Integration of antimicrobial peptide genes via CRISPR/Cas9 for disease resistance enhancement and reversible sterility in catfish

by

Jinhai Wang

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
July 21, 2023

Keywords: Genome engineering, Transgenesis, Cathelicidin, Cecropin, Luteinizing hormone, Myostatin, Melanocortin-4 receptor, Disease resistance, Reproduction, Catfish

Approved by

Rex A. Dunham, Chair, Professor, School of Fisheries, Aquaculture, and Aquatic Sciences
Charles Y. Chen, Professor, Crop, Soil and Environment Sciences
Ian A.E. Butts, Associate Professor, School of Fisheries, Aquaculture, and Aquatic Sciences
Timothy Bruce, Assistant Professor, School of Fisheries, Aquaculture, and Aquatic Sciences

Abstract

The CRISPR/Cas9 platform holds promise for modifying fish traits of interest as a precise and versatile tool for genome manipulation. To reduce introgression of transgenes and control reproduction, catfish were studied for upscaled disease resistance coupled with intervention of reproduction to lower the potential environmental risks of introgression of transgenic escapees.

To generate disease resistance and sterility in channel catfish (*Ictalurus punctatus*), CRISPR/Cas9 systems were utilized to integrate the cathelicidin gene from an alligator (*Alligator sinensis*; *As-Cath*) into the target luteinizing hormone (*lh*) locus of channel catfish using two delivery systems assisted by double-stranded DNA (dsDNA) and single-stranded oligodeoxynucleotides (ssODNs), respectively. High knock-in (KI) efficiency (22.38%, 64/286) but low on-target was achieved using the ssODN strategy, whereas adopting a dsDNA as the donor template led to an efficient on-target KI (10.80%, 23/213). On-target KI of *As-Cath* was instrumental in establishing the *lh* knockout (LH^{-}_As-$Cath^{+}$) catfish line, which displayed heightened disease resistance and reduced fecundity compared to the wild-type sibling fish. Furthermore, implantion with HCG and LHRHa restored the fecundity, spawnability and hatchability of the new transgenic fish line.

To establish disease resistance and sterility in blue catfish (*I. furcatus*), transgenic blue catfish of primarily Rio Grande strain ancestry were generated with site-specific KI of the *As-Cath* transgene into the *lh* locus via two CRISPR/Cas9-mediated KI systems, assisted by the linear dsDNA and double-cut plasmid (dcPlasmid), respectively. High integration rates were observed with linear dsDNA (16.67%, [13/78]) and dcPlasmid strategies (24.53%, [26/106]). In addition, the on-target KI efficiency of the dcPlasmid strategy (16.04%, [17/106]) was 1.67 times higher than that of the linear dsDNA strategy (10.26%, [8/78]) based on the odds ratio. The relative expression of the *As-Cath* transgene of P_1 founders was detected in nine tissues, dominated by the kidney, skin, and muscle (14.30-, 7.71- and 6.92-fold change, $P < 0.05$). Moreover, the *As-Cath* transgenic blue catfish showed a higher survival rate than that of wild-type controls (80% *vs*. 30%, $P < 0.05$) following *Flavobacterium covae* infection. Survival during culture supports

the challenge data as survival of *As-Cath* transgenic individuals was 97.1% while that of pooled non-transgenic individuals was observed to be less 87.0% (P = 0.15). The growth rates and external morphology of the transgenic and wild-type siblings were not different (P > 0.05), indicating no pleiotropic effects for growth of the *As-Cath* transgene integration at the *lh* locus was observed in the P_1 founders.

To generate transgenic channel catfish carrying two exogenous antimicrobial peptide genes (AMGs), CRISPR/Cas9-assisted microinjection of cecropin (*Cec*) and *As-Cath* was employed to create dual-AMG integrated (*_Cec^+/*_$Cath^+$) transgenic embryos with high integration rates. Additionally, a univariate-multiple logit regression model was fitted to determine the synergistic expression of transgenes and endogenous AMGs in the head kidney post-bacterial infection. Transgenic-embryo-based genome editing significantly increased the efficiency of dual-AMG integration from 17.6% to 37.3%. The survival rate of single-AMG (50% *vs.* 20%, P = 0.023) and dual-AMG (70% *vs.* 20%, P = 0.005) integrated fish was dramatically higher than that of wild-type fish (20%) following *Edwardsiella ictaluri* challenge. More dual-AMG fry survived than expected based on integration and inheritance rates of single-AMG transgenics compared to other genotypes. Logistic regression analysis indicated that individual body weight and gender did not affect survival, while the transgenes *Cec* and *As-Cath* contributed directly to the survival during the bacterial infection. Furthermore, transgenes enhanced fish disease resistance by regulating the expression of *TCP* and *NK-lysin* genes.

To establish transgenic sterile channel catfish lines with elevated disease resistance and fast growth rate, single-sgRNA-based genome editing (ssGE) and multi-sgRNA-based MGE (msMGE) were used to replace the *lh* and melanocortin-4 receptor (*mc4r*) genes with the *As-Cath* transgene and the myostatin (two target sites: *mstn1*, *mstn2*) gene with the *Cec* transgene, respectively. A total of 9,000 embryos were microinjected from three families, and 1,004 live fingerlings were generated and analyzed. There was no significant difference in hatchability (all P > 0.05) and fry survival (all P > 0.05) between ssGE and msMGE. Compared to ssGE, CRISPR/Cas9-mediated msMGE assisted by the mixture of dsDNA and dcPlasmid donors yielded a higher KI efficiency of *As-Cath* (19.93%, [59/296] *vs.* 12.96%, [45/347]; P = 0.018) and *Cec* (22.97%, [68/296] *vs.* 10.80%, [39/361]; P = 0.003) transgenes, respectively. The

msMGE strategy can be used to generate transgenic fish carrying two transgenes at multiple loci. In addition, double and quadruple mutant individuals can be produced with high efficiency (36.3% ~ 71.1%) in one-step microinjection.

Overall, the *lh* gene was replaced with the *As-Cath* transgene and then hormone therapy was administered to gain complete reproductive control of disease-resistant transgenic catfish in an environmentally sound manner. In addition, potential sterile catfish with enhanced disease resistance carrying two AMGs at multiple loci using transgenic-embryo-based genome editing or msMGE strategy was achieved. This strategy not only effectively improves the consumer-valued traits, but also guards against genetic contamination of wild populations. This is a breakthrough in aquaculture genetics to confine fish reproduction and prevent the establishment of transgenic or domestic genotypes in the natural environment.

Dedicated to my father Jianglin,

my mother Youzhen,

and my sister Haisi.

Acknowledgments

I would like to thank my advisor, Dr. Rex Dunham, for his guidance and mentorship during my graduate research tenure at Auburn University. I would also express my gratitude to my committee members: Dr. Charles Chen, Dr. Ian Butts, and Dr. Timothy Bruce for their expertise and helpful comments. My sincere gratitude also goes to Dr. Peng Zeng, for his valuable comments to improve my dissertation in statistical analysis, serving as my university reader.

I am grateful to Dr. Baofeng Su for teaching me the CRISPR system and guiding me in molecular experiments, Dr. Mei Shang for assisting me in sample collection. I would also like to thank other colleagues in the genetics lab for taking care of the fish and working in the pond.

I am forever grateful to my family, my mother Youzhen Wang, my father Jianglin Wang, and my sister Haisi Wang for their endless support, love, and encouragement.

Finally, I would like to thank myself for my perseverance.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AMP	Antimicrobial Peptide
AMG	Antimicrobial Peptide Gene
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
Cas9	CRISPR-associated Protein 9
sgRNA	Single Guide RNA
ssODN	Single-stranded Oligodeoxynucleotide
dsDNA	Double-Stranded Deoxyribonucleic Acid
HDR	Homology Directed Repair
NHEJ	Non-Homologous End Joining
LHRHa	Luteinizing Hormone-Releasing Hormone Analogue
HCG	Human Chorionic Gonadotropin

CHAPTER ONE

General Introduction

1. Introduction

Proliferating human populations and the demands for nutritious food have led to the realization that aquaculture is one of the most environmentally sustainable ways to produce food and protein. In addition, aquaculture directly drives year-round jobs for farmers and contributes to employment in other industries (Dunham and Elaswad, 2018), thereby providing economic growth (Naylor et al., 2021). In 2018, world aquaculture supplied around 82.1 million tons of fish, encompassing more than 425 farmed species, and bringing it to an all-time high (FAO, 2020). Despite impressive achievements, the aquaculture industry still faces persistent challenges. In particular, drug resistance and residues are one of the major concerns for sustainable aquaculture.

The drug residues in aquatic products are mainly due to the abuse and overuse of antibiotics, which have been associated with the emergence of multidrug-resistant foodborne pathogens, rendering treatment of infectious diseases in aquaculture and even humans ineffective (Dong et al., 2007; Karunasagar et al., 2020). The indiscriminate use of antibiotics lasted until the early 2000s, especially chloramphenicol, resulting in serious pathogen resistance (Cabello, 2006). Recently, FAO has published a detailed five-year plan (from 2021 to 2025) to restrict the use of antibiotics/antimicrobials in animal production (FAO, 2021). In addition, WHO also emphasizes avoiding the use of antibiotic-like drugs or medical-important antibiotics for preventive and therapeutic purposes in food-producing animals (WHO, 2017). Although antibiotic-free aquaculture might be overly optimistic as a goal, fortunately, numerous studies are being done to find alternatives to antibiotics.

Antimicrobial peptides (AMPs), promising substitutes of antibiotics (FAO, 2021), also known as cationic host defense peptides originated from virtually all domains of life, are naturally occurring innate-immune peptides with immunocompetence and homeostasis against bacteria, viruses, fungi, and parasites (Hancock, 2001; Wang et al., 2016; Wang et al., 2022). In addition

to natural AMPs, an increasing number of studies focus on developing shortened-synthetic AMPs to improve the stability and therapeutic index by harnessing rational design *in silico* (Torres et al., 2019; Okella et al., 2020). AMPs, whether natural or synthetical-modified, can eliminate pathogen invasion, and trigger the host's autoimmunity (Mookherjee et al., 2020). As wide-spectrum anti-pathogen activities and harmless candidates for antibiotics, versatile applications of AMPs have sparked great interest in the scientific community (Wang et al., 2022). Currently, AMPs are explored to be use in food preservation (Said et al., 2019), therapeutic drugs for human beings (Mookherjee et al., 2020), and feed additives as functional nutrition for animals (Silveira et al., 2021).

2. Categories and Properties of Catfish AMPs

2.1 Summary from APD3 database

The antimicrobial peptide database (APD3) has catalogued more than 3,000 natural AMPs from various organisms, including bacteria, archaea, protozoa, fungi, plants, and animals (Wang et al., 2016). A total of 3,167 AMP sequences were collected from the online APD3 (https://wangapd3.com/APD_sequence_release_09142020.fasta), which were employed as query sequences for putative AMP identification by BLASTP against channel catfish genome (Liu et al., 2016). A total of 605 putative AMPs were obtained, which can be divided into 51 classes according to annotation of AMPs in the APD3. Most of the identified catfish AMPs are homologous to mammals, except for 10 AMPs extracted and identified from catfish, which are species-specific. Histone-derived, thrombin-derived C-terminal peptides (TCPs), and chemokine-derived were the top three classes among them, accounting for 23.80%, 13.22% and 9.92% of the total AMP contents, respectively (Figure 1). Additionally, BPTI, Eotaxin, Reglllgamma and ShLvsG were found in the catfish genome as well, with percentages of the total AMPs of 2.81%, 2.31%, 2.15% and 1.65%, respectively. In addition, two AMPs, liver-expressed antimicrobial peptides-2 (LEAP-2) and hepcidin, which are low in content but extensively studied in catfish were observed.

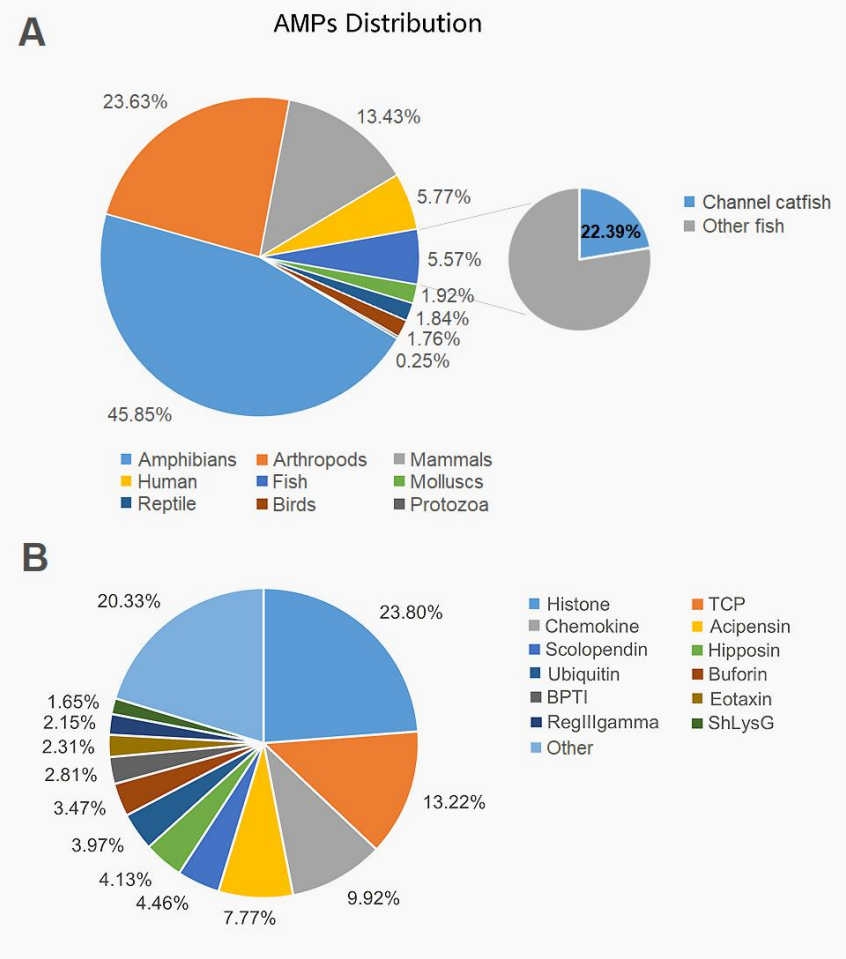


Figure 1. Distribution of the frequency of antimicrobial peptides (AMPs) discovered across various organisms. (**A)** The distribution of AMPs in animals. The number of AMPs from fish is similar to that from human. Channel catfish (*Ictalurus punctatus*) AMPs accounted for 22.39% of those from fish. (**B)** Proportion of the first 12 AMPs found in channel catfish based on the screening of database, relative to the total putative AMPs from the APD3 database screening.

2.2 Categories from the database

Based on the APD3 database, we categorized the AMPs in the channel catfish genome into the following 7 groups, including the first and largest histone-derived AMP family; buforin I, hipposin and acipensin (of which the first two originated from the N-terminus domain of histone H_2A); thrombin-derived C-terminal AMPs; chemokine-derived AMPs; scolopendins; ubiquitin-derived AMPs; and the last group BPTI, eotaxin, regIIIgamma and shlysG. They were introduced and discussed below based on the number of catfish AMPs found in each category.

2.2.1 Histone-derived AMPs

In general, histones are primarily involved in the regulation of DNA packaging, DNA replication and transcription (Strahl et al., 2000; Patat et al., 2004). Histones are the basic structural proteins of eukaryotic chromosomes, including four core proteins H_2A, H_2B, H_3, H_4 and one connexin H_1 (Strahl et al., 2000). The core histones are widely conserved, and the amino acid sequences are very similar even in the distantly related species, but histone H_1 is diversified (Ryu et al., 2021). During the past few decades, all five histones or histone-derived peptides as innate immune effectors have been evaluated from fish to humans. As early as 1958, Hirsch (1958) first reported that histones A and B purified from calf thymus exhibited antibacterial activity against various bacteria. Since then, such activity was described in aquatic animal tissues, such as Atlantic salmon liver (Richards et al., 2001), rainbow trout skin (Fernandes et al., 2002), Pacific white shrimp (*Litopenaeus vannamei*) (Patat et al., 2004), scallop hemocytes (*Chlamys farreri*) (Li et al., 2007), olive flounder (*P. olivaceus*) testes (Nam et al., 2012), and catfish skin (Park et al., 1998; Chen et al., 2017).

Fragments from histone H_1 of human epithelial cells were active against bacteria (Rose et al., 1998), they were rarely found in aquatic animals. Based on BLASTP, four core histones of catfish (accession numbers: XP_017348452.1, XP_017331187.1, XP_017345147.1 and XP_017336153.1) were found in the catfish genome with putative AMP fragments, and some of their functions have been verified by previous experiments. Robinette et al. (1998) determined that fragments from H_2B-like proteins in catfish skin had bactericidal and fungicidal activities. In addition to catfish H_2B-like proteins, the AMPs, such as parasin I, hipposin, and buforin derived from H_2A, were confirmed to possess antibacterial properties (Park et al., 1996, 1998; Birkemo et al., 2003). These studies indicated that complete histones may have some deficiencies when used as intracellular antibacterial factors due to their large molecular weight and functional areas that are not related to antibacterial activity. Therefore, the AMPs derived from histone as smaller fragments are more effective than the larger histone protein itself in animals.

2.2.2 Buforin I, Hipposin and Acipensin

AMPs can be derived from larger intact proteins with other biological functions or intrinsic antibacterial activities (Nguyen et al., 2011). For example, buforin I and hipposin are both derived from the N-terminus domain of histone H_2A (Park et al., 1996; Birkemo et al., 2003). Nguyen et al. (2011) have also reported that large proteins of overlap structure possess similar

functionality as AMPs. We found that the sequence of hipposin shows 96.1% similarity to acipensin 1; in addition, the same 3D structures are predicted due to the same domain with an overlay is present (Figure 2). The only difference between the two potential AMPs is an extra leucine residue at the C-terminal of hipposin, accompanied by an amino acid mutation (histidine mutates to glutamine). Similarly, we also observed buforin I (accession number: XP_017336154.1), hipposin (accession number: XP_017318461.1) and two variants of acipensin, acipensin 1 and acipensin 2 (accession numbers: XP_017318334.1 and XP_017318289.1) in the catfish proteome.

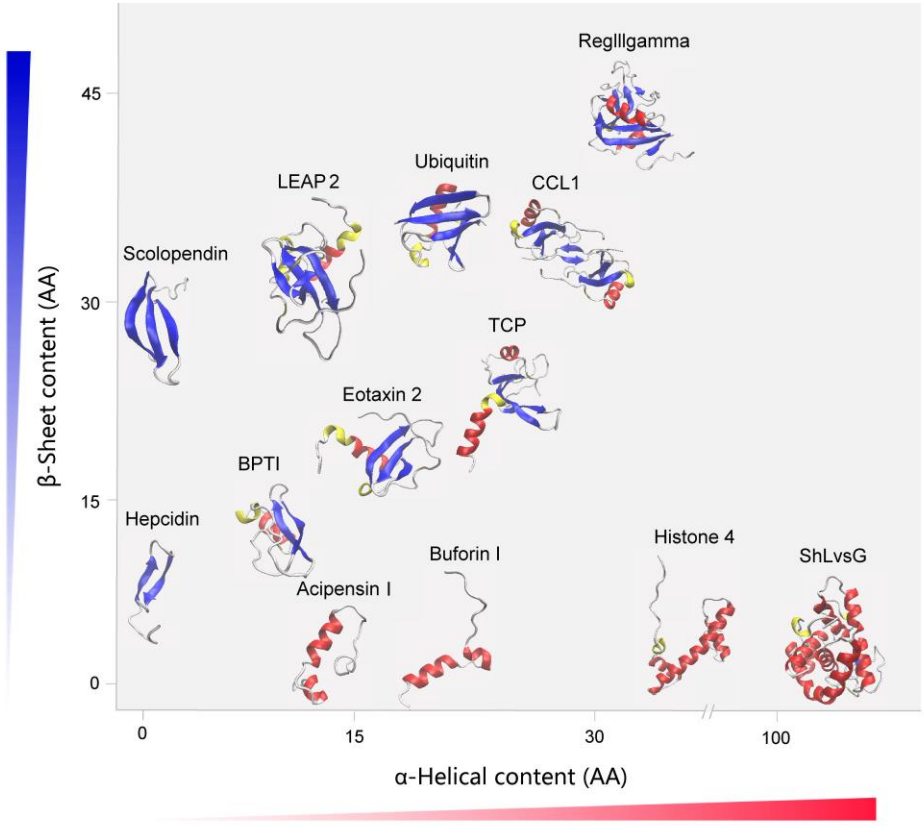


Figure 2. Examples of the diversity of catfish AMPs. Tertiary structures of the main peptides are arranged by secondary structure contents. α-helices are shown in red, β-strands shown in blue, and 3(10)-helices are shown in yellow. The unit of coordinate scale is the number of amino acids (AAs), with marginal legends indicating the number of secondary structures, α-helix in x-axis and number of β-sheet in y-axis. LEAP2, liver-expressed antimicrobial peptide 2; CCL1, chemokine, CC family; TCP, Thrombin-derived C-terminal peptide; BPTI, bovine pancreatic trypsin inhibitor; ShLvsG, G-type lysozyme; AA, amino acid. The 3D protein structural models were produced by SWISS-MODEL prediction algorithm (http://swissmodel.expasy.org) and visualized VMD software (https://www.ks.uiuc.edu/Research/vmd/).

