0 \bigcirc non-transgenic broodstock; 5, sample from fin of G_0 \bigcirc transgenic broodstock; 6, sample from fin of G_0 \bigcirc transgenic broodstock; 7, sample from fin of G_1 \bigcirc transgenic broodstock; 8, sample A from fin of G_2 \bigcirc transgenic broodstock; 10, sample A from fin of G_3 transgenic broodstock; 11, sample B from fin of G_3 transgenic broodstock; 12, sample C from fin of G_3 transgenic broodstock; 13, pCMV-*CgGH* plasmid; GH, growth hormone.

20 30 40 50 60 70 Cons.seq.CaGH-G0 -----AACTTATGGCTC--GAGTTTTGGTGCTGCTCYCWGTKGTGKTKG-CGAGTCTGTTCTTTAATCAA 62 Cons.seq.CgGH-G1 Cons.seq.CgGH-G2 CACTGCTATTATCGCTC--GAGTTTTGGTGCTGCTCWCWRTKGTGGTCKGCGAGTCTGTTCTTTAATCAA 68 Cons.seq.CgGH-G3 Clustal Consensus ***** ******* ********* 80 90 100 110 120 130 Cons.seq.CgGH-G0 GGCGCGGACATTTGAGACCCCAGCGGCTCTTCAACAACGCGGTCATCCGTGTGCAACACCTTCACCAACTGG 132 Cons.seq.CgGH-G1 GGCGCGACATTTGAGACCCAGCGGCTCTTCAACAACGCGGTCATCCGTGTGCAACACCTTCACCAACTGG 132 Cons.seq.CgGH-G2 GGCGCGACATTTGAGACCCAGCGGCTCTTCAACAACGCGGTCATCCGTGTGCAACACCTTCACCAACTGG 138 Cons.seq.CgGH-G3 GCCCCCACATTTGAGACCCAGCGCTCTTCAACAACGCGGTCATCCGTGTGCAACACCTTCACCAACTGG 132 Clustal Consensus 150 160 170 180 190 200 210 Cons.seq.CgGH-G0 CTGCCAAGATGATGGATGGATGACTTTGAAGAAGCTTTGTTACCTGAAGAACGCAAACAGCTGAGCAAGATCTT 202 Cons.seq.CgGH-G1 CTGCCAAGATGATGATGACTTTGAAGAAGCTTTGTTACCTGAAGAACGCAAACAGCTGAGCAAGATCTT 202 CTGCCAAGATGATGGATGACTTTGAAGAAGCTTTGTTACCTGAAGAACGCAAACAGCTGAGCAAGATCTT 208 Cons.seq.CqGH-G2 Cons.seq.CgGH-G3 CTGCCAAGATGATGGATGACTTTGAAGAAGCTTTGTTACCTGAAGAACGCAAACAGCTGAGCAAGATCTT 202 ***** 190
 220
 230
 240
 250
 260
 270
 280

 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ... Cons.seq.CqGH-G0 Cons.seq.CgGH-G1 Cons.seq.CgGH-G2 CCCCCTGTCATTCTGCAACTCGGACTCTATCGAGGCTCCGGCAGGCCAGGACGAGACCCAGAAAAGCTCC 278 Cons.seq.CgGH-G3 CCCCCTGTCATTCTGCAACTCGGACTCTATCGAGGCTCCGGCAGGCCAGGACGAGACCCAGAAAAGCTCC 272 Clustal Consensus 290 300 310 320 330 340 · · Cons.seq.CgGH-G0 GTGCTGAAACTGCTGCACACATCTTATCGTCTGATCGAGTCATGGGAGTTCCCCCAGCAAGAACCTGGGCA 342 Cons.seq.CgGH-G1 GTGCTGAAACTGCTGCACACATCTTATCGTCTGATCGAGTCATGGGAGTTCCCCAGCAAGAACCTGGGCA 342 GTGCTGAAACTGCTGCACACATCTTATCGTCTGATCGAGTCATGGGAGTTCCCCAGCAAGAACCTGGGCA 348 Cons.seq.CgGH-G2 Cons.seq.CgGH-G3 GTGCTGAAACTGCTGCACACATCCTATCGTCTGATCGAGTCATGGGAGTTCCCCCAGCAAGAACCTGGGCA 342 Clustal Consensus 360 370 380 390 400 410 420 ACCCTAACCATATCTCTGAAAAGCTGGCTGACCTGAAAATGGGCATCGGTGTGCTTATTGAGGGATGTGT 412 Cons.seq.CgGH-G0 Cons.seq.CgGH-G1 ACCCTAACCATATCTCTGAAAAGCTGGCTGACCTGAAAATGGGCATCGGTGTGCTTATTGAGGGATGTGT 412 Cons.seq.CgGH-G2 ACCCTAACCATATCTCTGAAAAGCTGGCTGACCTGAAAATGGGCATCGGTGTGCTTATTGAGGGATGTGT 418 Cons.seq.CgGH-G3 ACCCTAACCATATCTCTGAAAAGCTGGCTGACCTGAAAATGGGCATCGGTGTGCTTATTGAGGGATGTGT 412 Clustal Consensus *** 399 430 440 450 460 470 480 Cons.seq.CgGH-G0 GGATGGACAAACCAGCCTGGACGAGAATGACGCATTTGCTCCGCCCTTCGAGGATTTCTACCAGACCCTG 482 GGATGGACAAACCAGCCTGGACGAGAATGACGCATTTGCTCCGCCCTTCGAGGATTTCTACCAGACCCTG 482 Cons.seq.CgGH-G1 Cons.seq.CgGH-G2 GGATGGACAAACCAGCCTGGACGAGAATGACGCATTTGCTCCGCCCTTCGAGGATTTCTACCAGACCCTG 488 GGATGGACAAACCAGCCTGGACGAGAATGACGCATTTGCTCCGCCCTTCGAGGATTTCTACCAGACCCTG 482 Cons.seq.CgGH-G3 Clustal Consensus ** 469 500 510 520 530 540 550 5 Cons.seq.CgGH-G0 Cons.seq.CgGH-G1 Cons.seq.CqGH-G2 Cons.seq.CgGH-G3 Clustal Consensus ***** ************** 539 610
 570
 580
 590
 600
 610

 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ...
 ... 620 CTTATCTCAGCGTGGCCAAGTGCMGGWGATCCCTGGATTCCAACTGCACTCTG-TAGAGNAGCA----- 615 Cons.seq.CqGH-G0 CTTATCTCAGCGTGGCCAAGTSMAGAWGRTCCCTGGATTCCAACTGCACTCTG-TAGAGCAGCACCAAAA 621 Cons.seq.CgGH-G1 Cons.seq.CgGH-G2 CTTATCTCAGCGTGGCCAAGTGCAGAAGATCCCTGGATTCCAACTGCACTCTG-TAGAGCAGCACCAAAA 627 CTTATCTCAGCGTGGCCAAGTGCAGGAGATCCCTGGATTCCAACTGCACTCTGATAGAGCAGCACCAAAA 622 Cons.seq.CgGH-G3 595 Cons.seq.CgGH-G0 ----- 615 Cons.seq.CgGH-G1 CTCGAGCCCTTTATAAAAG----- 640 Cons.seq.CgGH-G2CTCGAGCCATATAAAAG----- 644 Cons.seq.CgGH-G3 CTCGAGGC-CATAAAA----- 637 Clustal Consensus

Fig. 2. Consensus alignment of the *CgGH* sequences in G₀, G₁, G₂, and G₃ (ClustalW multiple alignment analysis BioEdit). ATG, start codon; TAG, stop codon; GH, growth hormone.

sequence, and in nucleotides 594, 597, 600, 606 and 610 (Fig. 3A). At nucleotide numbers 594, 597, and 600 of G_2 fish there was a lack of cytosine, thymine and cytosine residues, and at nucleotide 606 of G_3 fish, there was a lack of thymine residue.

The start codon (ATG) and stop codon (TAG) were located at nucleotides 18–20 and 576–578, respectively. In G_3 fish of sample B (1st_BASE_3527730_B), there was a difference of 16 nucleotides in the first sequence of *CgGH* with G_2 fish (code