2.2.3 Thrombin-derived C-terminal Peptides

A key enzyme in the coagulation cascade, thrombin-derived C-terminal peptide (TCP) was first detected *in vivo* in human wounds, exerting potent anti-endotoxic effects (Papareddy et al., 2010), and interestingly, TCP homologues were also observed from the channel catfish genome (accession number: XP_017314082.1). Like most classical AMPs, TCPs exhibit physicochemical properties, such as amphipathicity, cationicity and α-helicity (Papareddy et al., 2010; Petrlova et al., 2017). Additionally, synthetic shorter peptides from the C-terminus of TCP are antibacterial and adopt a helical conformation in lipid environments (Wang et al., 2016). TCP has double roles during wounding and infection conditions; TCP helps fibrin clot in the coagulation system and also inhibits macrophage responses to bacterial lipopolysaccharides (LPS) by binding to LPS. Such binding facilitated the formation of helical structures and permeabilization of liposomes (Papareddy et al., 2010). A more recent study further revealed the significance of TCP in host defense as TCP can bind to LPS or gram-negative bacteria, leading to bacterial permeation and amyloidation/aggregation, which promotes a clearance by phagocytic uptake (Petrlova et al., 2017). As a potential application of a prototypic thrombin-derived peptide, GKY25 showed bactericidal activity against a mouse model of *Pseudomonas aeruginosa* via reduction of both systemic cytokine responses and excessive coagulation (Kalle et al., 2012). Furthermore, another homologue of GKY25 from C terminus of human thrombin negatively regulated endotoxin-induced responses, inhibition of TLR4- and TLR2-elicited NF-κB activation and reduction of proinflammatory cytokine production in macrophages and monocytes (Hansen et al., 2015).

2.2.4 Chemokine-derived AMPs

Chemokines are a group of small structurally related proteins that are important participants and modulators of a variety of physiological and pathological processes, including inflammation and immunity (Baggiolini, 1998; Rollins, 1998), regulating cell migration and activation, and acting as a bridge between innate and adaptive immunity (Zlotnik and Yoshie, 2000; Zlotnik et al., 2006; Raman et al., 2011). They are usually classified into four subfamilies: C, CC, CXC and CX3C (or CXCL, CCL, CX3CL and XCL) based on the arrangement of the conserved first four-cysteine motifs in N-terminal (Zlotnik and Yoshie, 2000). Notably, not all chemokines have

antibacterial properties, and antimicrobial chemokines usually form a large positively charged electrostatic patch on the surface of the molecule (Yang et al., 2003).

Initially, seven phylogenetic groups of channel catfish CC chemokines/[CC chemokine ligand (CCL)] were identified and characterized, including the CCL19/21/25 group, the CCL20 group, the CCL27/28 group, the CCL17/22 group, the macrophage inflammatory protein (MIP) group, the monocyte chemotactic protein (MCP) group and a fish-specific group (Peatman and Liu, 2007). With the availability of the channel catfish genome resource, Fu et al. (2017) found that 64 CCLs were divided into 11 distinct classes, making this species with the largest number of CCLs in fish. Additionally, based on our BLASTP result, the putative AMPs derived from catfish chemokines include CCL (CCL1, CCL8, CCL11, CCL13, CCL18, CCL19, CCL20, CCL21 and CCL 25) and CXCL (CXCL6, CXCL10, CXCL11, CXCL12 and CXCL13). Fu et al. (2017a) reported the expression of one of the CCL19 members, CCL19a.1, was up-regulated post *E. ictalurid* infection but significant upregulation was observed in all CLL19 members (CCL19a.1, CCL19a.2 and CCL19b) after *Flavobacterium columnare* infection, suggesting CCL19 contributes to fish immune responses. Similarly, another 17 CXC chemokine members were identified in the channel catfish genome and their expression were gene-specific after bacterial infections. CXCL11.3 and CXCL20.3 were two potential molecular markers for ESC resistance after ESC infection, while CXCL20.2 was suggestive as a molecular marker for columnaris infection (Fu et al., 2017b). Recently, CXCL20a/b in grass carp (*Ctenopharyngodon idella*) has been demonstrated to have potent antimicrobial activity *in vitro* and *in vivo*, and oral administration of CXCL20a nanopeptide can efficiently prevent bacterial infection in fish (Xiao et al., 2020; Zhang et al., 2021; Wang et al., 2022). Such diversity in chemokine molecules may provide enriched resources for discovery of new AMPs and attract more interest towards exploration of this gene family as immune regulators.

2.2.5 Scolopendins

Scolopendins were identified as a broad-spectrum AMP from adult centipedes using RNA sequencing, including scolopendin 1 and 2 (Choi et al., 2014; Lee et al., 2015). Scolopendin 1 contains 52 amino acid residues with sequence similarity to cecropin peptides, while scolopendin 2 is shorter and has 16 amino acids in length. Compared with scolopendin 2, scolopendin 1 exhibits stronger antibacterial activity, and neither induce haemolysis at high concentrations

(Wang et al., 2016). Although they are from the same animal species, they exhibit various mechanisms against microbial pathogens: scolopendin 1 exerts antifungal activity by triggering the apoptotic pathway (Lee et al., 2017) and scolopendin 2 kills bacteria by affecting cell membranes (Lee et al., 2015). At present, no such AMP has been found in aquatic animals except giant grouper (*Epinephelus lanceolatus*) (Wang et al., 2019); however, we identified both scolopendin 1 and scolopendin 2 (accession numbers: XP_017306585.1 and XP_017323132.1) as putative catfish AMPs based on BLASTP search.

2.2.6 Ubiquitin-derived AMPs

As an adenylate cyclase stimulating polypeptide, ubiquitin was first isolated from bovine thymus (Schlesinger and Goldstein, 1975). Like the chemokine family, it is highly conserved in living organisms from protozoans to vertebrates (Hicke, 2001). Recently, ubiquitin has been reported to play an important role in innate immune defenses via a bacteriostatic process (Seo et al., 2013), and several ubiquitin-derived AMPs have been isolated from the secretions of bovine-stimulated chromaffin cells, the human amniotic fluid and the gill of Pacific oyster (Kieffer et al., 2003; Kim et al., 2007; Seo et al., 2013). In fish, ubiquitination regulates cellular processes, which is vital for the generation of innate and adaptive immune responses to pathogens (Viswanathan et al., 2010; Huang et al., 2017). Here, the putative AMP derived from catfish ubiquitin (accession number: XP_017315121.1) is identical in sequence to oyster cgUbiquitin, implying a similar antibacterial effect.

2.2.7 BPTI, Eotaxin, RegIIIgamma and ShLysG

In addition to the potential AMPs mentioned above, another four AMPs were screened out. Bovine pancreatic trypsin inhibitor (BPTI) was originally determined to be an enzyme inhibitor with high stability related to temperature, acidity and high pH (Fritz and Wunderer, 1983). Recently, it was found to be an antifungal AMP with a novel mechanism of action by inhibiting Mg^{2+} uptake (Bleackley et al., 2014). Eosinophil-recruiting chemokines (Eotaxins), such as eotaxin-1 (CCL11), eotaxin-2 (CCL24) and eotaxin-3 (CCL26), have potent bactericidal activity against both Gram-positive and Gram-negative pathogens (Gela et al., 2015). As another broad-spectrum bactericidal peptide, regIIIgamma inhibits deleterious bacteria and regulates the spatial relationships between microbiota and host to promote host-bacterial mutualism (Vaishnava et al., 2011). ShLvsG is an ortholog of goose-type lysozyme, characterized from seahorse

(*Hippocampus abdominalis*) (Ko et al., 2016), and lytic activities of this AMP indicate it is an antimicrobial effector of the innate immune system by hydrolyzing the wall of bacterial cell (Hikima et al., 2001; Wang et al., 2012). Although these four AMPs have not been identified or isolated from fish species, we determined that there were four potential corresponding AMPs from the catfish proteome (accession numbers: XP_017313286.1, XP_017328327.1, XP_017350515.1 and XP_017328787.1).

2.3 Isolated AMPs from catfish

In addition to the putative AMPs obtained by using the APD3 database, we summarized the catfish-derived AMPs that have been cloned, isolated or identified. So far, 11 catfish AMPs/AMGs have been extracted or sequenced in the past two decades (Table 1).

The first purified AMP from catfish was histone-like (H_2B-like) antimicrobial protein (HLP-1) isolated from the skin, which was similar to trout histone H_2B with 82% homology (Robinette et al., 1998). Almost at the same time, another AMP extracted from the skin mucosa was parasin I, containing 19 amino acids (Park et al., 1998), and 18 of the 19 residues were identical to the N-terminal of buforin I due to the same origin from histone H_2A. The third AMP, haemoglobin-derived peptide (HbβP-1), was composed of 33 amino acid residues from the gill tissue of channel catfish, and its expression was upregulated in the skin and gills after parasitic infection (Ullal et al., 2008). Additionally, a novel catfish AMP, pelteobagrin, was identified and isolated from the epidermal mucus of yellow catfish (*Pelteobagrus fulvidraco*), which contained 20 amino acids and it was predicted to carry a positive charge of +2 with 60% hydrophobic amino acid content (Su, 2011). Subsequently, four important catfish AMGs were cloned and sequenced, namely bactericidal permeability-increasing protein (BPI), LEAP-2, natural killer lysin (NK-lysin) and hepcidin (hamp) (Bao et al., 2005, 2006; Wang et al., 2006). BPI is primarily resistant to Gram-negative bacteria as a lipid transfer/LPS-binding protein and has at least 57% homology with other teleost fish. Channel catfish has only one copy of BPI gene, whereas rainbow trout has two (Xu et al., 2005). Mature catfish LEAP-2 is a cysteine-rich peptide composed of 41 amino acids, which is similar to other fish, especially zebrafish and rainbow trout (Su, 2011). NK-lysin has three subtypes in catfish (Petrlova et al., 2012), which shares 44.6% amino acid identity with zebrafish NK-lysin sequence (Wang et al., 2006). Similar to these three sequenced genes, the catfish hepcidin only has one copy and is also conserved compared to other species. Some

Table 1. Catfish AMPs/AMGs and their antibacterial, antifungal and antiparasitic activities.

AMPs/AMGs	Fish species/Weight	Pathogen G^+ (Gram-positive)	G^- (Gram-negative)	Fungi/Parasites	MIC/PC or AMP gene expression	References
Cathepsin D (regulation of parasin I)	CC (32.5 g)		*Edwardsiella ictaluri*		Gene expression upregulated at 24 and 48 h after infection.	Gisbert et al., 2011
β-defensin	CC (62.5 ± 7.3 g)		*E. ictaluri*		Gene expression increased significantly at 48 h after infection.	Krogdahl, et al., 2005
HbβP-1	CC		*Vibrio alginolyticus* *Aeromonas hydrophila* *Escherichia coli*	*Ichthyophthirius multifiliis* (Ich)	MIC = 50 μg/ml (*V. alginolyticus*) MIC = 12.5 μg/ml (*A. hydrophila*) MIC = 12.5 μg/ml (*E. coli*) PC_{min} = 6.3 μg/ml (ich trophonts) PC_{100} = 12.5μg/ml (ich trophonts)	Yu et al., 2010
Hepcidin	CC (32.5 g) (larvae)		*E. ictaluri*		1. Gene expression upregulated at 4, 6, 12, 24 and 48 h after infection. 2. Gene expression upregulated 1–3 days after infection but returned to basal level at 7 days.	Pascual et al., 2005
HLP-1	CC		*V. alginolyticus* *A. hydrophila* *Escherichia coli*	*Saprolegnia* spp.	MIC = 50 μg/ml (*A. hydrophila*) MIC = 0.37 μM±0.023 μM (*E. coli*)	Hornick et al., 2000
LEAP-2	CC and BC (32.5 g) (larvae)		*E. ictaluri*		1. Gene expression not significantly changed. 2. Gene expression upregulated (less than two folds) at 4 h after infection.	Sacristán et al., 2016; Gisbert et al., 2011
Parasin I	AC	*Pseudomonas putida* *Bacillus subtilis* *Staphylococcus aureus* *Streptococcus mutans*	*E. coli* *Salmonella typhimurium* *Serratia* spp.	*Cryptococcus neoformans* *Saccharomyces cerevisiae* *Candida albicans*	MIC = 2 μg/ml (*P. putida, S. aureus S. typhimurium, C. neoformans, S. cerevisiae*) MIC = 1 μg/ml (*B. subtilis, S. mutans, E. coli, C. albicans*) MIC = 4 μg/ml (*S.* spp.)	Lemieux et al., 1999
Pelteobagrin	YC	*Staphylococcus aureus* *Bacillus subtilis*	*E. coli*	*C. albicans*	MIC = 2 μg/ml (*B. subtilis*) MIC = 4 μg/ml (*S. aureus*) MIC = 16 μg/ml (*E. coli*) MIC = 64 μg/ml (*C. albicans*)	Muhlia-Almazán et al., 2002
NK-lysin	CC (32.5 g)		*E. ictaluri*		Gene expression upregulated at 24 and 48 h after infection.	Gisbert et al., 2011
BPI	CC (32.5 g) (larvae)		*E. ictaluri*		1. Gene expression upregulated at 48 h after infection. 2. BPI upregulation peaked 3 days after infection.	Furné et al., 2008; Gisbert et al., 2011

*Sa*RpAMP	AC	*Bacillus subtilis* *Staphylococcus aureus* *Micrococcus luteus* *Streptococcus iniae*	*E. coli* *S. enterica* *A. hydrophila* *Pseudomonas aeruginoa* *Shigella sonnei* *Klebsiella pneumonia* *Enterobacter cloacae*	*C. albicans*	MIC = 7.46 μg/ml (*B. subtilis*) MIC = 7.94 μg/ml (*S. aureus*) MIC = 11.07 μg/ml (*M. luteus*) MIC = 5.07 μg/ml (*E. coli*) MIC = 8.07 μg/ml (*S. enterica*) MIC = 7.26 μg/ml (*S. sonnei*) MIC = 6.93 μg/ml (*E. cloace*) MIC = 8.22 μg/ml (*P. aeruginosa*) MIC = 7.15 μg/ml (*K. pneumoniae*) MIC = 9.05 μg/ml (*S. iniae*) MIC = 8.65 μg/ml (*A. hydrophila*) MIC = 64 μg/ml (*C. albicans*)	Sakyi et al., 2020

Note: AMPs, antimicrobial peptides; AMGs, antimicrobial peptide genes; CC, channel catfish (*Ictalurus punctatus*); BC, blue catfish (*I. furcatus*); AC, amur catfish (*Silurus asotus*); YC, yellow catfish (*Pylodictis olivaris*); HbβP-1, haemoglobin-derived peptide; HLP, histone-like protein; BPI, bactericidal permeability-increasing protein; LEAP-2: liver-expressed antimicrobial peptide-2. PC_{min}, the minimum protozoacidal concentration is the lowest concentration that causes the death of at least one parasite. PC_{100}, 100% protozoacidal concentration is the lowest concentration where all parasites die. MIC, the minimum inhibitory concentration is defined as the lowest concentration that reduces bacterial growth by more than 50%.

researchers argue that LEAP-2 and hepcidin are identical in other species. However, based on the genomic contents and characterization of channel catfish, the precursor of the former is on chromosome 23 (location: 16548089-16548089) while the latter is on chromosome 7 (location: 20401983-20403674). Therefore, we regard them as two different AMPs encoded by two genes (accession numbers: AY845143, AY834209), respectively (Liu et al., 2016).

All of these four sequenced AMGs exhibit disease-resistance properties by increasing their expression when invaded by pathogens. Recently, another catfish AMP *Sa*RpAMP, was isolated from the skin of Amur catfish (*Silurus asotus*), which originated from the C-terminal region of 60S ribosomal protein L27 and composed of 33 amino acids (Oh et al., 2020).

2.4 Structures

Based on the 3D structures, the following 14 main catfish AMPs can be divided into α-helical AMPs (acipensin 1, hipposin, buforin 1, histone 4 and ShLvsG), β-sheet AMPs (scolopendin and hepcidin), and mixed α/β AMPs (reglllgamma, ubiquitin, eotaxin 2, LEAP-2, CCL1, BPTI and TCP) (Figure 2). The basic protein architecture of the majority of these AMPs is a short 3(10)-helix. All these 14 AMPs, except for ShLvsG, which is not charged, are cationic. Besides, catfish skin peptides, chromogranin A, HbbetaP and beta-thymosin are also not charged. In addition, we found another 6 anionic catfish AMPs, including the N-terminal region of surfactant protein B (SP-BN), beta-amyloid peptide, beta2-microglobulin, vasostatin and TroTbeta4. This confirms that catfish have a diverse AMP repository. Katzenback (2015) stated that fish AMPs can be divided into five categories according to their structures: cathelicidins, β-defensins, piscidins, hepcidins, and histone derived peptides. However, based on our BLASTP result, the putative innate AMPs of catfish do not completely correspond to these classifications, mainly because these peptides originate from different protein fragments that do not necessarily fall into these categories.

2.5 Properties

Length, net charge, and hydrophobicity are important structural properties that directly affect the function of AMPs (Hancock and Scott, 2000; Yin et al., 2012). However, these characteristics have not been investigated in catfish-derived AMPs. Catfish AMPs range from 24 to 184 residues with an average of 62 amino acids. In general, most AMPs have been identified as small

peptides with around 50 amino acid residues (Valero et al., 2020), although recent research has found some exceptions (Mookherjee et al., 2020). Likewise, Hancock and Scott (2000) indicated that extended AMP fragments usually increase antimicrobial activity. Although catfish AMPs are charged with diversity, the average net charge is +5, which is consistent with the charge characteristics of most fish AMPs.

Hydrophobicity is usually kept within a certain range, and if it exceeds or falls below this limit, it causes hemolysis of the host cells (Mookherjee et al., 2020). For example, a fragment of cathelicidin 1 of Atlantic salmon is slight hemolytic with a 27% hydrophobic ratio (HR) (Wang et al., 2016), but H_2A is highly hydrophobic with a HR of 50% and exhibits some hemolytic properties when the concentration reaches 0.3 μM (Brunner et al., 2020). From our summary, the hydrophobicity of catfish AMPs is between these two, and the average HR is 35.59%. Catfish parasin I and pelteobagrin have been shown to be non-toxic and non-hemolytic (Park et al., 1998; Su, 2011). Theoretically, they reduce the possibility of hemolysis if they are developed to treat diseases across species. Thus far, the categories and properties of catfish AMPs have been determined mainly based on human and mammalian studies, combined with catfish whole genome analysis compared with the APD3, but more research should be conducted to further verify their functionalities.