(A)		10 20	30	40	50	60	70
()	lst_BASE_3044995_Cg_F	ACAATTGTGGTCTGCGAATGT	GTTCTTTAT CAGGC	GCGACATTTGAGA	CCCAGCGGCT	CTT CAACA ACG	2 70
	lst_BASE_3527728_A_Cg-F	CHINN NUMBER NUMBER NUMBER M	GTTCTTTATCAGGC C S L S G	GCGACATTTGAGA A T F E	CCCÃGCGGCT TQRI	CTT CAACA ACG	C 70
	lst_BASE_3044995_Cg_F	SC S	ICACCAACT GGCTG	CCAAG ATGAT GGA	120 TGACTTTGAR	AGAAGCTITGTI	140 1A 140
	lst_BASE_3527728_A_Cg-F	GGTC ATCCG TGTGCA ACACCT V I R V Q H L	ICACCAACT GGCTG H Q L A	CCAAGATGATGGA A K M M I	TGACTTTGAN D F E	GAAGCTITGIT E A L L	A 140
		150 160	170	180	190	200 :	230 I
	1st_BASE_3044995_Cg_F	P E E R K Q L	AGCAAGATCTTCCC	L S F C	N S D	S I E A	C 210
	lst_BASE_3527728_A_Cg-F	CCTGAAGAACGCAAACAGCTG P E E R K Q L 	AGCAAGATCTTCCC SKIFP		N S D	S I E A	C 210
		220 230	240 	250	260 -	270 :	280
	1st_BASE_3044995_Cg_F	P A G K D E T	Q K S S V	L K L L	H T S Y	R L I	E 280
	1st_BASE_3527726_A_Ug-F	P A G K D E T	Q K S S V	LKLL	H T S Y	R L I	E 280
		290 300	310 	320 	330	342 :	550
	1st_BASE_3044995_Cg_F	S W E F P S K	N L G N	P N H I S	E K L	A D L R	A 350
	130 <u>4885</u> 5517710 <u>17</u> 0g 1	SWEPPSK	NLGN	PNHIS	3 E K L -	A D L R	
	1st BASE 3044995 Co F	360 370	350	390	400	420 4 1 - 1	620 I IG 420
	lst BASE 3527728 A Co-F	M G I G V L I ATGGGCATCGGTGTGCTTATT	E G C V D	G Q T S	L D E	N D A F	G 420
	,-	MGIGVLI	EGCVD	GQTS	LDE	NDAF	_
	lst_BASE_3044995_Cg_F	430 440 	450 	450 GAGGG GAACT TG?	470 	450 A	490 1 10 10 10 10 10 10 10 10 10 10 10 10
	lst_BASE_3527728_A_Cg-F	A P P F E D F CTCCGCCCTTCGAGGATTTCT A P P F E D F	YQTLS ACCAGACCCTGAGC YQTLS	E G N L GAGGGGAACTTGA E G N L	R K S H AGGAAGAGCTI R K S H	? R L L CCGTCTGCTGT ? R L L	S C 490 S
	-	500 510	520	530	540	550 :	560
	<pre>lst_BASE_3044995_Cg_F</pre>	TTGCTTCAAGAAAGACATGCA	CAAAGTGGAGACTT	ATCTCAGCGTGGG	CANGTGCAG	AGATCCCTGGA	T 560
	lst_BASE_3527728_A_Cg-F	C F K K D M H	K V E T	ATCTCAGCGT GGC Y L S V 2	CAAGTGCAGG	R S L D	T 560
	1-5 BACK 2044025 C- F	570 550	590	600	610 - -		
	12C_DV9F_30443A2_Cd_E	S N C T L * S	S T K T	R A I +	K COS		
	1st_BASE_3527728_A_Cg-F	S N C T L + S	AGCACCAAAACTCC	FINGC CCATA - AT + P I I	TAAA 610		

Fig. 3. Continued.

(B)	<pre>lst_BASE_3044997_Cg_F lst_BASE_3527730_B_Cg-F</pre>	10 20 30 40 50 60 70 -CCA GTT GGTT IGCG AATG IGT TCTT TATCAGGC GCG ACAT ITGA GAC CCAG CGGC TCTT CAA CAAC GCG 69 M C S L S G A T F E T Q R L F N N A NINNIN NGGG NINNIN NATG IGT TCTT TATCAGGC GCG ACAT ITGA GAC CCAG CGGC TCTT CAA CAAC GCG 70 M C S L S G A T F E T Q R L F N N A
	<pre>lst_BASE_3044997_Cg_F lst_BASE_3527730_B_Cg-F</pre>	50 90 100 110 120 130 140 GTCATCCGIGT GCAACACCTTC ACCAACTG GCTG CCA AGAT GATG GAT GACTTTGA AGAA GCTTTGTTAC 139 V I R V Q H L H Q L A A K M D D F E A L L GTCATCCGTGT GCAA CACCTTC ACCAACTG GCTG CCA AGAT GATG GATG
	<pre>lst_BASE_3044997_Cg_F lst_BASE_3527730_B_Cg-F</pre>	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
	<pre>lst_BASE_3044997_Cg_F lst_BASE_3527730_B_Cg-F</pre>	ZZD ZSO Z40 ZSO ZSO ZFO ZFO
	<pre>lst_BASE_3044997_Cg_F lst_BASE_3527730_B_Cg-F</pre>	ZEC SOO SLO SZO SZO
	<pre>lst_BASE_3044997_Cg_F lst_BASE_3527730_B_Cg-F</pre>	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
	<pre>lst_BASE_3044997_Cg_F lst_BASE_3527730_B_Cg-F</pre>	$\begin{array}{cccccccccccccccccccccccccccccccccccc$
	<pre>lst_BASE_3044997_Cg_F lst_BASE_3527730_B_Cg-F</pre>	500 510 520 530 540 550 560 IGCTICA AGA A AGA CATGO CACA AAGT GOAG AAGT ATCTCAG COTGGOCC AAGT GOAG AAGA TCCCTGGATT IGCTICA AGA A AGA CATGO CACA AAGT GOAG AAGT CACA AAGT GOAG AAGT ATCTCAG COTGGOCC AAGT GOAG GAAGA TCCCTGGATT C F K K M H K V E T Y L S V A K C R S L D IGCTICA AGA AGA AGA CATGO ACA AAGT GOAG AAGT GOAG AAGT GOAG GAAGA AGA CATGO ACA AAGT GOAG ACTT ATCTCAG COTGGOCC AAGT GOAG GAAGA TCCCTGGATT 560 C F K K D M H K V E T Y L S V A K C R S L D C F K K M H K V E T Y L S V A K C R S L D
	<pre>lst_BASE_3044997_Cg_F lst_BASE_3527730_B_Cg-F</pre>	570 580 580 600 CCAACTG CACT CTGT AGAG CAG CACC AAAA CT CG AGC CCTTTATA AAAG 608 S N C T L S S K T R A L Y K CCAACTG CACT CTGT AGAG CAG CACC AAAA CT CG AGG CC

Fig. 3. Continued.

(-)		
	1st_BASE_3044995_Cg_F	ACAATTGTGGTCTGCGAATGTGTTCTTTATCAGGCGCGAATTTGAGACCCAGCGGCTCTTCAACAACGC 70 M C S L S G A T F E T O R L F N N A
	1st_BASE_3527732_C_Cg-F	-NNNNNNNNNGNGAATGTGTTCTTTATCAGGCGCGACATTTGAGACCCÀGCGGCTCTTCAACAACGC 66 MCSLSGATFEETQRLEFNNA
		50 90 100 110 120 130 140
	1st_BASE_3044995_Cg_F	GGTCATCCGTGTGCAACACCTTCACCAACTGGCTGCCCAAGATGATGGATG
	lst_BASE_3527732_C_Cg-F	GGTCATC CGTGT GCÂACACCTT CACCAACTGG CTGCCAAGATGATGATGACTTTGAAGAAGCTTTGTTA 136 V I R V Q H L H Q L A A K M M D D F E E A L L
		150 160 170 150 190 200 210
	1st_BASE_3044995_Cg_F	$ \begin{array}{c} \texttt{CCTGAAGAACGCAAACAGCTGAGCAAAGATCTT CCCCCTGTCATTCTGCAACTCGGACTCTATCGAAGGCTC 210} \\ \texttt{P} \texttt{E} \texttt{E} \texttt{R} \texttt{K} \texttt{Q} \texttt{L} \texttt{S} \texttt{K} \texttt{I} \texttt{F} \texttt{P} \texttt{L} \texttt{S} \texttt{F} \texttt{C} \texttt{N} \texttt{S} \texttt{D} \texttt{S} \texttt{I} \texttt{E} \texttt{A} \end{array} $
	1st_BASE_3527732_C_Cg-F	CCTGAAGAACGCAAACAGCTGAGCAAGATCTT CCCCCTGTCATTCTGCAACTCGGACTCTATCGAGGCTC 206 P E E R K Q L S K I F P L S F C N S D S I E A
		220 230 240 250 250 250 270 250
	lst_BASE_3044995_Cg_F	CGGCAGGCAAGGACGAGACCCAGAAAAGCTCCGTGCTGAAACTGCTGCACACATCTTATCGTCTGATCGA 280 P A G K D E T Q K S S V L K L L H T S Y R L I E
	1st_BASE_3527732_C_Cg-F	CGGCAGGCAAGGACGAGACCCAGAAAAGCTCCGTGCTGAAACTGCTGCACACACA
		280 300 310 320 330 340 350
	1st_BASE_3044995_Cg_F	GT CATGG GAGTT COCCAGCAAG AACCT GGGCAACCCTAACCAT ATCTC TGAAA AGCTG GCTGA CCTGAAA 350 S W E F P S K N L G N P N H I S E K L A D L K
	1st_BASE_3527732_C_Cg-F	GT CATGEGAGTT COCCAGCAAG AACCT GGGCAACCCTAACCAT ATCTC TGAAAAGCTGGCTGACCTGAAA 346 S W E P P S K N L G N P N H I S E K L A D L K
		380 370 380 390 400 410 420
	1st_BASE_3044995_Cg_F	AT GGGCATCGGT GTGCTTATTG AGGGATGTGT GGATGGACAAACCAGCCTGGACGAGAATGACGCCATTG 420 M G I G V L I E G C V D G Q T S L D E N D A F
	1st_BASE_3527732_C_Cg-F	AT GGGCATCGGT GTGCTTATTG AGGGATGTGGATGGACGACAAACCAGCCTGGACGAGAATGACGCCATTG 416 M G I G V L I E G C V D G Q T S L D E N D A F
		430 440 450 480 470 480 490
	1st_BASE_3044995_Cg_F	CT CCGCC CTTCG AGGATTTCTA CCAGA CCCTG AGCGAG GGGAA CTTGA GGAAG AGCTT CCGTC TGCTG CCGCC 490 A P P F E D F Y Q T L S E G N L R K S F R L L S
	1st_BASE_3527732_C_Cg-F	CT COGCC CTTCG AGGATTTCTA CCATA COCTG AGGAG GGGAA CTTGATGAAG AGCTT CCGTC TGCTG TC 48 6 A P P F E D F Y H T L S E G N L M K S F R L L S
		500 51.0 520 550 540 550 560
	1st_BASE_3044995_Cg_F	TTGCTTCAAGAAAGACATGCACAAAGTGGAGACTTATCTCAGCGTGGCCAAGTGCAGAAGATCCCTGGAT 560 C F K K D M H K V E T Y L S V A K C R R S L D
	1st_BASE_3527732_C_Cg-F	C F K K D M H K V E T Y L S V A K C R R S L D
		570 550 590 600
	1st_BASE_3044995_Cg_F	TCCAACTGCACTCTGTAGAGCAGCACCAAAACTCGAGCCATATAAAAG 608 S N C T L * S S T K T R A I * K
	lst_BASE_3527732_C_Cg-F	ICCAACTGCACTCTGTAGAGCAGCACCAAAACTCGAACC-AT-AAA- 600 SNCTL*SSTKTRTI