3. Antimicrobial activities of fish AMPs

3.1 Antibacterial activity

The antibacterial properties of AMPs have been intensively studied compared to other characteristics of AMPs, and of course, is of the greatest interest for potential aquaculture application. They protect the host from bacterial invasion via direct antimicrobial activity, which is attributed to their capacity to form pores and disrupt cell membranes (Olivieri et al., 2015; Mookherjee et al., 2020). Moreover, they also show an antibacterial function against antibiotic-resistant bacteria strains (Noga et al., 2009). Indeed, there are also many fish AMPs exhibiting antimicrobial functions against non-fish pathogens (review herein). The bactericidal activity of fish AMPs under *in vitro* conditions has been widely documented.

All the isolated catfish-derived AMPs or identified AMGs exhibited antibacterial properties against diverse pathogens (Table 1). Robinette et al. (1998) concluded that HLP-1 from the catfish skin had a strong bactericidal effect on aeromonads and *Aeromonas hydrophila* was

totally inhibited at 50 μg/ml, which was as efficient as magainin 1. However, the bactericidal efficacy of HLP-1 against *A. hydrophila* depends upon the strain of the bacterium, and different strains have different sensitivities. In addition, HLP-1 can kill *Escherichia coli* with the lethal concentration at 0.37 μM (Robinette et al., 1998).

Park et al. (1998) used ten pathogens to determine the minimum inhibitory concentration (MIC) of parasin I, an AMP derived from Amur catfish histone H_2A and it was an effective AMP for killing bacteria and fungi. Compared to buforin I or magainin 2, the MIC of parasin I was 1-4 μg/ml, which is only one-fourth or one-hundredth of the concentrations for buforin I or magainin 2, suggesting a stronger bactericidal efficacy. Although parasin I is homologous to histone H_2A, it is regulated by cathepsin D. Healthy catfish skin does not secrete parasin I due to cathepsin D inactivation (pro-cathepsin D). It is only when the skin is damaged that the injury stimulates a metalloprotease and the precursor of cathepsin D (pro-cathepsin D) to form the mature cathepsin D, which cleaves the Ser(19)-Arg(20) bond of H_2A to produce parasin I (Cho et al., 2002). This inducible mechanism has been observed in fish, including rainbow trout, loach (*Misgurnus anguillicandatus*) and eel (*Anguilla japonica*) (Cho et al., 2002). Molecular level studies suggested that the expression of cathepsin D was upregulated when catfish were infected by bacteria (Petrlova et al., 2012) with the purpose of producing more AMPs to resist the invasion of pathogens.

Similarly, the HbβP-1 from channel catfish also exhibited bactericidal function against *A. hydrophila* and *E. coli* with the concentrations of 12.5 μg/ml which was more efficient than HLP-1 (Ullal et al., 2008). Compared with parasin I, yellow catfish pelteobagrin can inhibit bacterial growth at higher concentrations against the same bacteria (*Staphylococcus aureus* and *Bacillus subtilis*), indicating that parasin I antibacterial activity is higher under these conditions. In addition, the most recently discovered highly effective catfish AMP, *Sa*RpAMP has been shown to inhibit 11 species of bacteria with the MIC at around 10 μg/ml (Oh et al., 2020).

3.2 Antifungal and antiparasitic activities

Although parasitic and fungal diseases are also common in fish, they have been less studied. In catfish infectious diseases, the most common parasite and fungi documented are *Ichthyophthirius multifiliis* (Ich) and *Candida albicans*, respectively. *C. albicans* is one of the most widely studied fungi in catfish, and parasin I, pelteobagrin and *Sa*RpAMP all inhibit this species with different

MICs: 1 μg/ml, 64 μg/ml and 64 μg/ml, respectively (Table 1). In addition, parasin I can kill fungi *Cryptococcus neoformans* and *Saccharomyces cerevisiae* at the concentration of 2 μg/ml, suggesting that fungi are sensitive to parasin I.

HLP-1 not only has bactericidal effect, but also inhibits fungi as it showed significant antifungal activity against *Saprolegnia* spp., which was stronger than that of magainin 2. HLP-1 significantly and continuously prevented *S. prasitica* germination to less than 5% in a 48-hour time course antifungal activity evaluation experiment (Zasloff, 1987; Robinette et al., 1998). Ullal et al. (2008) proved that Ich trophonts can be killed using HbβP-1 at the concentration of 6.3 to 12.5 μg/ml. Recently, piscidins derived from yellow croaker (*Larimichthys crocea*) have been shown to demonstrate parasiticidal activity against *Cryptocaryon irritans* by causing macronuclei swelling, cell membrane rupture and content leakage of theronts (Zheng et al., 2020, 2021).

3.3 Antiviral activity

AMPs are known as anti-viral agents, although reports in the literature focus more on their broad-spectrum anti-bacterial activity. At present, the study of antiviral properties mainly focuses on fish hepcidins, piscidins and defensins against a diversity of viruses. Chia et al. (2010) firstly reported the possible antiviral mechanism of fish AMPs by analyzing inactivation of nervous necrosis virus (NNV) by tilapia hepcidins, and they exhibited antiviral activity by agglutinating viral particles into clumps so that the viruses were blocked outside the cells. Additionally, the synthetic SA-hepcidin-2 derived from the spotted scat (*Scatophagus argus*) showed antiviral function against largemouth bass (*Micropterus salmoides*) reovirus (MsReV) and Chinese perch (*Siniperca chuatsi*) rhabdovirus (SCRV) (Gui et al., 2016). The effectiveness of piscidins purified from hybrid striped bass against channel catfish virus (CCV) and frog virus 3 (FV3) has been demonstrated by plaque assay at different concentrations (Chinchar et al., 2004). In the case of defensins, the β-defensin from the liver of the orange-spotted grouper was shown to reduce the Singapore grouper iridovirus (SGIV) and NNV infectivity (Guo et al., 2012). Furthermore, zebrafish β-defensin can reduce infection by the spring viraemia of carp virus (SVCV) accompanying the up-regulation of Mx (myxovirus resistance) gene expression after infection (García-Valtanen et al., 2014). Rainbow trout β-defensin exhibited strong virucidal

ability in regard to the viral haemorrhagic septicaemia virus (VHSV) when it was transmitted into the epithelioma papulosum cyprinid (EPC) cell line with a plasmid (Falco et al., 2008).

4. Factors affecting AMG expression

4.1 Pathogenic infection

The activation of AMPs is usually accompanied by the regulation of antimicrobial peptide gene (AMG) expression. An increasing number of studies have proved that AMG expression increases significantly when organisms are stressed by pathogens. As a genetic marker of fish-health and welfare, chronic stresses can lead to downregulation of AMPs of the fish (Noga et al., 2011). Therefore, the molecular response of AMGs is vital to fish survival and growth, especially in the context of infectious diseases and abiotic stress.

Hitherto, the expression of AMGs has been extensively researched in fish. As one of the most important members of AMG family, hepcidin has been widely investigated in fish, including medaka, Atlantic salmon, white bass, as well as catfish (Bao et al., 2005). In channel catfish, the expression of hepcidin gene was significantly upregulated in head kidney and spleen but not liver after 24-h *Edwardsiella ictalurid* challenge. However, after *Streptococus iniae* infection, the hepcidin of white bass was upregulated over 4500-fold in the liver (Shike et al., 2002), which indicated that the pattern of hepcidin expression is tissue-specific and infection specific. Bactericidal permeability-increasing protein (BPI) can be detected at the early stage of development and is widely expressed in various tissues of channel catfish. It was significantly upregulated within 48 h post-infection, and then remained at high levels for 3 to 7 days post-*E. ictalurid* treat (Xu et al., 2005; Petrlova et al., 2012). Similar to BPI, catfish LEAP-2 gene was expressed in a variety of tissues except brain and stomach, which was different from rainbow trout that showed expression in the liver and intestine (Zhang et al., 2004; Bao et al., 2006). This indicated that AMG expression was not only tissue specific but also organism specific. Cathepsins are a large group of proteases, which can be used as an enzyme to degrade damaged proteins to avoid the formation of cytotoxic aggregates (Zhou et al., 2018), and the expression of AMG effector cathepsin D was dramatically induced after infection of *E. ictaluri* and *F. columnare* (Li et al., 2012; Sun et al., 2014). Although the expression profiles of catfish AMGs are diverse, their expression is rapidly upregulated in a short period of time (12 h post-infection), suggesting that they are quickly involved in inflammatory responses to pathogen invasion in a

tissue-specific manner. Additionally, RNA-Seq data confirmed that two AMG families, cathepsins (Wang et al., 2013; Dong et al., 2016) and chemokines (Fu et al., 2017ab; Zhu et al., 2017) were induced after *E. ictaluri* infection in fish. Based on the summary, AMGs display tissue-specific or pathogen-specific patterns of expression, indicating distinct functions and disparate roles in immunity.

4.2 Abiotic stress

AMGs are not only induced by pathogen invasion, but a few studies indicate that they are also involved in physiological and genetic regulation under abiotic stress. This may be attributed to the increased disease incidents with exposures to environmental stresses, which induce innate AMPs to combat against them, especially when exposed to heat stress and hypoxia (Geng et al., 2014; Zhou et al., 2018). Heat shock proteins (HSPs) are produced by cells and are ascribed to heat stimuli, serving as powerful predictors of environmental stresses (Zhu et al., 2016). Members of HSP proteins, such as the HSP40 family, HSP70 family and the HSP90 family, were significantly up-regulated after heat challenge in fish (Liu et al., 2013). Similarly, one AMG family, cathepsins (cathepsin B, cathepsin D, cathepsin L and cathepsin Z) were slightly up-regulated to 2.2, 2.9, 4.5 and 2.6 folds in the gill after heat stimuli (Liu et al., 2013). Furthermore, the CC and CXC chemokine families were involved in the hypoxia as well as disease responses (Fu et al., 2017ab), indicating that similar pathways or mechanism of the responses to hypoxia and bacterial infection in fish, and this interdependence of hypoxic responses and innate immune responses have also been detected in mammals (Zinkernagel et al., 2007; Nizet and Johnson, 2009).

5. Applications of exogenous AMPs/AMGs in fish

AMPs/AMGs possess powerful potential to be applied in several industrial sectors, especially the health care industry, which is engaged in the treatment or prevention of many diseases in human as well as fish (Heuer et al., 2009; Cheung et al., 2015; Mookherjee et al., 2020; Valero et al., 2020). The current application of AMPs in reducing the cost of disease in the aquaculture sector mainly includes immunostimulants (Sakai, 1999).

Alternatively, it is theoretically feasible to transfer exogenous AMGs to the target organisms through transgenic technology to produce disease-resistant lines of specific species. Great

achievements have been accomplished in this direction in recent years. Some progress has been achieved by introducing exogenous AMGs of interest into fish of interest through genetic engineering or CRISPR/Cas9 mediated targeted gene insertion, which would be a sustainable prophylactic strategy for acquiring hereditary high disease resistance. In this sense, a new modified donor genome generated by CRISPR/Cas9 can insert coding sequences of foreign AMGs into the appropriate regions of target genome, potentiating AMP expression at the specific sites.

5.1 AMPs act as immunostimulants

The use of alternative feed ingredients may not only affect the fish welfare, health, nutrient utilization, growth, and fillet quality, but also intestinal microbiota. As nature provides a variety of bioactive peptides such as AMPs from bacteria, algae, insects, plants, crustaceans, crabs, mollusks, and fish as well as milk and milk products, researchers have been exploiting their potential as novel food preservatives and food additives (Madrazo and Campos, 2020). Commercialization of AMPs has experienced problems during the infant stages of clinical design, clinical trials, funding sources and regulatory hurdles (Fox, 2013).

In aquaculture, considerable attention has been paid to the use of immunostimulants to enhance activities of the immune system to improve the disease resistance of farmed fish (Sakai, 1999). As potential immunostimulants, external AMPs administered as feed additives to regulate non-specific immunity have been proved to play a vital role in the prevention of fish diseases (Sakai, 1999; Kumari et al., 2003; Abdel-Wahab et al., 2021).

In the previous literature, the most widely documented AMP is bovine lactoferrin (BLF), which has been used in 11 aquatic species, including two species of catfish, Asian catfish (*Clarias batrachus*) and channel catfish (*I. punctatus*) (Abdel-Wahab et al., 2021). The non-specific immunity and disease resistance against *A. hydrophila* of Asian catfish can be gained when the BLF was added to a regular diet at 100 mg/kg feeding for 1 week (Kumari et al., 2003). Welker et al. (Welker et al., 2010) established a minimal concentration of 1,136 mg/kg BLF in feed against *E. ictaluri* infection for juvenile channel catfish, and this level is in agreement with Kakuta (1996) who reported that dietary lactoferrin supplementation up to 1,000 mg/kg improved growth performance in goldfish (*Carassius auratus*) after 4 weeks. However, the amount and period of oral additives are dependent on the species and individual size of the fish,

and there is no uniform formulation at present. For instance, 400 mg/kg of dietary BLF can significantly improve survival rate of yellowfin sea bream (*Acanthopagrus latus*) upon challenge with *Vibrio harveyi*, and this level was optimal regarding the capability of killing bacteria compared to 800 and 1,200 mg/kg (Abdel-Wahab et al., 2021). Concentrations fluctuate for different fish species and various pathogens, indicating that the BLF as feed additive has species specificity.

As a dietary supplementation, the cecropin from *Hyalophora cecropia* (accession number: AAP93872.1) was added to common carp (*Cyprinus carpio*) basal diets at different concentrations to improve immunity by increasing expression of cytokines and IgM (Dong et al., 2015). Lin et al. (2015) reported that cecropin increased the disease resistance against *A. veronii* in tilapia when the diet was supplemented with 150 mg/kg cecropin. In addition to BLF and cecropin, a recombinant hepcidin peptide was added to grass carp feed formula, and protection against *F. columnare* was achieved via regulation of the iron distribution and immune gene expression (Wei et al., 2018). Pridgeon et al. (2013ab) developed two recombinant plasmids containing chicken-type (pcDNA-Lys-c) and goose-type (pcDNA-Lys-g) lysozyme, which can be used as novel immunostimulants to protect channel catfish against *A. hydrophila* infection. For these plasmid DNA immunostimulants, Pridgeon's team delivered these plasmid DNA immunostimulants to the fish via intraperitoneal injection and the challenge experiment revealed that these novel immunostimulants offered 100% protection against *A. hydrophila* at 2 days post injection (Pridgeon et al., 2013ab).

5.2 AMGs act as transgenes

Genome editing tools have revolutionized the ability to accelerate the pace of aquaculture breeding, bringing edible products, including growth hormone gene transgenic Atlantic salmon, AquAdvantage salmon (Ledford, 2015; Waltz, 2017), leptin receptor gene edited tiger puffer and myostatin-disrupted red sea bream (https://doi.org/10.1038/s41587-021-01197-8) as commercial applications. In addition, gene edited tilapia has been exempted from genetically modified (GM) regulation in Argentina (https://thefishsite.com/articles/gene-edited-tilapia-secures-gmo-exemption). This initial commercialization may lead to increased momentum for GM fish to reach the table in the future. As food additives, AMPs can effectively improve the disease resistance of aquatic animals in a short term, and their integrity and stability depend on the

concentrations and physiological status of fish (Kakuta, 1996; Welker et al., 2010; Abdel-Wahab et al., 2021). However, from a genetic perspective, if germline of fish of interest harboring the external AMG that confers the host more antimicrobial activities and the resistance is multi-generational, this would momentously increase the output of any aquacultured fish production, relatively lower the mortality and cost, and increase the profit. Currently, the introduction of exogenous AMGs via genetic engineering and CRISPR/Cas9 has attracted the attention of researchers.

5.2.1 Transgenes

The AMGs introduced into aquatic animals via transgenic technology to improve disease resistance primarily focus upon cecropin (Dunham et al., 2002; Sarmasik et al., 2002; Chiou et al., 2014), hepcidin (Hsieh et al., 2010), epinecidin (Lee et al., 2013), piscidin (Lin et al., 2016), lysozyme (Yazawa et al., 2006) and lactoferrin (Zhong et al., 2002; Mao et al., 2004; Lin et al., 2010) and these AMG transgenes have been studied in nine species of fish, including two model fish, zebrafish (*Danio rerio*) and medaka (*Oryzias latipas*) (Sarmasik et al., 2002; Yazawa et al., 2006; Lin et al., 2010) (Figure 3).

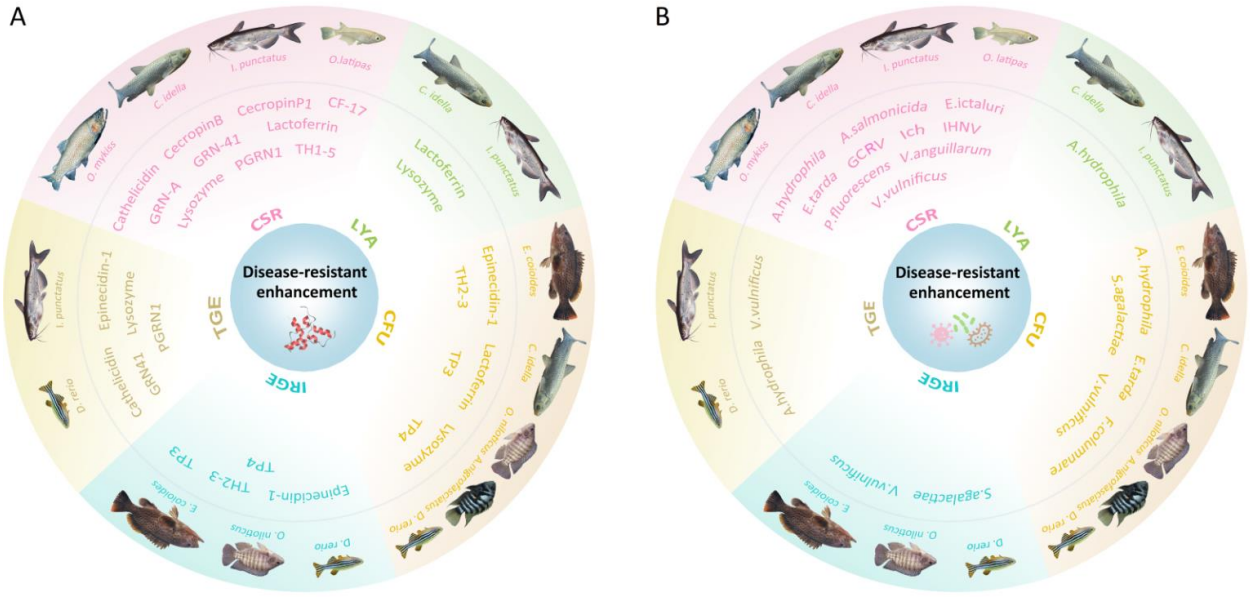


Figure 3. A summary of the current applications for enhancement of disease resistance in aquaculture using antimicrobial peptide genes (AMGs). These applications included 23 studies, 8 fish species, 15 AMGs, and 12 diseases. **(A)** Different AMGs were applied for each parameter (CFU, LYA,

CSR, TGE and IRGE). For example, two AMGs (lactoferrin and lysozyme) were used as transgenes in grass carp (*C. idella*) and channel catfish (*I. punctatus*) for evaluation of LYA. **(B)** A variety of pathogens were involved for each parameter. For example, one bacterial species (*A. hydrophila*) was used as a pathogenic infection in grass carp and channel catfish for evaluation of LYA. Ich, Ichthyophthirius multifiliis; GCRV, grass carp reovirus; IHNV, infectious hematopoietic necrosis virus; TP3, tilapia piscidin 3; TP4, tilapia piscidin 4; TH2-3, tilapia hepcidin 2-3; TH1-5, tilapia hepcidin 1-5; PGRN1, a type of progranulin gene from Mozambique tilapia; GRN-41/GRN-A, AMGs from Mozambique tilapia to produce secreted GRN peptides; CF-17, a synthetic cecropin B analog; IRGE, the expression of immune-related genes; TGE, the expression of exogenous AMGs; CSR, cumulative survival rate; LYA, lysozyme activity; CFU, colony-forming unit of bacteria.