Fig. 3. *CgGH* functional domain in G_2 and G_3 transgenic mutiara catfish. (A) Forward alignment of the *CgGH* in G_2 (1st_ BASE_3044995 sample (A) with those in G_3 sample A (1st_BASE_3527728_A). (B) Forward alignment of the *CgGH* in G_2 (1st_ BASE_3044997 sample (B) with those in G3 sample B (1st_BASE_3527730_B). (C) Forward alignment of *CgGH* in G_2 (1st_ BASE_3044995 sample (A) with those in G3 sample C (1st_BASE_3527732_C). CNSDSIEAPAGKDETQKSSVLKLLHTSYRLIE SW, somatotropin-1 site; CFKKDMH KVETYLSVAKC, somatotropin-2 site; LFNNAVIRVQHLHQ LAAKMMDDFEEALLP, Helix-1; TSYRLIESWEFPSKNLGNPNHIS, Helix-2; GIGVLIEGRVDGQTSLDENDAFAPPF, Helix-3; KDMHKVETYL SVAKCRRSLDSNCT, Helix-4; NCTL, N-Glicosylation; CCCC, 4-Cysteine residue; GH, growth hormone. 1st_BASE_3044997 sample B). In G₃ fish, nucleotides 600–602 lacked three thymine residues and nucleotide 609 lacked a guanine residue (Fig. 3B). The start codon, ATG, was located at nucleotides 17–20, and the stop codon, TAG, at nucleotides 575–577. The results of the alignment of the *CgGH* sequence of G₂ fish (code 1st_BASE_3044995 sample A) with that of the G₃ fish of sample C (1st_BASE_3527732_C) showed a difference of 14 nucleotides in the first sequence, where the start codon was located at nucleotides 18–20 and the stop codon at nucleotides 571–573 (Fig. 3C). In sample G₃, nucleotides 600, 601, 604, and 608 lacked adenine, thymine, adenine and guanine residues, respectively. The amino acid residues in Figs 3A–C are underlined and marked with coloured boxes after conversion using SWISS-MODEL to form a three-dimensional structure of the *GH* protein molecule.

Overall alignment of the sites of somatotropin-1 (nucleotide 186-287), somatotropin-2 (nucleotide 492-548), N-glycosylation (nucleotide 564-575), 4-residue cysteine (nucleotide 181, 493, 544, 598), helix -1 (nucleotide 56-143), helix-2 (nucleotide 261-329), helix-3 (nucleotide 354-431) and helix-4 (nucleotide 501-572) were contained in the CgGH sequence in both G2 of sample A (1st_Base_304495) and G3 fish sample A (1st_BASE_3527728_ A) were located on the same nucleotide (nt.) (Fig. 3A). Meanwhile, the alignment of the CgGH functional domains between in G2 of the sample B (1st_Base_3044997) and the G3 fish sample B (1st_Base_3527730) were located on different nucleotides. In G2 of the sample B, the somatotropin-1 site at nt. 186–286, somatotropin-2 at nt. 491-547, N-glycosylation at nt. 563-574, 4-cysteine residue at nt. 186, 492, 543, 562, helix-1 at nt. 56-141, helix-2 at nt. 260-328, helix-3 at nt. 353-430 and helix-4 at nt. 500–574, while in G3 of the sample B, the somatotropin-1 site at nt. 187-287, somatotropin-2 at nt. 492-548, N-glycosylation at nt. 564-575, 4-cysteine residue at nt. 187, 493, 544, 563, helix-1 at nt. 57-142, helix-2 at nt. 261-329, helix-3 at nt. 354-431 and helix-4 at nt. 501-575 (Fig. 3B). The alignment of the position of the CgGH functional domain between in G2 of the sample A (1st_ Base_3044995) and the G3 fish sample C (1st_Base_3527732) was also located on different nucleotides. Somatotropin-1 (nt. 187–287), somatotropin-2 (nt. 492–548), N-glycosylation site (nt. 564-575), 4-cysteine residue (nt. 187, 497, 544, 568), helix-1 (nt. 57-143), helix-2 (nt. 261-329), helix-3 (nt. 354-431) and helix-4 (nt. 501-572) in G2 of the sample A, while in G3 fish sample C is located at different nucleotides, namely somatotropin-1 (nt. 183-283), somatotropin-2 (nt. 488-544), N-glycosylation (nt. 560-571), 4-cysteine residue (nt. 183, 493, 540, 564), helix-1 (nt. 53-139), helix-2 (nt. 257-325), helix-3 (nt. 350-427) and helix-4 (nt. 493-564) (Fig. 3C).

Discussion

The presence of CgGH in four generations of transgenic mutiara catfish indicates that the exogenous GH gene is inherited stably in each generation of GH-transgenic catfish. The rate of CgGH transmission in G₂ transgenic mutiara catfish was 50% and increased in G₃ to 70% (Buwono et al., 2021). This indicates the potential for increased CgGH transmission in crosses between G4 transgenic catfish as a consequence of the stability of CgGH inheritance in transgenic catfish offspring. Homozygous transgenic fish need to be produced to obtain stable transgene inheritance (Iwai et al., 2009). Homozygous fish were produced when crossing between heterozygous G₂ mud loach (*Misgurnus mizolepis*) transgenic fish (carrying CMV-H2B-GFP) to produce 50% homozygous G₃ progeny (Nam et al., 2000).

To confirm its stable inheritance, the stability of the CgGH sequence needs to be analysed for similarity as an indication that its copies in G₀, G₁, G₂, and G₃ transgenic mutiara catfish have high similarities between generations. Yang et al. (2018) also explained that the coding region in the gene sequences are generally conserved and have high similarities with those of related fish species. There was a high homology of the gene encoding the hormone oxytocin, which regulates GH release in ricefield eel (Monopterus albus), being 84.6% identical to that of Anguilliformes (Anguilla bicolor). High homology was shown in the gene sequences encoding GH in C. gariepinus G_0 fish compared with those in G_1 fish (93.76%), G_2 fish (93.78%) and G_3 fish (95.15%), indicating that the nucleotide sequence of CgGH did not change much and was conserved (Table 1). In addition, there was a tendency for an increase in the homology of CgGH sequences between G₁ and G₂ and G₃ fish by 97.06% and 96.42%, respectively. The results of another study also indicated that the GH sequences of blue gourami (Trichogaster trichopterus) and pearl gourami (T. leeri) showed high homology as conserved sequences, at 97% and 96%, respectively (Degani et al., 2006). The same study also showed that the Indian catfish (Heteropneustes fossilis) GH sequence had high homology (98%) with the Siluridae and Clariidae groups (Anathy et al., 2001). It was shown that the CgGH sequence was conserved with high homology (93.76%–97.06%) in four generations of transgenic mutiara catfish, which was required for stable exogenous GH inheritance in the transgenic fish generations. The consistency of CgGH consensus sequence homology in G_0 , G_1 , G_2 , and G_3 fish, especially at the start codon (ATG) and stop codon (TAG), is shown in Fig. 2, indicating that the coding sequence of CgGH is conserved in the generation of transgenic mutiara catfish.