The cecropin B gene, first discovered from the moth (*Hyalophora cecropia*), was the first exogenous AMG to be integrated into the genome of fish. The disease resistance against *F. columnare* and *E. ictaluri* of F_1 channel catfish possessing cecropin gene was increased by 27.3% and 14.8%, respectively (Dunham et al., 2002), and cecropin-transgenic channel catfish conferred the similar resistance to Ich in hybrid catfish (channel catfish female × blue catfish male) (Elaswad et al., 2019). Sarmasik et al. (2002) applied the same construct into medaka, the result was in agreement with Dunham et al. (2002), as the F_2 progeny of cecropin-transgenic medaka acquired enhanced resistance to *Pseudomonas fluorescens* and *V. anguillarum* infection. In addition, two cecropin constructs, porcine cecropin and one synthetic cecropin analog, were used to generate transgenic rainbow trout, exhibiting resistance against both bacteria (*A. salmonicida*) and virus (infectious hematopoietic necrosis virus; IHNV) (Chiou et al., 2014). These results clearly indicated that cecropins as antimicrobial peptide genes were powerful candidate genes for improving disease resistance of aquatic animals for aquaculture.

In addition to being used as immunomodulators or additives, lactoferrin and lysozyme were also integrated into fish to improve disease resistance. As a transgene, the human lactoferrin (hLF) gene has been widely introduced into plants, such as tobacco (Mitra and Zhang, 1994). For fish, hLF-transgenic grass carp exhibited enhanced resistance against *A. hydrophila* infection by improving the phagocytic activity (Mao et al., 2004), and it also has been proved to be an antiviral against GCH virus (Zhong et al., 2002). In addition to hLF, bovine lactoferrin (LFB) was also studied in fish, and a disease-resistant line of zebrafish carrying LFB was produced which was more resistant to *E. coli*, *E. tarda* and *A. hydrophila* (Lin et al., 2010). Lysozyme is present in a wide variety of organisms, including aquatic animals. In order to verify that the foreign

lysozyme has a higher bactericidal effect than the innate one, the hen egg white (HEW) lysozyme gene was introduced into zebrafish, and the results showed that the disease resistance was greatly improved in F_2 transgenic individuals (Yazawa et al., 2006).

In addition to lysozyme, piscidin and epinecidin are also fish-derived AMPs. Interestingly, when the fish muscle was injected with endogenous piscidin or epinecidin-1 via electroporation, it can greatly increase the overexpression of immune-related genes in the body, thereby improving disease resistance (Lee et al., 2013; Lin et al., 2016). Although the construct containing AMGs electroporated intramuscularly can remain in the body for up to one year (Mir et al., 1999), their heritability was not tested. In addition to these AMGs applied in antimicrobial studies, a potent antiviral gene, Mx, has also been used in antiviral studies in fish. Mx protein possesses direct antiviral activity and inhibits multiple viruses by blocking the viral genome replication cycle (Haller et al., 2007), which has been proved in mice and human (Pavlovic et al., 1990; Arnheiter et al., 2007). The first attempt to apply the Mx gene to aquatic animals was to conquer the grass carp reovirus (GCRV) in a rare minnow (*Gobiocypris rarus*), and the offspring of genetically modified fish have greatly improved anti-virus ability (Su et al., 2009).

Obviously, AMGs can be transferred from one species of fish to another, or from other organisms, and can enhance disease resistance by expressing corresponding proteins or promoting innate immune responses in the body. Compared to food additives or immunomodulators, the transgenic approach has the advantages of allowing the fish to always be ready to respond to disease, the resistance is multi-generational, and once naturally incorporated, no additional cost is incurred. All approaches mentioned would be considered an application of a drug by the US Food and Drug Administration, and all have this disadvantage of having to clear the approval process for commercial application.

5.2.2 Genome editing

To achieve site-specific repair or gene modification to enhance several traits simultaneously, genome editing technology, including ZFN (Kim et al., 1996; Doyon et al., 2008), TALEN (Miller et al., 2011) and CRISPR/Cas9 (Jinek et al., 2012; Cong et al., 2013), has developed rapidly and applied in aquatic animal breeding to improve commercially relevant traits, including reproduction and fertility (Qin et al., 2016), fatty acid composition (Cheng et al., 2014; Huang et

al., 2021), growth rate (Zhong et al., 2016; Khalil et al., 2017; Zhang et al., 2020) and disease resistance (Elaswad et al., 2018; Simora et al., 2020a).

As the principal family of AMPs, cathelicidins play a pivotal role in immune defense against microbial invasions, and are widely present in various organisms, including humans, plants and aquatic animals, but they are absent in catfish (Maier et al., 2008). Recent studies have shown that alligators are rich in cathelicidins, and these peptides of the crocodilian are more powerful than those from other species (Chen et al., 2017). Moreover, compared to moth cecropin B, sea snake cathelicidin, flounder pleurocidin and ampicillin, alligator cathelicidin shows highest effectiveness against bacterial pathogens (Simora et al., 2021). In order to transfer this powerful disease-resistant AMP into non-cathelicidin species, some pioneer work has been done via CRISPR/Ca9 technique. Simora et al. (2020b) investigated the bactericidal ability of the cathelicidin gene derived from American alligator (*Alligator mississippiensis*) and compared it with cecropin and pleurocidin. The results showed that alligator cathelicidin had strong bactericidal effect against *E. ictalurid* infection. Subsequently, it was integrated into the non-coding region of channel catfish genome via CRISPR/Cas9, these cathelicidin-transgenic fish possessed potential anti-viral and antimicrobial activities (Simora et al., 2020a). Additionally, cathelicidin gene was introduced into the genomes of another two catfish species, blue catfish (*I. furcatus*) and white catfish (*Ameiurus catus*) at *lh* locus (Wang et al. unpublished data).

Unlike inserting foreign AMGs, knocking out immune-related genes to regulate gene expression or disturb pathways can also improve disease resistance. A few other genes, TICAM1/RBL in channel catfish (our lab), TLR22 in rohu carp (*Labeo rohita*), stat2 in a Chinook salmon embryo cell line (*O. tshawytscha*), JAM-A in a grass carp cell line and PoMaf1 in olive flounder (*P. olivaceus*) cell line, summarized in Gratacap et al. (Gratacap et al., 2019) and Blix et al. (Blix et al., 2021) have been mutated, modifying the immunity of the fish and increasing the host's disease resistance. Using CRISPR/Cas9 tool to edit TLR22 in *L. rohita* carp abolished its mRNA expression, but phenotypic traits evaluation on the mutant genotypes had not been performed (Chakrapani et al., 2016). Expression of TICAM1 gene in channel catfish was dramatically upregulated against with *E. ictalurid* infection (Baoprasertkul et al., 2006). Based on this assumption, Elaswad et al. (Elaswad et al., 2018) attempted to enhance the disease resistance by knocking out TICAM1 and RBL. Mutant catfish of P_1 and F_1 generation have been produced but

not thoroughly evaluated. This strategy has been successfully applied against grass carp reovirus (GCRV). As an immunoglobulin member, JAM-A serves as a receptor for reovirus (Antar et al., 2009), and JAM-A expression can support reovirus infection (Barton et al., 2001). In this way, Ma et al. (Ma et al., 2018) exploited CRISPR/Cas9 knockout grass carp JAM-A gene, inducing disease resistance against reovirus infection in cell lines. In a second case of Chinook salmon cell line, knockout of stat2 increased viral resistance as seen the increased production of viral particles of the DNA virus, epizootic hematopoietic necrosis virus (EHNV), and the RNA virus, salmon pancreatic disease virus (SPDV), but less ssRNA viral particles, viral hemorrhagic septicemia virus (VHSV) (Dehler et al., 2019). In a third hirame natural embryo cell line of Japanese flounder, knockout PoMaf1 (an RNA polymerase (pol) III-associated transcription repressor) using CRISPR/Cas9 increased the VHSV glycoprotein (G) mRNA levels during viral hemorrhagic septicemia virus infection (Kim et al., 2020). Although CRISPR/cas9 technology has been successfully applied in the research of disease-resistant enhancement of aquatic animals, further pathogen infection experiments *in vivo* are needed to verify its value regarding knock out of AMGs and other immune-associated/responsive genes that could potentially increase disease resistance for specialized cases.

The rapid development of biotechnology, especially CRISPR-Cas 9 technique opens the possibility for novel application approaches of AMGs. In this vein, we focused on reviewing the application of AMPs/AMGs to improve disease resistance in aquaculture through genetic engineering and genome editing. Although an increasing number of fish AMPs/AMGs have been isolated and identified, only a small portion has been studied for potential applications because of several insurmountable issues. In general, whether testing the physicochemical properties or bactericidal potency of a novel AMP or applying it to actual production, the need for sufficient quantity of AMPs for adequate testing is the first step to achieve these goals. There are three main ways to accomplish this need, direct isolation and purification from natural products, chemical synthesis, and biotechnological production. However, the source of natural AMPs is limited, and the chemical synthesis has the problems of high cost and difficulty in batch production. Fortunately, based on molecular biology, especially genetic engineering technology, it is a feasible way to generate purified peptides harnessing the power of plant-based production systems. For instance, plant production of recombinant fish cytokine interleukin-22 (IL-22) has been achieved in channel catfish, and the plant-made IL-22 is of high quality at sufficient yield

(5.4 mg/kg fresh tobacco leaf) (Elkins and Dolan, 2021). Except for plant-based production systems, bacterial and yeast expression systems are also applied, such as grass carp CXCL20a (Zhang et al., 2021; Wang et al., 2022). On the one hand, this provides the possibility to produce a large amount of high-purity AMPs at low cost and solves the shortage of some specific AMPs. On the other hand, once genetically modified concepts are accepted by the public and the market, highly disease-resistant aquatic animals carrying AMGs created by genetic engineering or genome editing could have many practical and economic advantages.

Although, public acceptance is an issue, partially because of problems with public education, transparency in regulatory approaches and governance, establishment of trust and other factors (Dehler et al., 2019; Dunham, 2022), consumers in China, the US and around the world are likely to be increasingly accepting to genetically engineered (GE) food when it lowers food costs or increases nutritional quality (Bai et al., 2004; Curtis et al., 2004; Hallerman et al., 2022). The public is becoming increasingly aware of and concerned about animal welfare. Insertion of genes that enhance health and well-being of aquatic organisms will likely be among the first strategies to gain public approval.

6. Objectives

The long-term goal is to genetically improve overall catfish performance, including disease resistance, revisable sterility and growth rate in both channel catfish, *Ictalurus punctatus*, and blue catfish, *I. furcatus*. Our supportive/specific objectives are 1) Replace the luteinizing hormone (*lh*) gene with the cathelicidin transgene from the alligator (*Alligator sinensis*, *As-Cath*) followed by hormone therapy to gain total reproductive control and confinement of disease resistant transgenic channel catfish. 2) Integrate the *As-Cath* transgene into the target *lh* locus of blue catfish using two CRISPR/Cas9-mediated delivery systems assisted by linear double-stranded DNA (dsDNA) and double-cut plasmid (dcPlasmid), respectively. 3) Generate dual-gene integrated genetic lines using transgenic channel catfish founders coupled with a CRISPR/Cas9-mediated system for the introduction of a second gene to improve disease resistance. 4) Establish the linear dsDNA and double-cut plasmid (dcPlasmid) combination-assisted multiplex genome editing in channel catfish, allowing combinational deletion mutagenesis and transgene knock-in at multiple sites via the NHEJ/HDR pathway in parallel.

7. References

Abdel-Wahab MM, Taha NM, Lebda MA, Elfeky MS, Abdel-Latif HMR. Effects of bovine lactoferrin and chitosan nanoparticles on serum biochemical indices, antioxidative enzymes, transcriptomic responses, and resistance of Nile tilapia against *Aeromonas hydrophila*. *Fish Shellfish Immunol*. 2021;111:160-169. doi:10.1016/j.fsi.2021.01.017

Antar AAR, Konopka JL, Campbell JA, et al. Junctional adhesion molecule-A is required for hematogenous dissemination of reovirus. *Cell Host Microbe*. 2009;5(1):59-71. doi:10.1016/j.chom.2008.12.001

Arnheiter H, Skuntz S, Noteborn M, Chang S, Meier E. Transgenic mice with intracellular immunity to influenza-virus. *Cell*. 1990; 62(1):51-61. doi:10.1016/0092-8674(90)90239-B

Baggiolini M. Chemokines and leukocyte traffic. *Nature*. 1998;392:565-568. doi:10.1038/33340

Bai J, Zhang C, Huang J, Hallman W, Pray CE, Aquino HL. Consumer acceptance of genetically modified foods: A comparison between the US and China. *American Agricultural Economics Association*. Annual Meeting, Denver CO, 2004;1-4. doi: 10.22004/ag.econ.20026

Bao B, Peatman E, Li P, He C, Liu Z. Catfish hepcidin gene is expressed in a wide range of tissues and exhibits tissue-specific upregulation after bacterial infection. *Dev Comp Immunol*. 2005;29(11):939-950. doi:10.1016/j.dci.2005.03.006

Bao B, Peatman E, Xu P, et al. The catfish liver-expressed antimicrobial peptide 2 (LEAP-2) gene is expressed in a wide range of tissues and developmentally regulated. *Mol Immunol*. 2006; 43(4):367-377. doi:10.1016/j.molimm.2005.02.014

Baoprasertkul P, Peatman E, Somridhivej B, Liu Z. Toll-like receptor 3 and TICAM genes in catfish: species-specific expression profiles following infection with *Edwardsiella ictaluri*. *Immunogenetics*. 2006;58:817-830. doi:10.1007/s00251-006-0144-z

Barton ES, Forrest JC, Connolly JL, et al. Junction adhesion molecule is a receptor for reovirus. *Cell*. 2001;104(3):441-451. doi:10.1016/S0092-8674(01)00231-8

Birkemo GA, Luders T, Andersen Ø, Nes IF, Nissen-Meyer J. Hipposin, a histone-derived antimicrobial peptide in Atlantic halibut (*Hippoglossus hippoglossus* L.). *BBA-Proteins Proteomics*. 2003;1646(1-2):207-215. doi:10.1016/S1570-9639(03)00018-9

Bleackley MR, Hayes BM, Parisi K, et al. Bovine pancreatic trypsin inhibitor is a new antifungal peptide that inhibits cellular magnesium uptake. *Mol Microbiol*. 2014;92(6):1188-1197. doi:10.1111/mmi.12621

Blix TB, Dalmo RA, Wargelius A, Myhr AI. Genome editing on finfish: Current status and implications for sustainability. *Rev Aquac*. 2021;3:2344-2363. doi:org/10.1111/raq.12571

Brunner SR, Varga JFA, Dixon B. Antimicrobial peptides of salmonid fish: from form to function. *Biology-Basel*. 2020;9(8):233. doi:10.3390/biology9080233

Cabello FC. Heavy use of prophylactic antibiotics in aquaculture: a growing problem for human and animal health and for the environment. *Environ Microbiol.* 2006;8:1137-44. doi:10.1111/j.1462-2920.2006.01054.x.

Chakrapani V, Patra SK, Panda RP, Rasal KD, Jayasankar P, Barman HK. Establishing targeted carp TLR22 gene disruption via homologous recombination using CRISPR/Cas9. *Dev Comp Immunol*. 2016;61:242-247. doi:10.1016/j.dci.2016.04.009

Chen Y, Cai S, Qiao X, et al. As-CATH1-6, novel cathelicidins with potent antimicrobial and immunomodulatory properties from *Alligator sinensis*, play pivotal roles in host antimicrobial immune responses. *Biochem J*. 2017;474(16):2861-2885. doi:10.1042/BCJ20170334

Cheng Q, Su B, Qin Z, et al. Interaction of diet and the masou salmon Delta 5-desaturase transgene on Delta 6-desaturase and stearoyl-CoA desaturase gene expression and N-3 fatty acid level in common carp (*Cyprinus carpio*). *Transgenic Res*. 2014;23:729-742. doi:10.1007/s11248-014-9812-1

Cheung RCF, Ng TB, Wong JH. Marine peptides: Bioactivities and applications. *Mar Drugs*. 2015;13(7):4006-4043. doi:10.3390/md13074006

Chia T-J, Wu Y-C, Chen J-Y, Chi S-C. Antimicrobial peptides (AMP) with antiviral activity against fish nodavirus. *Fish Shellfish Immunol*. 2010;28(3):434-439. doi:10.1016/j.fsi.2009.11.020

Chinchar VG, Bryan L, Silphadaung U, Noga E, Wade D, Rollins-Smith L. Inactivation of viruses infecting ectothermic animals by amphibian and piscine antimicrobial peptides. *Virology*. 2004;323(2):268-275. doi:10.1016/j.virol.2004.02.029

Chiou PP, Chen MJ, Lin C-M, et al. Production of homozygous transgenic rainbow trout with enhanced disease resistance. *Mar Biotechnol*. 2014;16:299-308. doi:10.1007/s10126-013-9550-z

Cho JH, Park IY, Kim HS, Lee WT, Kim MS, Kim SC. Cathepsin D produces antimicrobial peptide parasin I from histone H2A in the skin mucosa of fish. *Faseb J*. 2002;16(3):429-431. doi:10.1096/fj.01-0736fje

Choi H, Hwang J-S, Lee DG. Identification of a novel antimicrobial peptide, scolopendin 1, derived from centipede *Scolopendra subspinipes mutilans* and its antifungal mechanism. *Insect Mol Biol*. 2014;23(6):788-799. doi:10.1111/imb.12124

Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339(6121):819-823. doi:10.1126/science.1231143

Curtis KR, McCluskey JJ, Wahl TI. Consumer acceptance of genetically modified food products in the developing world. *AgBioForum*. 2004;7(1&2):70-75. http://hdl.handle.net/10355/177

Dehler CE, Lester K, Della Pelle G, et al. Viral resistance and IFN signaling in STAT2 knockout fish cells. *J Immunol*. 2019;203(2):465-475. doi:10.4049/jimmunol.1801376

Dong X, Ye Z, Song L, et al. Expression profile analysis of two cathepsin S in channel catfish (*Ictalurus punctatus*) mucosal tissues following bacterial challenge. *Fish Shellfish Immunol*. 2016;48:112-118. doi:10.1016/j.fsi.2015.11.030

Dong X-Q, Zhang D-M, Chen Y-K, Wang Q-J, Wang Y-Y. Effects of antimicrobial peptides (AMPs) on blood biochemical parameters, antioxidase activity, and immune function in the common carp (*Cyprinus carpio*). *Fish Shellfish Immunol*. 2015;47(1):429-434. doi:10.1016/j.fsi.2015.09.030

Dong Y-H, Wang L-H, Zhang L-H. Quorum-quenching microbial infections: mechanisms and implications. *Philos Trans R Soc B-Biol Sci*. 2007;362:1201-11. doi:10.1098/rstb.2007.2045.

Doyon Y, McCammon JM, Miller JC, et al. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat Biotechnol.* 2008;26:702-708. doi:10.1038/nbt1409

Dunham RA, Elaswad A. Catfish Biology and Farming. In: Lewin HA, Roberts RM, editors. *Annual Review of Animal Biosciences*, vol 6, Palo Alto: Annual Reviews; 2018, p. 305–25. doi:10.1146/annurev-animal-030117-014646.

Dunham RA, Warr GW, Nichols A, et al. Enhanced bacterial disease resistance of transgenic channel catfish *Ictalurus punctatus* possessing cecropin genes. *Mar Biotechnol*. 2002;4:338-344. doi:10.1007/s10126-002-0024-y

Dunham RA. Aquaculture and Fisheries Biotechnology: Genetic Approaches. 3rd Edition. CABI, Wallingford, UK, 2022. In press.

Elaswad A, Khalil K, Ye Z, et al. Effects of cecropin transgenesis and interspecific hybridization on the resistance to *Ichthyophthirius multifiliis* in channel catfish and female channel catfish × male blue catfish hybrids. *N Am J Aqualcult*. 2019;81(3):242-252. doi:10.1002/naaq.10096

Elaswad A, Khalil K, Ye Z, et al. Effects of CRISPR/Cas9 dosage on TICAM1 and RBL gene mutation rate, embryonic development, hatchability and fry survival in channel catfish. *Sci Rep*. 2018;8:16499. doi:10.1038/s41598-018-34738-4

Elkins LL, Dolan MC. Plant production and functional characterization of catfish interleukin-22 as a natural immune stimulant for aquaculture fish. *J Biotechnol*. 2021;325:233-240. doi:10.1016/j.jbiotec.2020.10.017

Falco A, Chico V, Marroqui L, Perez L, Coll JM, Estepa A. Expression and antiviral activity of a beta-defensin-like peptide identified in the rainbow trout (*Oncorhynchus mykiss*) EST sequences. *Mol Immunol*. 2008;45(3):757-765. doi:10.1016/j.molimm.2007.06.358

FAO. The FAO action plan on antimicrobial resistance 2021–2025. Rome; 2021, p. 1–46. doi:10.4060/cb5545en.