Functional domains are conserved sequences that characterise a particular gene group consisting of 40-700 amino acid residues (Xiong, 2006). Generally, five functional domains characterise fish GH sequences (somatotropin-1 and somatotropin-2, N-glycosylation, four a-helix structure and four cysteine residues), which are homologous (Anathy et al., 2001; Pinheiro et al., 2008). The results of SWISS-MODEL processing showed that the four characteristic sites of the GH molecule (somatotropin-1, somatotropin-2, a helix-1 to a helix-4, N-glycosylation and four cysteine residues) in the CgGH sequences of G₂ and G₃ fish were conserved and located at the same base pairs. According to Anathy et al. (2001) and Pinheiro et al. (2008), ahelix-1 is encoded by amino acid residues LFNNAVIRVQHL-HQLAAKMMDDFEEALLP (underlined in blue), a helix-2 by TSYRLIESWEFPSKNLGNPNHIS (underlined in gold), a helix-3 by GIGVLIEGRVDGQTSLDENDAFAPPF (underlined in red) and a helix-4 by KDMHKVETYLSVAKCRRSLDSNCT (underlined in green). These four helix structures are bound by four cysteine residues (marked with red circles). The α -helix site is a domain that indicates the formation of a secondary structure of GH protein, namely, α -helix sites 1 to 4, which are important for the functional activity of these hormones (Pinheiro et al., 2008). Generally, this domain has relatively high homology among the GH of freshwater fish (including the catfish group). The amino acid residues CNSDSIEAPAGKDE-TQKSSVLKLLHTSYRLIESW (marked with purple box) are the functional domain of somatotropin-1, and the amino acid residues CFKKDMHKVETYLSVAKC (marked with yellow box) are the functional domain of somatotropin-2. The existence of these two functional domains is related to GH activity and synthesis of insulin-like growth factor-1and prolactin for tissue growth. The somatotropin sites (1 and 2) present in the GH sequences of transgenic mutiara catfish (G_2 and G_3) and in Indian catfish (Anathy et al., 2001) are both conserved. N-glycosylation site domains encoded with NCTL amino acid residues (marked with pink boxes) in GH protein sequences were also found to be conserved in fish (including transgenic mutiara catfish) and act as signals for protein transport to the cell surface (Degani et al., 2006). Another important site is four cysteine residues (C) in the protein-coding GH gene, which are involved in the formation of two disulphide bonds for the structural integrity and biological activity of the hormone (Anathy et al., 2001); the five functional domain sites were found to be conserved in CgGH of G_2 and G_3 transgenic mutiara catfish (Figs 3A–C). The molecular structure of GH protein in G_3 mutiara catfish (samples A–C), shown in Fig. 4A–C, was confirmed as a GH protein molecule (Swiss model analysis), as shown in the Siluriformes group GH protein molecule (Vaz et al., 2010).

These results indicate that CgGH from G_s broodstock had been successfully inherited in up to three generations (G_1 , G_2 , and G_3) with a high degree of similarity and confirmed as fish *GH* protein. This verification was based on analysis of the functional domains of fish *GH* molecules composed of somatotropin-1 and somatotropin-2, four-helix structures, N-glycosylation and four cysteine residues that bind to the helix structure (Pinheiro et al., 2008; Vaz et al., 2010).

Conclusion

CgGH (600 bp) can be inherited in G₁, G₂, and G₃ transgenic mutiara catfish through reproduction. The consensus sequence similarity of CgGH between transgenic fish in G₁ to G₃ ranged



Fig. 4. Three-dimensional model of the GH protein molecule from G₃. (A) G₃ sample A; (B) G₃ sample B; (C) G3 sample C; (D) GH protein molecule from *Siluriformes*. Adapted from Vaz et al. (2010) with CC-BY -NC-SA. GH, growth hormone.

from 93.76% to 97.06%, and they had five fish *GH* protein functional domain sites (somatotropin-1, somatotropin-2, four α -helix, N-glycosylation and four cysteine residues).

Competing interests

No potential conflict of interest relevant to this article was reported.

Funding sources

This work was supported by a grant from the Directorate of Research and Community Service of Universitas Padjadjaran for the research costs through the Internal Research Grant of Universitas Padjadjaran Number: 1959/UN6.3.1/PT.00/2021.

Acknowledgements

The authors are thankful to the Directorate of Research and Community Service of the Universitas Padjadjaran for support of research. The author would like to thank to the research team and research assistants (Rahayu Ramadhayanti and Ayuniar Puteri) thank you for the technical support during the research.

Availability of data and materials

Upon reasonable request, the datasets of this study can be available from the corresponding author.

Ethics approval and consent to participate

This article does not require IRB/IACUC approval because there are no human and animal participants.