FAO. The state of world fisheries and aquaculture 2020. Sustainability in action. Rome; 2020, p. 1-244. doi:10.4060/ca9229en.

Fernandes JMO, Kemp GD, Molle MG, Smith VJ. Anti-microbial properties of histone H2A from skin secretions of rainbow trout, *Oncorhynchus mykiss*. *Biochem J*. 2002;368(2):611-620. doi:10.1042/BJ20020980

Fox JL. Antimicrobial peptides stage a comeback. *Nat Biotechnol*. 2013;31(5):379-382. doi:10.1038/nbt.2572

Fritz H, Wunderer G. Biochemistry and applications of aprotinin, the kallikrein inhibitor from bovine organs. *Arzneimittelforschung-Drug Res*. 1983;33(4):479-494

Fu Q, Yang Y, Li C, et al. The chemokinome superfamily: II. The 64 CC chemokines in channel catfish and their involvement in disease and hypoxia responses. *Dev Comp Immunol*. 2017a;73:97-108. doi:10.1016/j.dci.2017.03.012

Fu Q, Zeng Q, Li Y, et al. The chemokinome superfamily in channel catfish: I. CXC subfamily and their involvement in disease defense and hypoxia responses. *Fish Shellfish Immunol*. 2017b;60:380-390. doi:10.1016/j.fsi.2016.12.004

García-Valtanen P, Martinez-Lopez A, Ortega-Villaizan M, Perez L, Coll JM, Estepa A. In addition to its antiviral and immunomodulatory properties, the zebrafish beta-defensin 2 (zfBD2) is a potent viral DNA vaccine molecular adjuvant. *Antiviral Res*. 2014;101:136-147. doi.o:10.1016/j.antiviral.2013.11.009

Gela A, Kasetty G, Jovic S, et al. Eotaxin-3 (CCL26) exerts innate host defense activities that are modulated by mast cell proteases. *Allergy*. 2015;70(2):161-170. doi:10.1111/all.12542

Geng X, Feng J, Liu S, Wang Y, Arias C, Liu Z. Transcriptional regulation of hypoxia inducible factors alpha (HIF-alpha) and their inhibiting factor (FIH-1) of channel catfish (*Ictalurus punctatus*) under hypoxia. *Comp Biochem Physiol B-Biochem Mol Biol*. 2014;169:38-50. doi:10.1016/j.cbpb.2013.12.007

Gratacap RL, Wargelius A, Edvardsen RB, Houston RD. Potential of genome editing to improve aquaculture breeding and production. *Trends Genet*. 2019;35(9):672-684. doi:10.1016/j.tig.2019.06.006

Gui L, Zhang P, Zhang Q, Zhang J. Two hepcidins from spotted scat (*Scatophagus argus*) possess antibacterial and antiviral functions *in vitro*. *Fish Shellfish Immunol*. 2016;50:191-199. doi:10.1016/j.fsi.2016.01.038

Guo M, Wei J, Huang X, Huang Y, Qin Q. Antiviral effects of beta-defensin derived from orange-spotted grouper (*Epinephelus coioides*). *Fish Shellfish Immunol*. 2012;32(5):828-838. doi:10.1016/j.fsi.2012.02.005

Haller O, Staeheli P, Kochs G. Interferon-induced Mx proteins in antiviral host defense. *Biochimie*. 2007;89(6-7):812-818. doi:10.1016/j.biochi.2007.04.015

Hallerman EM, Bredlau JP, Camargo LSA, et al. Towards progressive regulatory approaches for agricultural applications of animal biotechnology. *Transgenic Res*. 2022. doi:10.1007/s11248-021-00294-3.

Hancock REW, Scott MG. The role of antimicrobial peptides in animal defenses. *Proc Natl Acad Sci USA*. 2000;97(16):8856-8861. doi:10.1073/pnas.97.16.8856

Hancock REW. Cationic peptides: effectors in innate immunity and novel antimicrobial peptide. *Lancet Infect Dis*. 2001;1:156-64. doi:10.1016/S1473-3099(01)00092-5.

Hansen FC, Kalle-Brune M, van der Plas MJA, et al. The thrombin-derived host defense peptide GKY25 inhibits endotoxin-induced responses through interactions with lipopolysaccharide and macrophages/monocytes. *J Immunol*. 2015;194(11):5397-5406. doi:10.4049/jimmunol.1403009

Heuer OE, Kruse H, Grave K, Collignon P, Karunasagar I, Angulo FJ. Human health consequences of use of antimicrobial agents in aquaculture. *Clin Infect Dis*. 2009;49(8):1248-1253. doi:10.1086/605667

Hicke L. Protein regulation by monoubiquitin. *Nat Rev Mol Cell Biol*. 2001;2:195-201. doi:10.1038/35056583

Hikima J-I, Minagawa S, Hirono I, Aoki T. Molecular cloning, expression and evolution of the Japanese flounder goose-type lysozyme gene, and the lytic activity of its recombinant protein. *Biochim Biophys Acta-Gene Struct Expression*. 2001;1520(1):35-44. doi:10.1016/S0167-4781(01)00248-2

Hirsch J. Bactericidal action of histone. *J Exp Med*. 1958;108(6):925-944. doi:10.1084/jem.108.6.925

Hsieh J-C, Pan C-Y, Chen J-Y. Tilapia hepcidin (TH)2-3 as a transgene in transgenic fish enhances resistance to *Vibrio vulnificus* infection and causes variations in immune-related genes after infection by different bacterial species. *Fish Shellfish Immunol*. 2010;29(3):430-439. doi:10.1016/j.fsi.2010.05.001

Huang R, Zhang J, Zhu G, He J, Xie J. The core ubiquitin system of mandarin fish, *Siniperca chuatsi*, can be utilized by infectious spleen and kidney necrosis virus. *Fish Shellfish Immunol*. 2017;70:293-301. doi:10.1016/j.fsi.2017.09.017

Huang Y, Bugg W, Bangs M, et al. Direct and pleiotropic effects of the masou salmon Delta-5 desaturase transgene in F1 channel catfish (*Ictalurus punctatus*). *Transgenic Res*. 2021;30:185-200. doi:10.1007/s11248-021-00242-1

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable Dual-RNA-Guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337(6096):816-821. doi:10.1126/science.1225829

Kakuta I. Effects of orally administrated bovine lactoferrin on growth and blood properties of goldfish. *Suisan Zoshoku*. 1996;44(4): 419-426. doi:10.11233/aquaculturesci1953.44.419

Kalle M, Papareddy P, Kasetty G, et al. Host defense peptides of thrombin modulate inflammation and coagulation in endotoxin-mediated shock and *Pseudomonas aeruginosa* sepsis. *PLoS One*. 2012;7:e51313. doi:10.1371/journal.pone.0051313

Karunasagar ID, Karunasagar IN, Bondad-Reantaso MG. Complexities involved in source attribution of antimicrobial resistance genes found in aquaculture products. *Asian Fish Sci*. 2020;33:16-21. doi:10.33997/j.afs.2020.33.S1.003.

Katzenback BA. Antimicrobial peptides as mediators of innate immunity in teleosts. *Biology-Basel*. 2015;4(4):607-639. doi:10.3390/biology4040607

Khalil K, Elayat M, Khalifa E, et al. Generation of myostatin gene-edited channel catfish (*Ictalurus punctatus*) via zygote injection of CRISPR/Cas9 system. *Sci Rep*. 2017;7:7301. doi:10.1038/s41598-017-07223-7

Kieffer A-E, Goumon Y, Ruh O, et al. The N- and C-terminal fragments of ubiquitin are important for the antimicrobial activities. *Faseb J*. 2003;17(6):776-778. doi:10.1096/fj.02-0699fje

Kim J, Cho JY, Kim J-W, et al. Molecular characterization of *Paralichthys olivaceus* MAF1 and its potential role as an anti-viral hemorrhagic septicaemia virus factor in hirame natural embryo cells. *Int J Mol Sci*. 2020;22(3):1353. doi:10.3390/ijms22031353

Kim J-Y, Lee SY, Park S-C, et al. Purification and antimicrobial activity studies of the N-terminal fragment of ubiquitin from human amniotic fluid. *BBA-Proteins Proteomics*. 2007;1774(9):1221-1226. doi:10.1016/j.bbapap.2007.06.013

Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci USA*. 1996;93(3):1156-1160. doi:10.1073/pnas.93.3.1156

Ko J, Wan Q, Bathige SDNK, Lee J. Molecular characterization, transcriptional profiling, and antibacterial potential of G-type lysozyme from seahorse (*Hippocampus abdominalis*). *Fish Shellfish Immunol*. 2016;58:622-630. doi:10.1016/j.fsi.2016.10.014

Kumari J, Swain T, Sahoo PK. Dietary bovine lactoferrin induces changes in immunity level and disease resistance in Asian catfish *Clarias batrachus*. *Vet Immunol Immunopathol*. 2003;94(1-2):1-9. doi:10.1016/S0165-2427(03)00065-5

Ledford H. Salmon approval heralds rethink of transgenic animals. *Nature*. 2015;527:417-418. doi:10.1038/527417a

Lee H, Hwang JS, Lee DG. Scolopendin, an antimicrobial peptide from centipede, attenuates mitochondrial functions and triggers apoptosis in *Candida albicans*. *Biochem J*. 2017;474(5):635-645. doi:10.1042/BCJ20161039

Lee H, Hwang J-S, Lee J, Kim J, Lee DG. Scolopendin 2, a cationic antimicrobial peptide from centipede, and its membrane-active mechanism. *Biochim Biophys Acta-Biomembr*. 2015;1848(2):634-642. doi:10.1016/j.bbamem.2014.11.016

Lee L-H, Hui C-F, Chuang C-M, Chen J-Y. Electrotransfer of the epinecidin-1 gene into skeletal muscle enhances the antibacterial and immunomodulatory functions of a marine fish, grouper (*Epinephelus coioides*). *Fish Shellfish Immunol*. 2013;35(5):1359-1368. doi:10.1016/j.fsi.2013.07.050

Li C, Song L, Zhao J, et al. Preliminary study on a potential antibacterial peptide derived from histone H2A in hemocytes of scallop *Chlamys farreri*. *Fish Shellfish Immunol*. 2007;22(6):663-672. doi:10.1016/j.fsi.2006.08.013

Li C, Zhang Y, Wang R, et al. RNA-seq analysis of mucosal immune responses reveals signatures of intestinal barrier disruption and pathogen entry following *Edwardsiella ictaluri* infection in channel catfish, *Ictalurus punctatus*. *Fish Shellfish Immunol*. 2012;32(5):816-827. doi:10.1016/j.fsi.2012.02.004

Lin C-Y, Yang P-H, Kao C-L, Huang H-I, Tsai H-J. Transgenic zebrafish eggs containing bactericidal peptide is a novel food supplement enhancing resistance to pathogenic infection of fish. *Fish Shellfish Immunol*. 2010;28(3):419-427. doi:10.1016/j.fsi.2009.11.019

Lin W-C, Chang H-Y, Chen J-Y. Electrotransfer of the tilapia piscidin 3 and tilapia piscidin 4 genes into skeletal muscle enhances the antibacterial and immunomodulatory functions of *Oreochromis niloticus*. *Fish Shellfish Immunol*. 2016;50:200-209. doi:10.1016/j.fsi.2016.01.034

Lin X, Chen W, Lin S, Luo L. Effects of dietary cecropin on growth, non-specific immunity and disease resistance of tilapia (*Oreochromis niloticus* × *O. aureus*). *Aquac Res*. 2015;46:2999-3007. doi:10.1111/are.12457

Liu S, Wang X, Sun F, et al. RNA-Seq reveals expression signatures of genes involved in oxygen transport, protein synthesis, folding, and degradation in response to heat stress in catfish. *Physiol Genomics*. 2013;45:462-476. doi:10.1152/physiolgenomics.00026.2013

Liu Z, Liu S, Yao J, et al. The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts. *Nat Commun*. 2016;7:11757. doi:10.1038/ncomms11757

Ma J, Fan Y, Zhou Y, et al. Efficient resistance to grass carp reovirus infection in JAM-A knockout cells using CRISPR/Cas9. *Fish Shellfish Immunol*. 2018;76:206-215. doi:10.1016/j.fsi.2018.02.039

Madrazo AL, Campos MRS. Review of antimicrobial peptides as promoters of food safety: Limitations and possibilities within the food industry. *J Food Saf*. 2020;40(6):e12854. doi:10.1111/jfs.12854

Maier VH, Dorn KV, Gudmundsdottir BK, Gudmundsson GH. Characterisation of cathelicidin gene family members in divergent fish species. *Mol Immunol*. 2008;45(14):3723-3730. doi:10.1016/j.molimm.2008.06.002

Mao WF, Wang YP, Wang WB, Wu B, Feng JX, Zhu ZY. Enhanced resistance to *Aeromonas hydrophila* infection and enhanced phagocytic activities in human lactoferrin-transgenic grass

carp (*Ctenopharyngodon idellus*). *Aquaculture*. 2004;242(1-4):93-103. doi:10.1016/j.aquaculture.2004.07.020

Miller JC, Tan S, Qiao G, et al. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol*. 2011;29:143-148. doi:10.1038/nbt.1755

Mir LM, Bureau MF, Gehl J, et al. High-efficiency gene transfer into skeletal muscle mediated by electric pulses. *Proc Natl Acad Sci USA*. 1999;96(8):4262-4267. doi:10.1073/pnas.96.8.4262

Mitra A, Zhang Z. Expression of a human lactoferrin cDNA in tobacco cells produces antibacterial protein(s). *Plant Physiol*. 1994;106(3):977-981. doi:10.1104/pp.106.3.977

Mookherjee N, Anderson MA, Haagsman HP, Davidson DJ. Antimicrobial host defence peptides: functions and clinical potential. *Nat Rev Drug Discov*. 2020;19:311-332. doi:10.1038/s41573-019-0058-8

Nam BH, Seo JK, Go HJ, et al. Purification and characterization of an antimicrobial histone H1-like protein and its gene from the testes of olive flounder, *Paralichthys olivaceus*. *Fish Shellfish Immunol*. 2012;33(1):92-98. doi:10.1016/j.fsi.2012.04.006

Naylor RL, Hardy RW, Buschmann AH, Bush SR, Cao L, Klinger DH, et al. A 20-year retrospective review of global aquaculture. Nature 2021;591:551-63. doi:10.1038/s41586-021-03308-6.

Nguyen LT, Haney EF, Vogel HJ. The expanding scope of antimicrobial peptide structures and their modes of action. *Trends Biotechnol*. 2011;29(9):464-472. doi:10.1016/j.tibtech.2011.05.001

Nizet V, Johnson RS. Interdependence of hypoxic and innate immune responses. *Nat Rev Immunol*. 2009;9:609-617. doi:10.1038/nri2607

Noga EJ, Silphaduang U, Park NG, Seo JK, Stephenson J, Kozlowicz S. Piscidin 4, a novel member of the piscidin family of antimicrobial peptides. *Comp Biochem Physiol B-Biochem Mol Biol*. 2009;152(4):299-305. doi:10.1016/j.cbpb.2008.12.018

Noga EJ, Ullal AJ, Corrales J, Fernandes JMO. Application of antimicrobial polypeptide host defenses to aquaculture: Exploitation of downregulation and upregulation responses. *Comp Biochem Physiol D-Genomics Proteomics*. 2011;6(1):44-54. doi.o:10.1016/j.cbd.2010.06.001

Oh HY, Go H-J, Park NG. Identification and characterization of *Sa*RpAMP, a 60S ribosomal protein L27-derived antimicrobial peptide from amur catfish, *Silurus asotus*. *Fish Shellfish Immunol*. 2020;106:480-490. doi:10.1016/j.fsi.2020.06.038

Olivieri C, Buonocore F, Picchietti S, et al. Structure and membrane interactions of chionodracine, a piscidin-like antimicrobial peptide from the icefish *Chionodraco hamatus*. *Biochim Biophys Acta-Biomembr*. 2015;1848(6):1285-1293. doi:10.1016/j.bbamem.2015.02.030

Papareddy P, Rydengard V, Pasupuleti M, et al. Proteolysis of human thrombin generates novel host defense peptides. *PLoS Pathog*. 2010;6:e1000857. doi:10.1371/journal.ppat.1000857

Park CB, Kim MS, Kim SC. Novel antimicrobial peptide from *Bufo bufo gargarizans*. *Biochem Biophys Res Commun*. 1996;218(1):408-413. doi:10.1006/bbrc.1996.0071

Park IY, Park CB, Kim MS, Kim SC. Parasin I, an antimicrobial peptide derived from histone H2A in the catfish, *Parasilurus asotus*. *FEBS Lett*. 1998;437(3):258-262. doi:10.1016/S0014-5793(98)01238-1

Patat SA, Carnegie RB, Kingsbury C, Gross PS, Chapman R, Schey KL. Antimicrobial activity of histones from hemocytes of the Pacific white shrimp. *Eur J Biochem*. 2004;271:4825-4833. doi:10.1111/j.1432-1033.2004.04448.x

Pavlovic J, Zurcher T, Haller O, Staeheli. Resistance to influenza-virus and vesicular stomatitis-virus conferred by expression of human Mxa protein. *J Virol*. 1990;64:3370-3375. doi:10.1128/JVI.64.7.3370-3375.1990

Peatman E, Liu Z. Evolution of CC chemokines in teleost fish: a case study in gene duplication and implications for immune diversity. *Immunogenetics*. 2007;59:613-623. doi:10.1007/s00251-007-0228-4

Petrlova JW, Hansen FC, van der Plas MJA, et al. Aggregation of thrombin-derived C-terminal fragments as a previously undisclosed host defense mechanism. *Proc Natl Acad Sci USA*. 2017;114(21):E4213-E4222. doi:10.1073/pnas.1619609114

Petrlova JW, Mu X, Klesius PH. Expression profiles of seven channel catfish antimicrobial peptides in response to *Edwardsiella ictaluri* infection. *J Fish Dis*. 2012;35(3):227-237. doi:10.1111/j.1365-2761.2011.01343.x

Pridgeon JW, Klesius PH, Dominowski PJ, Yancey RJ, Kievit MS. Chicken-type lysozyme in channel catfish: Expression analysis, lysozyme activity, and efficacy as immunostimulant against *Aeromonas hydrophila* infection. *Fish Shellfish Immunol*. 2013a;35(3):680-688. doi:10.1016/j.fsi.2013.05.018

Pridgeon JW, Klesius PH, Dominowski PJ, Yancey RJ, Kievit MS. Recombinant goose-type lysozyme in channel catfish: Lysozyme activity and efficacy as plasmid DNA immunostimulant against *Aeromonas hydrophila* infection. *Fish Shellfish Immunol*. 2013b;35(4):1309-1319. doi:10.1016/j.fsi.2013.08.015

Qin Z, Li Y, Su B, et al. Editing of the luteinizing hormone gene to sterilize channel catfish, *Ictalurus punctatus*, using a modified zinc finger nuclease technology with electroporation. *Mar Biotechnol*. 2016;18:255-263. doi:10.1007/s10126-016-9687-7

Raman D, Sobolik-Delmaire T, Richmond A. Chemokines in health and disease. *Exp Cell Res*. 2011;317(5):575-589. doi:10.1016/j.yexcr.2011.01.005

Richards RC, O'Neil DB, Thibault P, Ewart KV. Histone H1: An antimicrobial protein of Atlantic salmon (*Salmo salar*). *Biochem Biophys Res Commun*. 2001;284(3):549-555. doi:10.1006/bbrc.2001.5020

Robinette D, Wada S, Arroll T, Levy MG, Miller WL, Noga EJ. Antimicrobial activity in the skin of the channel catfish *Ictalurus punctatus*: characterization of broad-spectrum histone-like antimicrobial proteins. *CMLS Cell Mol Life Sci*. 1998;54:467-475. doi:10.1007/s000180050175

Rollins BJ. Chemokines. *Blood*. 1997;90(3):909-928. doi:10.1182/blood.V90.3.909

Rose F, Bailey K, Keyte JW, Chan WC, Greenwood D, Mahida. Potential role of epithelial cell-derived histone H1 proteins in innate antimicrobial defense in the human gastrointestinal tract. *Infect Immun*. 1998;66(7):3255-3263. doi:10.1128/IAI.66.7.3255-3263.1998

Ryu HY, Hochstrasser M. Histone sumoylation and chromatin dynamics. *Nucleic Acids Res*. 2021;49(21):6043-6052. doi:10.1093/nar/gkab280

Said LB, Fliss I, Offret C, Beaulieu L. 2019. Antimicrobial Peptides: The New Generation of Food Additives. In: Melton L, Shahidi F, Varelis P, editors. *Encyclopedia of Food Chemistry*, vol 3, Elsevier; 2019, p. 576-582. doi:10.1016/B978-0-08-100596-5.21718-6.