ORCID

Ibnu Dwi Buwonohttps://orcid.org/0000-0002-5203-1509Roffi Grandiosahttps://orcid.org/0000-0003-0498-7220Yuniar Mulyanihttps://orcid.org/0000-0002-5521-049X

References

- Anathy V, Venugopal T, Koteeswaran R, Pandian TJ, Mathavan S. Cloning, sequencing and expression of cDNA encoding growth hormone from Indian catfish (*Heteropneustes fossilis*). J Biosci. 2001;26:315-24.
- Buwono ID, Iskandar, Agung MUK, Subhan U. Construction of transgenic catfish (*Clarias* sp.) using sperm electrophoration technique. J Biol. 2016;20:17-28.
- Buwono ID, Iskandar I, Grandiosa R. Growth hormone transgenesis and feed composition influence growth and protein

and amino acid content in transgenic G₃ mutiara catfish (*Clarias gariepinus*). Aquac Int. 2021;29:431-51.

- Buwono ID, Junianto J, Iskandar I, Alimuddin A. Growth and expression level of growth hormone in transgenic mutiara catfish second generation. J Biotech Res. 2019a;10:102-9.
- Buwono ID, Junianto J, Iskandar I, Alimuddin A. Reproduction performance of transgenic mutiara catfish (G_1) comprising the growth hormone gene. J Biotech Res. 2019b;10:199-212.
- Degani G, Jackson K, Yom-Din S, Goldberg D. cDNA cloning and mRNA expression of growth hormone in Belontiidae (*Anabantoidei suborder*). Isr J Aquac Bamidgeh. 2006;58:124-36.
- Hew CL, Davies PL, Fletcher G. Antifreeze protein gene transfer in Atlantic salmon. Mol Mar Biol Biotechnol. 1992;1:309-17.
- Hinits Y, Moav B. Growth performance studies in transgenic *Cyprinus carpio*. Aquaculture. 1999;173:285-96.
- Iwai T, Inoue S, Kotani T, Yamashita M. Production of transgenic medaka fish carrying fluorescent nuclei and chromosomes. Zool Sci. 2009;26:9-16.
- Mori T, Devlin RH. Transgene and host growth hormone gene expression in pituitary and nonpituitary tissues of normal and growth hormone transgenic salmon. Mol Cell Endocrinol. 1999;149:129-39.
- Nam YK, Cho YS, Chang YJ, Jo JY, Kim DS. Generation of transgenic homozygous line carrying the CAT gene in mud loach *Misgurnus mizolepis*. Fish Sci. 2000;66:58-62.
- Nam YK, Noh JK, Cho YS, Cho HJ, Cho KN, Kim CG, et al. Dramatically accelerated growth and extraordinary gigantism of transgenic mud loach *Misgurnus mizolepis*. Transgenic Res. 2001;10:353-62.
- Peyush P, Moriyama S, Takahashi A, Kawauchi H. Molecular cloning of growth hormone complementary DNA in barfin flounder (*Verasper moseri*). Mar Biotechnol. 2000;2:21-6.
- Pinheiro JS, Wolff JLC, Araújo RC, Hilsdorf AWS. Molecular cloning and sequence analysis of growth hormone cDNA of neotropical freshwater fish pacu (*Piaractus mesopotamicus*). Genet Mol Biol. 2008;31:381-4.
- Rahman MA, Maclean N. Growth performance of transgenic tilapia containing an exogenous piscine growth hormone gene. Aquaculture. 1999;173:333-46.
- Rajesh R, Majumdar KC. A comparative account of the structure of the growth hormone encoding gene and genetic interrelationship in six species of the genus Labeo. Fish Physiol Biochem. 2007;33:311-33.

- Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci. 1977;74:5463-7.
- Tewari R, Michard-Vanhee C, Perrot E, Chourrout D. Mendelian transmission, structure and expression of transgenes following their injection into the cytoplasm of trout eggs. Transgenic Res. 1992;1:250-60.
- Uh M, Khattra J, Devlin RH. 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. 2006;15:711-27.
- Vaz BS, Cerqueira GM, Silva JC, Manzke VHB, Moreira CGA, Moreira HLM. Sequence analysis of the growth hormone gene of the South American catfish *Rhamdia quelen*. Genet Mol Res. 2010;9:2184-90.
- Xiong J. 15 Protein tertiary structure prediction. In Xiong J, editor. Essential bioinformatics. Cambridge, UK: Cambridge University Press; 2006. p. 214-30.
- Yang W, Zhang N, Shi B, Zhang S, Zhang L, Zhang W. Isotocin regulates growth hormone but not prolactin release from the pituitary of ricefield eels. Front Endocrinol. 2018;9:166.
- Zhang M, Chen C, Guo Y, Guo J, Wang X. Cloning and sequence analysis of full-length growth hormone cDNA from *Clarias gariepinus*. Acta Agric Boreal Sin. 2009;24:27-32.