Sakai M. Current research status of fish immunostimulants. *Aquaculture*. 1999;172(1-2):63-92. doi:10.1016/S0044-8486(98)00436-0

Sarmasik A, Warr G, Chen TT. Production of transgenic medaka with increased resistance to bacterial pathogens. *Mar Biotechnol*. 2002;4:310-322. doi:10.1007/s10126-002-0023-z

Schlesinger D, Goldstein G. Molecular conservation of 74 amino-acid sequence of ubiquitin between cattle and man. *Nature*. 1975;255:423-424. doi:10.1038/255423a0

Seo J-K, Lee MJ, Go H-J, et al. Purification and antimicrobial function of ubiquitin isolated from the gill of Pacific oyster, *Crassostrea gigas*. *Mol Immunol*. 2013;53(1-2):88-98. doi:10.1016/j.molimm.2012.07.003

Shike H, Lauth X, Westerman ME, et al. Bass hepcidin is a novel antimicrobial peptide induced by bacterial challenge. *Eur J Biochem*. 2002;269(8):2232-2237. doi:10.1046/j.1432-1033.2002.02881.x

Silveira RF, Roque-Borda CA, Vicente EF. Antimicrobial peptides as a feed additive alternative to animal production, food safety and public health implications: An overview. *Anim Nutr*. 2021;7:896-904. doi:10.1016/j.aninu.2021.01.004.

Simora RMC, Xing D, Bangs XR, et al. CRISPR/Cas9-mediated knock-in of alligator cathelicidin gene in a non-coding region of channel catfish genome. *Sci Rep*. 2020a;10:22271. doi:10.1038/s41598-020-79409-5

Simora RMC, Li S, Abass NY, Terhune JS, Dunham RA. Cathelicidins enhance protection of channel catfish, *Ictalurus punctatus*, and channel catfish ♀ × blue catfish, *Ictalurus furcatus* ♂ hybrid catfish against *Edwarsiella ictaluri* infection. *J Fish Dis*. 2020b;43(12):1553-1562. doi:10.1111/jfd.13257

Simora RMC, Wang W, Coogan M, Husseini N EI, Terhune JS, Dunham RA. Effectiveness of cathelicidin antimicrobial peptide against Ictalurid catfish bacterial pathogens. *J Aquat Anim Health*. 2021;33(3):178-189. doi:10.1002/aah.10131

Strahl BD, Allis CD. The language of covalent histone modifications. *Nature*. 2000;403:41-45. doi:10.1038/47412

Su J, Yang C, Zhu Z, Wang Y, Jang S, Liao L. Enhanced grass carp reovirus resistance of Mx-transgenic rare minnow (*Gobiocypris rarus*). *Fish Shellfish Immunol*. 2009;26(6):828-835. doi:10.1016/j.fsi.2008.12.007

Su Y. Isolation and identification of pelteobagrin, a novel antimicrobial peptide from the skin mucus of yellow catfish (*Pelteobagrus fulvidraco*). *Comp Biochem Physiol B-Biochem Mol Biol*. 2011;158(2):149-154. doi:10.1016/j.cbpb.2010.11.002

Sun L, Liu S, Wang R, Li C, Zhang J, Liu Z. Pathogen recognition receptors in channel catfish: IV. Identification, phylogeny and expression analysis of peptidoglycan recognition proteins. *Dev Comp Immunol*. 2014;46(2):291-299. doi:10.1016/j.dci.2014.04.018

Torres MDT, Sothiselvam S, Lu TK, de la Fuente-Nunez C. Peptide design principles for antimicrobial applications. *J Mol Biol*. 2019;431:3547-3567. doi:10.1016/j.jmb.2018.12.015.

Ullal AJ, Litaker RW, Noga EJ. Antimicrobial peptides derived from hemoglobin are expressed in epithelium of channel catfish (*Ictalurus punctatus*, Rafinesque). *Dev Comp Immunol*. 2008;32(11):1301-1312. doi:10.1016/j.dci.2008.04.005

Vaishnava S, Yamamoto M, Severson KM, et al. The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine. *Science*. 2011;334(6053): 255-258. doi:10.1126/science.1209791

Valero Y, Saraiva-Fraga M, Costas B, Guardiola FA. Antimicrobial peptides from fish: beyond the fight against pathogens. *Rev Aquac*. 2020;12(1):224-253. doi:10.1111/raq.12314

Viswanathan K, Frueh K, DeFilippis V. Viral hijacking of the host ubiquitin system to evade interferon responses. *Curr Opin Microbiol*. 2010;13(4):517-523. doi:10.1016/j.mib.2010.05.012

Waltz E. First genetically engineered salmon sold in Canada. *Nature*. 2017;548:148. doi:10.1038/nature.2017.22116

Wang D, Chen X, Zhang X, et al. Whole genome sequencing of the giant grouper (*Epinephelus lanceolatus*) and high-throughput screening of putative antimicrobial peptide genes. *Mar Drugs*. 2019;17(9):503. doi:10.3390/md17090503

Wang G, Li X, Wang Z. APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res*. 2016;44:D1087-1093. doi:10.1093/nar/gkv1278.

Wang G, Li X, Wang Z. APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res*. 2016;44(D1):D1087-D1093. doi:10.1093/nar/gkv1278

Wang J, Su B, Dunham RA. Genome-wide identification of catfish antimicrobial peptides: A new perspective to enhance fish disease resistance. *Rev Aquac*. 2022;(4):2002-2022. doi:10.1111/raq.12684

Wang Q, Bao B, Wang Y, Peatman E, Liu Z. Characterization of a NK-lysin antimicrobial peptide gene from channel catfish. *Fish Shellfish Immunol*. 2006;20(3):419-426. doi:10.1016/j.fsi.2005.05.005

Wang Q, Zhang L, Zhao J, You L, Wu H. Two goose-type lysozymes in *Mytilus galloprovincialis*: possible function diversification and adaptive evolution. *PLoS One*. 2012;7:e45148. doi:10.1371/journal.pone.0045148

Wang R, Sun L, Bao L, et al. Bulk segregant RNA-seq reveals expression and positional candidate genes and allele-specific expression for disease resistance against enteric septicemia of catfish. *BMC Genomics*. 2013;14:929. doi:10.1186/1471-2164-14-929

Wang Z, Huo X, Zhang Y, Gao Y, Su J. Carboxymethyl chitosan nanoparticles loaded with bioactive protein CiCXCL20a effectively prevent bacterial disease in grass carp (*Ctenopharyngodon idella*). *Aquaculture.* 2022;549(25):737745. doi:10.1016/j.aquaculture.2021.737745

Wei X, Babu VS, Lin L, et al. Hepcidin protects grass carp (*Ctenopharyngodon idellus*) against *Flavobacterium columnare* infection via regulating iron distribution and immune gene expression. *Fish Shellfish Immunol*. 2018;75:274-283. doi:10.1016/j.fsi.2018.02.023

Welker TL, Lim C, Yildirim-Aksoy M, Klesius PH. Dietary bovine lactoferrin increases resistance of juvenile channel catfish, *Ictalurus punctatus*, to enteric septicemia. *J World Aquacult Soc*. 2010;41(S1):28-39. doi:10.1111/j.1749-7345.2009.00330.x

WHO. Guidelines on use of medically important antimicrobials in food-producing animals. Geneva: World Health Organization 2017, p.1-67. http://apps.who.int/iris/bitstream/handle/10665/258970/9789241550130-eng.pdf;jsessionid=20765A3F28517C612028E9B85B000874?sequence=1.

Xiao X, Zhang Y, Liao Z, Su J, et al. Characterization and antimicrobial activity of the teleost chemokine CXCL20b. *Antibiotics.* 2020;9(2):78. doi:10.3390/antibiotics9020078

Xu P, Bao B, He Q, Peatman E, He C, Liu Z. Characterization and expression analysis of bactericidal permeability-increasing protein (BPI) antimicrobial peptide gene from channel catfish. *Dev Comp Immunol*. 2005;29(10):865-878. doi:10.1016/j.dci.2005.03.004

Yang D, Chen Q, Hoover DM, et al. Many chemokines including CCL20/MIP-3 alpha display antimicrobial activity. *J Leukoc Biol*. 2003;74(3):448-455. doi:10.1189/jlb.0103024

Yazawa R, Hirono I, Aoki T. Transgenic zebrafish expressing chicken lysozyme show resistance against bacterial diseases. *Transgenic Res*. 2006;15:385-391. doi:10.1007/s11248-006-0009-0

Yin LM, Edwards MA, Li J, Yip CM, Deber CM. Roles of hydrophobicity and charge distribution of cationic antimicrobial peptides in peptide-membrane interactions. *J Biol Chem*. 2012;28(10)7:7738-7745. doi:10.1074/jbc.M111.303602

Zasloff M. Magainins, a class of antimicrobial peptides from xenopus skin-Isolation, characterization of 2 active forms, and partial cDNA sequence of a precursor. *Proc Natl Acad Sci USA*. 1987;84(15):5449-5453. doi:10.1073/pnas.84.15.5449

Zhang X, Wang F, Dong Z, et al. A new strain of yellow catfish carrying genome edited myostatin alleles exhibits double muscling phenotype with hyperplasia. *Aquaculture*. 2020;523:735187. doi:10.1016/j.aquaculture.2020.735187

Zhang Y, Xiao X, Hu Y, et al. CXCL20a, a teleost-specific chemokine that orchestrates direct bactericidal, chemotactic, and phagocytosis-killing–Promoting functions, contributes to clearance of bacterial infections. *J Immunol*. 2021;207(7):1911-1925. doi:10.4049/jimmunol.2100300

Zhang Y-A, Zou J, Chang C-I, Secombes CJ. Discovery and characterization of two types of liver-expressed antimicrobial peptide 2 (LEAP-2) genes in rainbow trout. *Vet Immunol Immunopathol*. 2004;101(3-4):259-269. doi:10.1016/j.vetimm.2004.05.005

Zheng L, Qiu J, Liu H, Shi H, Chi C, Pan Y. Molecular characterization and antiparasitic activity analysis of a novel piscidin 5-like type 4 from *Larimichthys crocea*. *Mol Immunol*. 2021;129:12-20. doi:10.1016/j.molimm.2020.11.015

Zheng L-B, Hong Y-Q, Sun K-H, Wang J, Hong Y-J. Characteristics delineation of piscidin 5 like from *Larimichthys crocea* with evidence for the potent antiparasitic activity. *Dev Comp Immunol*. 2020;113:103778. doi:10.1016/j.dci.2020.103778

Zhong JY, Wang YP, Zhu ZY. Introduction of the human lactoferrin gene into grass carp (*Ctenopharyngodon idellus*) to increase resistance against GCH virus. *Aquaculture*. 2002;214(1-4):93-101. doi:10.1016/S0044-8486(02)00395-2

Zhong Z, Niu P, Wang M, et al. Targeted disruption of sp7 and myostatin with CRISPR-Cas9 results in severe bone defects and more muscular cells in common carp. *Sci Rep*. 2016;6:22953. doi:10.1038/srep22953

Zhou T, Yuan Z, Tan S, et al. A review of molecular responses of catfish to bacterial diseases and abiotic stresses. *Front Physiol*. 2018;9:1113. doi:10.3389/fphys.2018.01113

Zhu Q, Zhang L, Li L, Que H, Zhang G. Expression characterization of stress genes under high and low temperature stresses in the Pacific oyster, *Crassostrea gigas*. *Mar Biotechnol*. 2016;18:176-188. doi:10.1007/s10126-015-9678-0

Zhu R, Liu X-X, Lv X, et al. Deciphering transcriptome profile of the yellow catfish (*Pelteobagrus fulvidraco*) in response to *Edwardsiella ictaluri*. *Fish Shellfish Immunol*. 2017;70:593-608. doi:10.1016/j.fsi.2017.08.040

Zinkernagel AS, Johnson RS, Nizet V. Hypoxia inducible factor (HIF) function in innate immunity and infection. *J Mol Med*. 2007;85:1339-1346. doi:10.1007/s00109-007-0282-2

Zlotnik A, Yoshie O, Nomiyama H. The chemokine and chemokine receptor superfamilies and their molecular evolution. *Genome Biol*. 2006;7:243. doi:10.1186/gb-2006-7-12-243

Zlotnik A, Yoshie O. Chemokines: A new classification system and their role in immunity. *Immunity*. 2000;12(2):121-127. doi:10.1016/S1074-7613(00)80165-X

CHAPTER TWO

Generation of eco-friendly channel catfish, *Ictalurus punctatus,* harboring alligator cathelicidin gene with robust disease resistance by harnessing different CRISPR/Cas9-mediated systems

Abstract

The CRISPR/Cas9 platform holds promise for modifying fish traits of interest as a precise and versatile tool for genome manipulation. To reduce introgression of transgene and control reproduction, catfish species have been studied for upscaled disease resistance and intervening of reproduction to lower the potential environmental risks of introgression of escapees' as transgenic animals. Taking advantage of the CRISPR/Cas9-mediated system, we succeeded in integrating the cathelicidin gene from an alligator (*Alligator sinensis*; *As-Cath*) into the target luteinizing hormone (*lh*) locus of channel catfish (*Ictalurus punctatus*) using two delivery systems assisted by double-stranded DNA (dsDNA) and single-stranded oligodeoxynucleotides (ssODNs), respectively. In this study, high knock-in (KI) efficiency (22.38%, 64/286) but low on-target was achieved using the ssODN strategy, whereas adopting a dsDNA as the donor template led to an efficient on-target KI (10.80%, 23/213). On-target KI of *As-Cath* was instrumental in establishing the *lh* knockout (LH^{-}_As-$Cath^{+}$) catfish line, which displayed heightened disease resistance and reduced fecundity compared to the wild-type sibling fish. Furthermore, implanting with HCG and LHRHa can restore the fecundity, spawnability and hatchability of the new transgenic fish line. Overall, we replaced the *lh* gene with an alligator cathelicidin transgene and then administered hormone therapy to gain complete reproductive control of disease-resistant transgenic catfish in an environmentally sound manner. This strategy not only effectively improves the consumer-valued traits, but also guards against genetic contamination. This is a breakthrough in aquaculture genetics to confine fish reproduction and prevent the establishment of transgenic or domestic genotypes in the natural environment.

Keywords:

Genome editing, ssODN, dsDNA, antimicrobial peptide, immune, reproductive confinement, aquaculture

1. Introduction

Innovative biotechnologies continuously develop as science advances, benefiting food production, quality as well as animal and human welfare. Since its inception, CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9) has served as a prototype in genome engineering, paving the way for new possibilities in transgenesis and breeding. Two mechanisms are involved for DNA repair when double strand breaks are induced by the CRISPR/Cas9 complex: non-homologous end joining (NHEJ) and homology-directed repair (HDR) (Doudna and Charpentier, 2014). Both mechanism-mediated strategies have been employed in aquaculture to improve the consumer-valued qualities targeted within genetic breeding programs. These harness the NHEJ repair pathway to knock out (KO)/disrupt functional genes or knock in (KI) exogenous genes of interest via HDR at the expected locus to improve the target traits.

Numerous CRISPR/Cas9 systems have emerged recently to improve target-editing efficiency for KI via the HDR pathway. Success has been observed in model animals have been shown successes using ssODN-mediated KIs for the targeted insertions of small DNA fragments since single-stranded oligodeoxynucleotides (ssODNs) act as templates for repairing DNA damage (Storici et al., 2006; Chen et al., 2011; Wefers et al., 2014). Yoshimi et al. (2016) have optimized the ssODN-mediated approach to knock-in larger sequences by the combination of CRISPR/Cas9 system with two 80-bp ssODNs in length. In contrast to conventional plasmid donors, the donor vector used in this system does not require homologous arms (HAs), enabling the insertion of a large vector (CAG-GFP, 4.8 kb) into the designated site (*rRosa26*) with a ~10% integration rate in rat zygotes (Yoshimi et al., 2016). Later, using the CRISPR/Cas9-ssODNs mediated KI system, a 10.96% KI efficiency in sheep zygotes was determined (Mehravara et al., 2019). Boel et al. (2018) first applied this optimized system to a fish model, zebrafish (*Danio rerio*), and sequencing results revealed that erroneous repair was more likely to occur when ssODNs were used as repair templates. Alternatively, the modified donor plasmid containing two HAs flanked by two single guide (sgRNA)-targeted sequences (double-cut donors) typically results in a site-specific KI with a high integration rate (Hisano et al., 2015; Zhang et al., 2017), and this HA-medicated KI has been adapted to zebrafish and medaka (*Oryzias latipes*) (Murakami et al., 2017; Zhang et al., 2017). Theoretically, if we directly offer a linear double-stranded DNA (dsDNA) flanked by two HAs derived from 5'- and 3'- ends of the targeted site

and ignore the difference in stability between circular DNA and dsDNA donors, the KI efficiency will increase by convention. In addition to the type of donors, a proper concentration of each component of the CRISPR/Cas9 system has a great positive impact on KI by reducing off-target events and embryo lethality. In this regard, we anticipate achieving extremely efficient KI if a reliable delivery system and an optimized dosage of components are chosen in a non-model fish species.

Currently, transgenesis and CRISPR/Cas9-mediated genome editing have revolutionized traditional theories to accelerate the pace of aquaculture breeding programs, and delivered edible commercial products, such as the genetically modified AquAdvantage salmon (Ledford, 2015; Waltz, 2017), gene-edited tiger puffer fish and red sea bream (https://doi.org/10.1038/s41587-021-01197-8, 2022). Although the NHEJ strategy predominates in altering the consumer-focused traits of fish species, including growth, coloration, and reproduction, the HDR-mediated KI is an effective way to improve the omega-3 fatty acid content and disease resistance (Wang et al., 2022; Xing et al., 2022ab). In comparison to the non-insertion of KO mutations, the integration of foreign genes by harnessing the HDR pathway usually raises concerns about low KI efficiency and introgression, which directly impact the advocacy of this method and the consumer acceptance of gene-inserted fish (Dunham and Su, 2020). As a result, it is imperative to devise a strategy for both improving the desired traits and preventing introgression to alleviate public concerns about gene-inserted animals. Fortunately, numerous genome-editing-based studies have demonstrated that it is possible to render fish reproductively sterile by altering/disrupting key genes involved in reproduction via the NHEJ repair pathway. Thus, potentially reducing negative environmental effects associated with genetically modified fish (Blix et al., 2021; Qin et al., 2022; Yang et al., 2022). Luteinizing hormone (*lh*) gene regulates gametogenesis and gestation through binding the receptor (Grier, 1981; Yamaguchi et al., 2021). LH-deficient female zebrafish are infertile, whereas the mutant males are fertile, indicating that the *lh* gene facilitates fish oocyte maturation and triggers ovulation (Chu et al., 2014). In addition, interruption of the *lh* gene in channel catfish and white-edged rockfish (*Sebastes taczanowskii*) can result in the production of sterile lines (Qin et al., 2016; Yamaguchi et al., 2021).

Large-scale disease outbreaks are inevitable, and methods of disease control need to be improved. Antimicrobial peptides (AMPs) are polypeptides that serve as substitutes for antibiotics in a

variety of species' initial line of defense (innate immunity) against microbial invasions without developing considerable antibiotic resistance (Akira et al., 2006; Wang et al., 2016). AMPs and antimicrobial peptide genes (AMGs) including cecropin, hepcidin, piscidin, epinecidin-1, lysozyme, and lactoferrin have been used for decades to improve disease resistance in a variety of aquatic animals, as feed supplements or transgenes (Wang et al., 2022; Wang et al., 2023). Cathelicidins are a particularly important AMP family, sharing the common cathelin-like domain (Mookherjee et al., 2020) and exhibiting broad-spectrum antimicrobial and immune-modulating activities (Hilchie et al., 2013). Recent investigations have shown that alligator-derived cathelicidin (*As-Cath*) inhibits fish pathogens both *in vivo* and *in vitro* (Chen et al., 2017; Simora et al., 2020; Simora et al., 2021). Therefore, integrating the AMG into the genomic DNA has broad prospects for establishing novel disease-resistant fish lines.

Fish transgenic for AMGs could provide a significant option to address disease problems, however, and additional goal would be to prevent the possibility of breeding of escapees with wild populations. Hypothetically, a reproductive gene such as *lh*, responsible for gametogenesis and gestation could be knocked out at the DNA level with the replacement of a cathelicidin gene, leading to sterile fish with heightened disease resistance. Genome-edited sterilized fish from this approach would have fertility temporarily restored with hormone therapy used for artificial spawning of fish, and it is achievable to produce environmentally-compatible and disease-resistant fish lines. In this study, two CRISPR/Cas9 delivery systems: HA- and ssODN-mediated KI were employed to insert the *As-Cath* gene at the channel catfish (*Ictalurus punctatus*) *lh* locus to develop a reversibly sterile and disease-resistant line. We compared the KI efficiency, hatchability and fry survival from various systems, and then restored the fertility of *As-Cath*-integrated sterile of P_0 founders through hormone therapy. In addition, the bacterial resistance of P_0 and F_1 individuals from the new fish line was further evaluated.

2. Materials and methods

2.1 Ethical approval

The care and use of animals followed the applicable guidelines from expert training courses. Experimental protocols in the current study were approved by the Auburn University Institutional Animal Care and Use Committee (AU-IACUC). All fish studies were conducted in compliance with the procedures and standards established by the Association for Assessment and

Accreditation of Laboratory Animal Care (AAALAC).

2.2 Target locus for gene insertion

As the target integration site, we selected the *lh* gene, which is widely expressed in the theca cells of the ovary and aids in egg maturation and ovulation during gonadal development (Chu et al., 2014). Based on the published genome of channel catfish (Liu et al., 2016), the chosen *lh* site for sgRNA targeting was located in the middle of exon 2 (Figure 4A-B). The inserted segment was derived from the coding sequence (CDS) of the cathelicidin gene of *Alligator sinensis* (*As-Cath*, GeneBank accession number XM_006037211.3) (Chen et al., 2017).

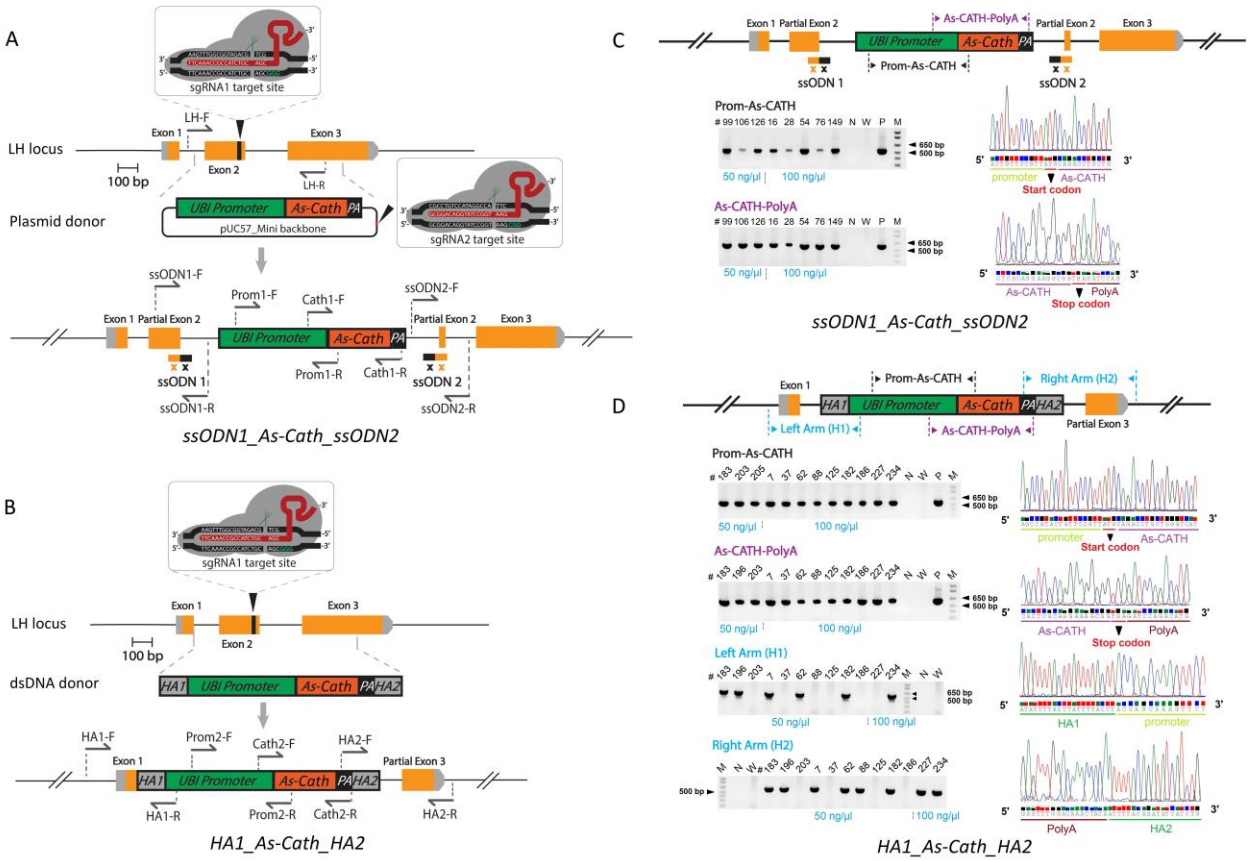


Figure 4. Single-stranded oligodeoxynucleotide (ssODN) and linear double-stranded DNA (dsDNA) with CRISPR/Cas9 mediating knock-in (KI) at the *luteinizing hormone* (*lh*) locus of channel catfish. (A) Schematic illustration of the insert-specific region for the cathelicidin gene from *Alligator sinensis* (*As-Cath*) KI via the two-hit two-oligo (2H2OP) system assisted by ssODNs at the *lh* locus, named as the ssODN1_As-Cath_ssODN2 construct. The structure of the *lh* gene's exons is constructed by yellow bars, sgRNAs-targeted sites are indicated by black triangles, and the target sequences are detailed in rectangular boxes. The protospacer-adjacent motif (PAM) is highlighted in green. Primer sets are

illustrated, showing the strategy to test *lh* mutation, ssODN1/ssODN2 junctions, the UBI promoter region and the insert-specific region of the *As-Cath* gene using PCR amplifications. **(B)** Schematic diagram of the *As-Cath* KI via the dsDNA system, named HA1_As-Cath_HA2 donor. Primers show the strategy to test the HA junctions, UBI promoter region, and *As-Cath* gene region. **(C)** TAE agarose gel of PCR amplicons showing off-target positive detection of the ssODN1_As-Cath_ssODN2 construct using 2H2OP method. The promoter region (Prom-As-CATH, 519 bp) and *As-Cath* region (As-CATH-PolyA, 591 bp) were illustrated with sequencing results. **(D)** TAE agarose gel of PCR amplicons showing on-target positive detection of the HA1_As-Cath_HA2 construct using dsDNA method. The targeted gene regions (Prom-As-CATH, 542 bp and As-CATH-PolyA, 597 bp) and the junctional regions (HA1, 573 bp and HA2, 598 bp) were determined with sequencing results. The numbers on the top of the gel images indicate the sample IDs of the fish. Lane N, negative control using water as template; Lane W, wild-type control (nCT); Lane P, positive (plasmid or dsDNA donor) control; Lane M, DNA marker (1 kb), 500 and 650-bp bands are highlighted with black triangles; 50 and 100 ng/μL show the different doses of donors: plasmid or dsDNA.

2.3 Design of donor DNA, sgRNA and CRISPR/Cas9 system

Gene-targeted KI can be engineered via HDR using the dsDNAs or ssODNs as donor templates. In the current study, we employed two CRISPR/Cas9-mediated systems to conduct targeted KI of the *As-Cath* fragment at the *lh* locus. For the first system, the CDS of the *As-Cath* gene was cloned into the pUC57_mini vector at the *Eco*RV enzyme digestion site to create the ssODN1_As-Cath_ssODN2 construct as a plasmid donor. Two sgRNAs (sgRNA1 and sgRNA2) were co-injected to operate as “scissors”, cutting the *lh* gene and linearizing the plasmid donor, respectively, and provided two short ssODNs to ligate the ends of both cut sites, labeled as the 2H2OP system (Figure 4A). ssODN1 consists of 80 bp, of which the upstream 40 bp are derived from partial exon 2 of *lh* gene and the remaining 40 bp are homologous to pUC57_mini backbone. For ssODN2, the upstream 40 bp are from the pUC57_mini backbone, while downstream 40 bps come from a portion of exon 2 of the *lh* gene. The dsDNA donor was created by constructing the *As-Cath* CDS sequence flanked with two homology arms (HAs) of 300 bp derived from the *lh* gene of channel catfish on either side of the insert DNA, and we tagged the second construct as HA1_As-Cath_HA2. More specifically, 163 bp of HA1 (the left homology arm) are derived from the upstream of exon 2; 136 bp are identical to intron 1, and 1 bp originated from exon 1. HA2 (the right homology arm) contains 21 bps from exon 2’s downstream; 85 bps from intron 2 and 194 bps from upstream of exon 3 (Appendix 4). Here, we used one sgRNA (sgRNA1) to cut the *lh* site in the channel catfish genomic DNA and provided a linear dsDNA as the donor template, and this system was labeled as dsDNA (Figure 4B). For both constructs, the expression of the *As-Cath* gene was driven by the zebrafish ubiquitin (UBI)

promoter (Mosimann et al., 2011). The linear dsDNA, circular plasmid and ssODNs were synthesized by Genewiz (South Plainfield, NJ).

The sgRNAs were selected via the CRISPR design online tool (CRISPR Guide RNA Design Tool, Benching, https://zlab.bio/guide-design-resources) that targeted the *lh* gene of channel catfish and the donor plasmid. Candidate sgRNA sequences were compared to the whole genome of channel catfish via the Basic Local Alignment Search Tool to avoid cleavage of off-target sites. In addition, putative off-target sites were excluded using the online tool Cas-OFFinder (http://www.rgenome.net/cas-offinder/) (Bae et al., 2014). Eventually, sgRNA1 for *lh* locus and sgRNA2 for donor plasmid were obtained. The Maxiscript T7 kit (Thermo Fisher Scientific, Waltham, MA) was used to generate sgRNAs *in vitro*, according to the instructions. Then purified sgRNAs were prepared using the RNA Clean and Concentrator Kit (Zymo Research, Irvine, CA). The concentration and quality of sgRNAs were detected with Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and 1% agarose gel with 1 × tris-borate-EDTA (TBE) buffer, respectively. The synthetic sgRNAs were diluted to a concentration of ~ 300 ng/μL and then divided into PCR tubes (2 μL/tube) and stored at − 80 ℃ until use. The Cas9 protein powder was purchased from PNA BIO Inc. (Newbury Park, CA), and was diluted with DNase/RNase-free water to 50 ng/μL, keeping at − 20 ℃ until use. Single guide RNAs and universal primer used in this study are listed in Table 2. Two different dosages of the donor DNA template and two control groups were set up: 50 ng/μL, 100 ng/μL, sham-injected control (iCT, only the 10% phenol red solution was injected) and non-injected control (nCT, no injection) for each KI system.

Table 2. Target sequences of sgRNAs and the universal primer used in the present study. Underlined sequences represent the protospacer adjacent motif.

sgRNA	Targeted sequence for sgRNA (5'-3')
sgRNA1	5'- TTCAAACCGCCATCTGCAGCGGG -3'
sgRNA2	5'- GCGGACAGGTATCCGGTAAGCGG -3'
Universal primer	TTTTGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAG CCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC

2.4 Transgenic fish production and rearing

Mature channel catfish females and males were paired for artificial spawning according to

Elaswad et al. (2018) with some modifications. Briefly, we selected individuals weighing more than 1.5 kilograms for spawning. Female channel catfish were implanted with 75 μg/kg of luteinizing hormone-releasing hormone analog (LHRHa) to induce ovulation, then eggs were gently stripped in a 20-cm greased spawning pan. Mature males were euthanized; testes were collected, rinsed, weighed, and crushed; and sperm were prepared in 0.9 % saline solution (g:v = 1:10). Two milliliters of sperm solution were added to approximately 300 eggs and gently mixed. After a one-minute mixing, sufficient pond water was added to the eggs to activate the sperm, then the sperm/egg mixture was gently swirled for 30 s. More water was added, and the embryos were kept in a single layer in the pan, and the embryos were allowed to harden for 15 min before microinjection.

The CRISPR/Cas9 system used for KI microinjections was combined with Cas9 protein, sgRNA and donor template in the ratio of 2:1:1, including one component of phenol red as an indicator. For the ssODN1_As-Cath_ssODN2 construct (2H2OP system), 8 μL of Cas9 protein (50 ng/μL), 2 μL of sgRNA1/sgRNA2 (300 ng/μL), 2 μL of donor plasmid (50 ng/μL, 100 ng/μL), 2 μL of ssODN1/ssODN2 (50 ng/μL, 100 ng/μL) and 2 μL of phenol red solution were mixed for microinjection (Total 8 + 2 + 2 + 2 + 2 + 2 + 2 = 20 μL). With respect to the HA1_As-Cath_HA2 construct (dsDNA system), 4 μL of Cas9 protein (50 ng/μL), 2 μL of sgRNA1 (300 ng/μL), 2 μL of donor dsDNA (50 ng/μL, 100 ng/μL), 2 μL of phenol red and 10 μL of DNase-free water were mixed to bring it up to 20 μL in total. For each mixture of the CRISPR/Cas9 system, we mixed Cas9 protein and sgRNA first and incubated them on ice for 10 min, then the donor templates were supplemented. For the iCT group, we only injected phenol red (diluted with 0.9 % saline). The mixed solution for each treatment was microinjected into one-cell stage embryos as previously described (Khalil et al., 2017). Every 6 μL of the mixture was loaded into a 1.0 mm OD borosilicate glass capillary that was pulled into a needle by a vertical needle puller (David Kopf Instruments, Tujunga, CA), and injected into 600 embryos. We injected 1,000 embryos dividing them into 5 random replicates for each treatment, and another 200 embryos with 3 replicates were prepared for each control group, respectively. All these embryos were from the same parents, and the microinjection was terminated after 90 min post-fertilization.

All injected and control embryos were transferred into 10-L tubs filled with 7-L Holtfreter's solution (59 mmol NaCl, 2.4 mmol $NaHCO_3$, 1.67 mmol $MgSO_4$, 0.76 mmol $CaCl_2$, 0.67 mmol

KCl) (Armstrong and Malacinski, 1989) and 10 − 12 ppm doxycycline for hatching immediately after microinjection. All tubs were placed in the same flow-through hatching trough and a heater was put upstream of the trough to ensure that the water temperature was 26 − 28 ℃ while dissolved oxygen levels were > 5 ppm via continuous aeration with airstones. Holtfreter's solution was replaced twice per day and dead embryos/fry were collected and recorded daily during hatching to analyze hatchability, fry survival rate and genotype. The hatched fry were transferred to a Holtfreter's solution without doxycycline and fed with live *Artemia* nauplii four times per day. After one week of culture in tubs, all fry from each treatment were stocked separately into 60 L aquaria (120 fish/tank) in a recirculating system for growth experiments. Feed pellet size was adjusted according to the size of the fish's mouth as the fish grew. In detail, fry in tanks fed with Purina® AquaMax® powdered feed (50% crude protein, 17% crude fat, 3% crude fiber, and 12% ash) four times per day for two months. Then fingerlings were fed with Aquaxcel WW Fish Starter 4512 (45% crude protein, 12% crude fat, 3% crude fiber, and 1% phosphorus) twice a day for two months. Juvenile fish were fed with WW 4010 Transition feed (40% crude protein, 10% crude fat, 4% crude fiber, and 1% phosphorus) once a day (Coogan et al., 2022). All fish were fed to satiation.

2.5 Integration analysis and mutation detection

After a 4-month culture, all fingerlings (20 − 40 g) were pit-tagged (Biomark Inc., Boise, Idaho, USA) to distinguish each individual, the fish from different treatments were then mixed together and randomly dispersed into two circular tanks (1,200 L volume filled with ~800 L of water) with the same density (120 fish/tank) for growth comparison monthly. Meanwhile, the pelvic fin clip and barbel were taken from anesthetized fish for DNA extraction and genotypic identification. During this phase, all fish received WW 4010 Transition feed once a day to satiation. Different genotyping strategies were involved for these two constructs: ssODN1_As-Cath_ssODN2, the CDS region of *As-Cath* was amplified to confirm gene insertion using primers Cath1-F/R (forward and reverse), and the promoter region was amplified via primers Prom1-F/R. As for the junctions, ssODN1 and ssODN2 regions were amplified using primers ssODN1-F/R and ssODN2-F/R to determine whether it was a target-site insertion. With respect to the HA1_As-Cath_HA2 construct, the *As-Cath* and promoter regions were detected using primers Cath2-F/R and Prom2-F/R, respectively. Then the left HA and right HA junctions were

amplified via primers HA1-F/R and HA2-F/R. Primers were designed using the online software Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) and listed in Table S1 (Appendix 1). PCR was performed in a 10-μL system and PCR products were resolved and visualized by running 1.0% agarose gel with 1 × tris-acetate-EDTA (TAE) buffer, and a bright band of each region with the corresponding length indicated an on-target positive (LH^-_As-$Cath^+$). Here, if we can determine that some individuals have been inserted with the As-Cath fragment, but we can not detect the junctional regions (HA- or ssODN-region), we then conclude them as potential off-target positives (LH^+_As-$Cath^+$).

With respect to the LH^+_As-$Cath^+$ fish, we selected 60 individuals to be tested for *lh* mutations. In this case, PCR was performed in a 20 μL-volume system using Expand High Fidelity^{PLUS} PCR System (Roche Diagnostics, Indianapolis, IN, USA) according to Elaswad et al. (2018), and LH-F/R primers were used in both constructs. Then, the surveyor mutation detection assay was performed via Surveyor Mutation Detection Kit (Integrated DNA Technologies, IDT, Coralville, Iowa, USA) according to the detailed instructions (Qiu et al., 2014). A negative control reaction was included in the assay by using genomic DNA from the nCT group. Surveyor-digested DNA samples were electrophoresed for 1 hour in a 2% agarose gel using 1 × TBE buffer and compared to wild-type samples.

2.6 DNA sequencing

For the integrated *As-Cath*, promoter and junction sequences, PCR of positive samples was performed in a 50 μL-volume system. Then the PCR products were purified using the QIAquick^R PCR Product Purification Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. Before sequencing, all purified DNA samples were quantitated and identified using Nanodrop and by running 1.0% agarose gel. Primers Cath1-F/Cath2-F and Prom1-F/Prom-2F were used for sequencing of *As-Cath* and promoter regions for HA1_As-Cath_HA2 and ssODN1_As-Cath_ssODN2 constructs, respectively; primers HA1-F/HA2-F and ssODN1-F/ssODN2-F were used for sequencing of junctional regions for these two constructs, respectively.

Regarding *lh* mutations, we cloned the PCR products of putative mutant individuals using TOPO

TA Cloning Kit (Invitrogen, Carlsbad, CA) before sequencing following the instructions with some modifications. Briefly, PCR was performed on each mutant individual that was previously identified with Surveyor assay using the primers LH-F/R for the next cloning steps. In addition, the DNA of three wild-type individuals from the nCT group was prepared using the same primers and procedures, then combined into one reaction and cloned as a wild-type control for sequencing. After cloning, we transformed the pCR^{TM}4-TOPO vector containing the PCR products into One Shot TOP10 Electrocomp^{TM} *E. coli* (Invitrogen, Carlsbad, CA) as previously described (Elaswad et al., 2018). Then 15 single colonies were randomly picked up to perform Colony PCR, and LH-F primer was used for the sequencing of *lh* mutant samples.

2.7 Determination of mosaicism and transgene expression

Five 12-month-old on-target positive fish and five sham-injected control fish were randomly chosen and sacrificed. Fourteen tissues, including skin, liver, kidney, spleen, blood, intestine, gill, stomach, fin, barbel, muscle, eye, brain, and gonad of each individual were collected in 1.5 mL tubes and immediately transferred into liquid nitrogen for DNA and RNA isolation. PCR and quantitative real-time PCR (qRT-PCR) were conducted to detect the *As-Cath* gene's potential mosaicism and mRNA level. Total RNAs were isolated from various tissues using TRIzol reagent (Thermo Fisher Scientific) and were reverse transcribed to cDNA using iScript™ Synthesis Kit (Bio-Rad, Hercules, CA) following the manufacture protocols.

qRT-PCR was performed on a C1000 Thermal Cycler using SsoFast™ EvaGreen Supermix kit (Bio-Rad, Hercules, CA) according to the instructions. Concentrations of the cDNA products were diluted to 250 ng/μL, and 1 μL template was used in a 10 μL PCR reaction volume. The mRNA level of 18S rRNA was used as an internal control, and the detailed qRT-PCR procedure was set up according to Coogan et al. (2022). The primers (Cath_RT-F and Cath_RT-R) used for qRT-PCR are listed in Table S1 (Appendix 1). The CFX Manager Software (version 1.6, Bio-Rad) was used to collect the raw crossing-point (C_t) values. The expression level of a target gene to the 18S rRNA gene from transgenic fish against non-transgenic sibling fish was converted to fold differences. Each sample was analyzed in triplicate using the formula $2^{(-\Delta\Delta CT)}$, which sets the zero expression of the non-transgenic full-siblings to 1× for comparison.

2.8 Reproductive evaluation and restoration of parental KI fish

All P_0 fish were stocked into a 0.04-ha earthen pond at Fish Genetics at Auburn University for growth and maturation. At the age of two years, some P_0 individuals are expected to reach sexual maturity (Davis, 2099). To evaluate the reproduction of two-year-old KI founders, on-target positive (LH^-_As-$Cath^+$), off-target positive (LH^+_As-$Cath^+$), and wild-type (WT) fish were selected to conduct a three-round mating experiment. Firstly, 3 pairs of WT, 6 pairs of LH^-_As-$Cath^+$, and 4 pairs of LH^+_As-$Cath^+$ mature parents were randomly placed into 13 tanks (60 × 45 × 30 cm^3) for a two-week natural spawning to evaluate the spawnability of each genotype, and egg masses were collected from the spawnable parents. Then we primed the males with a 50 μg/kg LHRHa implant and 1600 IU/kg human chorionic gonadotropin (HCG) in the unspawned groups with a one-week observation to determine if LH^-_As-$Cath^+$ females were fertile. After this period, we recruited 6 more pairs of LH^-_As-$Cath^+$ fish to perform a 3 × 4 factorial design with 3 dosages of a combination of HCG and LHRHa implant (1200 IU/kg HCG + 50 μg/kg LHRHa, 1600 IU/kg HCG + 50 μg/kg LHRHa, 2000 IU/kg HCG + 50 μg/kg LHRHa) and 0.85% NaCl injected control group to assess the effects of hormone therapy. A 30-g egg mass for each genotype with 3 replicates was collected to calculate the fecundity (eggs/kg body weight [BW]). The masses were then transferred into tubs for hatchability and fry survival determination. Fish were fed ad libitum throughout the experiment.

2.9 Generation and genotype analysis for F_1 fish

All the fry were separated into 60 L tanks by different genotypes. After 4 months of culture, fin clips and barbels were collected for DNA extraction from 60 F_1 individuals of each genotype except the control groups. The same culture and genotyping procedures as described above were applied to the F_1 generation.

2.10 Experimental challenge with *Flavobacterium covae* and *Edwardsiella ictaluri*

Gene-edited channel catfish were cultured in 60 L aquariums in the greenhouse of the Fish Genetics Laboratory at Auburn University (approved by AU-IACUC). To determine the resistance against pathogens, both P_0 and F_1 fish were challenged by *F. covae* and *E. ictaluri*.

F. covae challenge. Healthy P_0 fingerlings with body weight 150.62 ± 4.24 g (mean ± SEM), including four genotypes (15 fish/genotype): LH^-_As-$Cath^+$, LH^+_As-$Cath^+$, negative LH^+_As-$Cath^-$ (negative fish without *As-Cath* insertion or *lh* mutation) and WT were mixed and acclimated in one hatching trough for five days and then transferred to a 1,800-L tank in the challenge room for acclimation for another 24 to 48 h prior to bacterial infections. All fish were randomly/equally separated into two 60-L buckets (30 L water). Briefly, a revived *F. covae* isolate (strain ALG-00-530) on modified Shieh agar (MSA) was inoculated into multiple cultures of 12 mL of modified Shieh broth (MSB) in 50-mL sterile flasks and grown in a shaker incubator at 150 rpm for 12 hours at 28°C. These cultures were then expanded into 200 mL cultures (5 mL additions) in 500 mL flasks and grown for another 12 h. The optical density was adjusted to OD_{540} = 0.731 and then spread plate dilutions were performed to determine the final inoculum concentration. One hundred microliters of each inoculum were serially diluted and spread onto MSA agar plates in duplicate and incubated at 28 °C for 48 h to quantify the concentration of the inoculum. Two flasks containing 325 mL of inocula (4.55×10^8 CFU/mL) were immediately added to two 60 L buckets with fish following preparation, respectively. Then the fish were immersed statically in buckets for 1.5 hours at ~28 °C (immersion dose: 2.46×10^6 CFU/mL); afterward, all fish were gently moved back into the 1,800-L tank containing 1,000-L water and water flow was resumed. Meanwhile, a mock-challenged tank was used as the control but incorporated another 40 fish in 30 L of rearing water for 1.5 hours with sterile modified Shieh broth (325 mL) instead of the bacterial culture. With respect to the challenge of F_1 fry (3.15 ± 0.24 g), four families of F_1 fry (45 fish/family): LH^-_As-$Cath^+$, LH^+_As-$Cath^+$, LH^+_As-$Cath^-$ and WT were selected, and each family was randomly divided into three replicates with 15 fish per basket. The same challenge procedure and strain of *F. covae* with a dose of 4.75×10^8 CFU/mL (immersion dose: 2.57×10^6 CFU/mL) were implanted for the F_1 generation.

E. ictaluri challenge. Sixty P_0 fish (142.62 ± 3.72 g) including the above four genotypes, were prepared for the *E. ictaluri* challenge. *E. ictaluri* (S97-773) was provided by the USDA-ARS, Aquatic Animal Health Research Unit, Auburn, AL. The detailed procedures of the *E. ictaluri* challenge were performed according to Simora et al. (2020) with some modifications. Briefly, 1 mL of frozen glycerol stock of *E. ictaluri* was inoculated into 20-mL brain–heart infusion broth (BHIB; Hardy Diagnostics) at 26°C in a shaker incubator at 180 rpm for 24 hours. And then bacteria were subcultured into 1-L BHIB for another 24 hours at the same condition until the cell

density reached ~1×10^8 CFU/mL based on the OD_{600} value. All 60 P_0 individuals were transferred into one 1,800-L tank for the challenge. Before starting *E. ictaluri* infection, water was lowered to a total of 100 L, then one liter of *E. ictaluri* suspension containing 3.2×10^8 CFU/mL cells was added to the tank resulting in a final immersion dose of 3.2×10^6 CFU/mL. Fish were immersed statically for 2 hours with aeration > 5 ppm, then water was restored. In addition to infected groups, one control tank containing 30 fish received only BHIB as a mock-challenged group. With respect to the challenge of F_1 fingerlings (54.27 ± 1.49 g), a total of four genotypes containing 60 fish were selected, and the same challenge procedure and strain of *E. ictaluri* with a dose of 2.8×10^8 CFU/mL (immersion dose: 2.8×10^6 CFU/mL) were implanted for the F_1 generation.

During the first 72 h of the experiment, we checked for mortality every four hours and then three times daily. Challenged fish were continuously monitored for 10 days for external clinical signs of *F. covae*/*E. ictaluri* and confirmation of bacteria colony growth by isolating bacteria from the kidney and liver to determine the cause of death, and dead individuals were recorded over time.

2.11 Statistical analysis

Spawnability, hatchability, fecundity, fry survival rate, and growth data were analyzed using one-way ANOVA/Tukey's multiple comparisons test to determine the mean differences among treatments. To compare the KI efficiency of different groups, one-way ANOVA/Tukey's multiple comparisons and odds ratio (OR) were adopted. The survival curves of challenge experiments from different genotypes were compared by the Kaplan-Meier plots followed by Log-rank (Mantel-Cox) test. All statistical analysis was achieved via GraphPad Prism 9.4.1 (GraphPad Software, LLC). Gene expression between transgenic and non-transgenic fish was analyzed with an unpaired Student's two-sample *t*-test. Statistical significance was set at $P < 0.05$, and all data were presented as the mean ± standard error (SEM).

3. Results

3.1 Targeted KI of *As-Cath* gene into the *lh* locus

Both the 2H2OP and dsDNA systems can induce As-Cath-integrated catfish lines with high

integrated ratios, but the 2H2OP system had significant off-target effects (Figure 4CD, Figure S1-S4 in Appendix 2). More specifically, the 2H2OP system containing 50 ng/μL of donors (2H2OP50) showed the highest KI efficiency at 27.61% (37/134), followed by the groups 2H2OP100 (17.76%, 27/152), dsDNA50 (12.21%, 26/213) and dsDNA100 (10.25%, 25/244) (Table S2 in Appendix 1). Although the 2H2OP50 group can introduce the highest KI efficiency ($P < 0.01$) (Figure 5A), and 2H2OP system or 50 ng/μL of donors bring a significantly higher KI efficiency than the dsDNA method ($P = 0.0001$) or 100 ng/μL of donors ($P = 0.00469$) (Figure 5BC). However, the dsDNA with 50 ng/μL donors demonstrated the highest on-target KI efficiency (10.80%, 23/213) compared to other treatments ($P < 0.01$) (Figure 5D). In contrast, only one on-target KI case was observed in the 2H2OP system, which was significantly lower than that in the dsDNA ($P < 0.0001$) (Figure 5E). Although different dosages of donors exhibited a significant effect on the total KI efficiency, our results indicated that this difference was not significant in the on-target KI ($P = 0.3577$) (Figure 5F).

According to the odds ratio, the 2H2OP system and low dosage tended to bear a higher total integrated rate which was 2.30 and 1.47 times than that of the dsDNA (OR = 2.30 for 2H2OP *vs.* dsDNA) and high dosage (OR = 1.47 for 50 *vs.* 100 ng/μL), respectively. Nonetheless, dsDNA had an overwhelming surpriority in on-target integration, which was more than 20 times greater than that in the 2H2OP system (OR = 26.70) (Table S3 in Appendix 1). Taken together, the dsDNA system accompanied by a dosage of 50 ng/μL of donors tends to yield the highest on-target KI efficiency in our current study.

Given the non-*As*-*Cath*-integrated fish, we did detect individuals with only the *lh* mutation. Specifically, 5.56% (3/54), 6.67% (4/60), 3.33% (2/60), and 3.33% (2/60) of fish with *lh* deficiency in the 2H2OP50, 2H2OP100, dsDNA50 and dsDNA100 groups, respectively, were detected by the Surveyor mutation test (Table S2 in Appendix 1). The sequencing results revealed that 2, 2, 1 and 3 types of mutations in 4 *lh*-mutant individuals from the 2H2OP100 group (Figure S5 in Appendix 2).

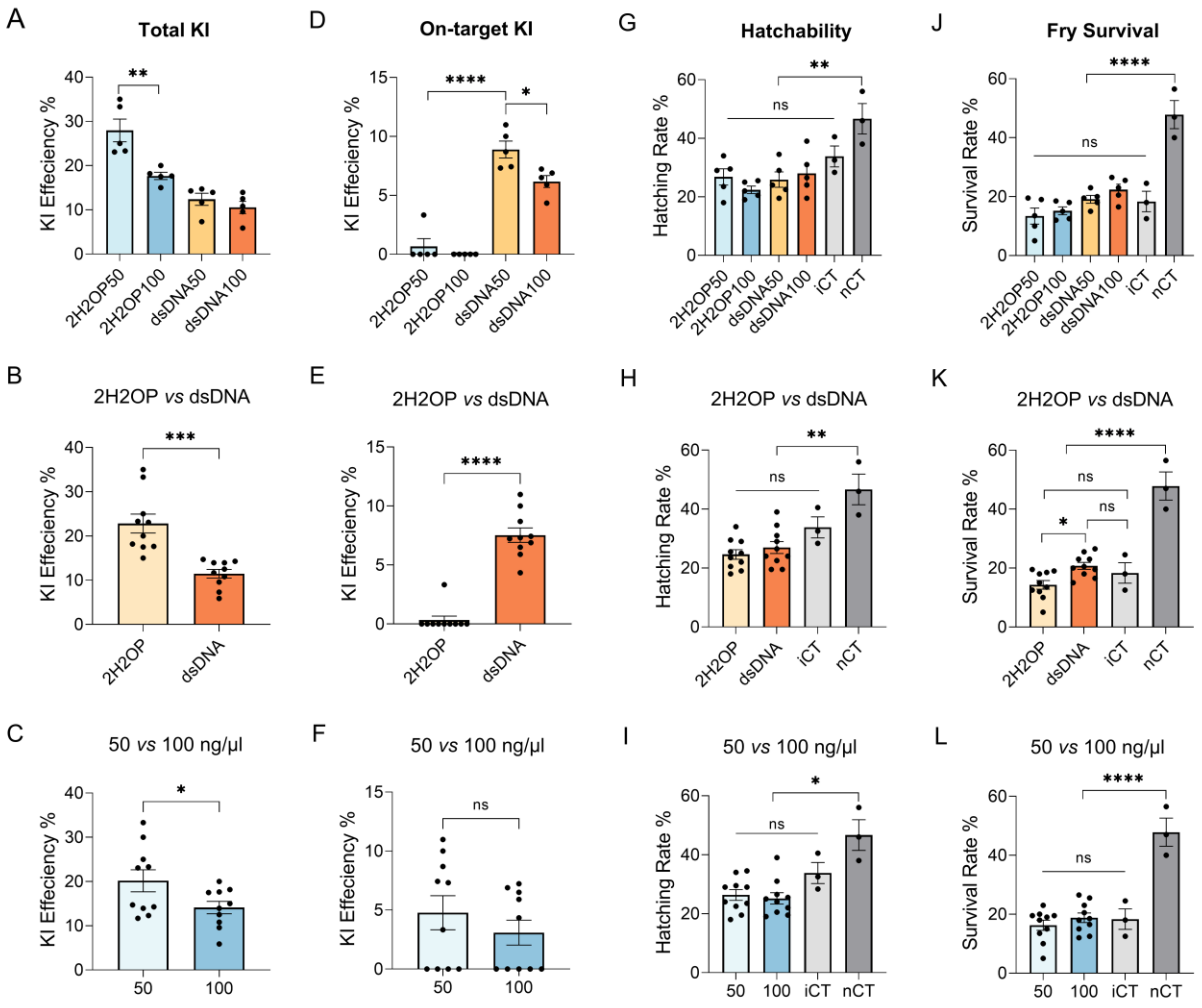


Figure 5. Effects of different CRISPR/Cas9-mediated systems (2H2OP *vs* dsDNA) with various dosages of donors (50 *vs* 100 ng/μL) on the knock-in (KI) efficiency, hatchability and fry survival rate. (A) Total KI efficiency of different CRISPR/Cas9-mediated systems and dosage combinations. **(B, C)** Comparison of total KI efficiency for different systems or dosages of donors. **(D)** On-target KI efficiency of different CRISPR/Cas9-mediated systems and dosage combinations. **(E, F)** Comparison of on-target KI efficiency of different systems or dosages. **(G)** Effect of different CRISPR/Cas9-mediated systems and dosage combinations on hatchability. **(H, I)** Comparison of the hatchability for different systems or dosages. **(J)** Effect of different CRISPR/Cas9-mediated systems and dosage combinations on fry survival. **(K, L)** Comparison of the fry survival rate for different systems or dosages. iCT, sham-injected control; nCT, non-injected control; 2H2OP(50/100), the CRISPR/Ca9-medicated system with ssODN1_As-Cath_ssODN2 construct (with a pUC57_mini plasmid and ssODN donor as 50/100 ng/μL); dsDNA(50/100), the CRISPR/Ca9-medicated system with HA1_As-Cath_HA2 donor DNA (with a dsDNA donor as 50/100 ng/μL); * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$; **** = $P < 0.0001$; ns = not significant, by unpaired student's *t*-test or one-way ANOVA.

3.2 Effects of the dosage and CRISPR/Cas9 system

Different donor dosages and CRISPR/Cas9-mediated systems exhibited toxicity to fish embryos by decreasing the hatchability and fry survival rate. Although there were no significant differences in hatching rates among these four CRISPR/Cas9-mediated injected groups compared to the iCT group ($P = 0.1630$), the hatching rate was lower than the nCT group ($P < 0.01$) (Figure 5G). Moreover, the lethality of embryos was consistent across different donor dosages (50 *vs.* 100 ng/μL) ($P = 0.1080$) or CRISPR/Cas9-mediated systems (2H2OP *vs.* dsDNA) ($P = 0.0796$), which was significantly higher than that in the nCT group (Figure 5H). For the fry survival, the survival rate of the microinjection group was significantly lower compared with the nCT group ($P < 0.0001$) (Figure 5J). In addition, the dsDNA system induced a higher survival rate of fry ($P = 0.0031$) (Figure 5K) than the 2H2OP system. Still, donor dosages showed no significant differences in fry survival after hatching ($P = 0.2923$) (Figure 5L).

3.3 Mosaicism and *As-Cath* expression

PCR and RT-PCR were used to detect the *As-Cath* transgene and its expression of different tissues in on-target positive fish. The results revealed that three of the five LH^{-}_As-$Cath^{+}$ fish showed the expression of the *As-Cath* in all 14 sampled tissues (skin, liver, kidney, spleen, blood, intestine, gill, stomach, fin, barbel, muscle, eye, brain and gonad) (Figure 6AB), but one of them had expression observed in 11 tissues (except barbel, muscle and gill) and another one in 8 tissues (skin, liver, blood, intestine, gill, barbel, muscle and gonad) (Figure S6 in Appendix 2), suggesting mosaicism in the on-target positive individuals. We found that the expression of *As-Cath* was detected even without pathogenic infections for the three on-target positive individuals. The three highest mRNA levels were determined in the kidney (28.91 fold changed), skin (24.30 fold), and gill (8.445 fold), followed by the muscle (7.430 fold), spleen (6.047 fold) and barbel (4.808 fold). However, the eye (1.327 fold), intestine (1.589 fold), and fin (1.608 fold) had the lowest expression compared to other tissues (Figure 6C).

In addition, compared to the WT individuals, the mRNA level of *lh* in gonads was down-regulated in LH^{-}_As-$Cath^{+}$ females at the age of one year ($P = 0.0016$), but there was no significant difference in that of males ($P = 0.5817$) (Figure 6D).

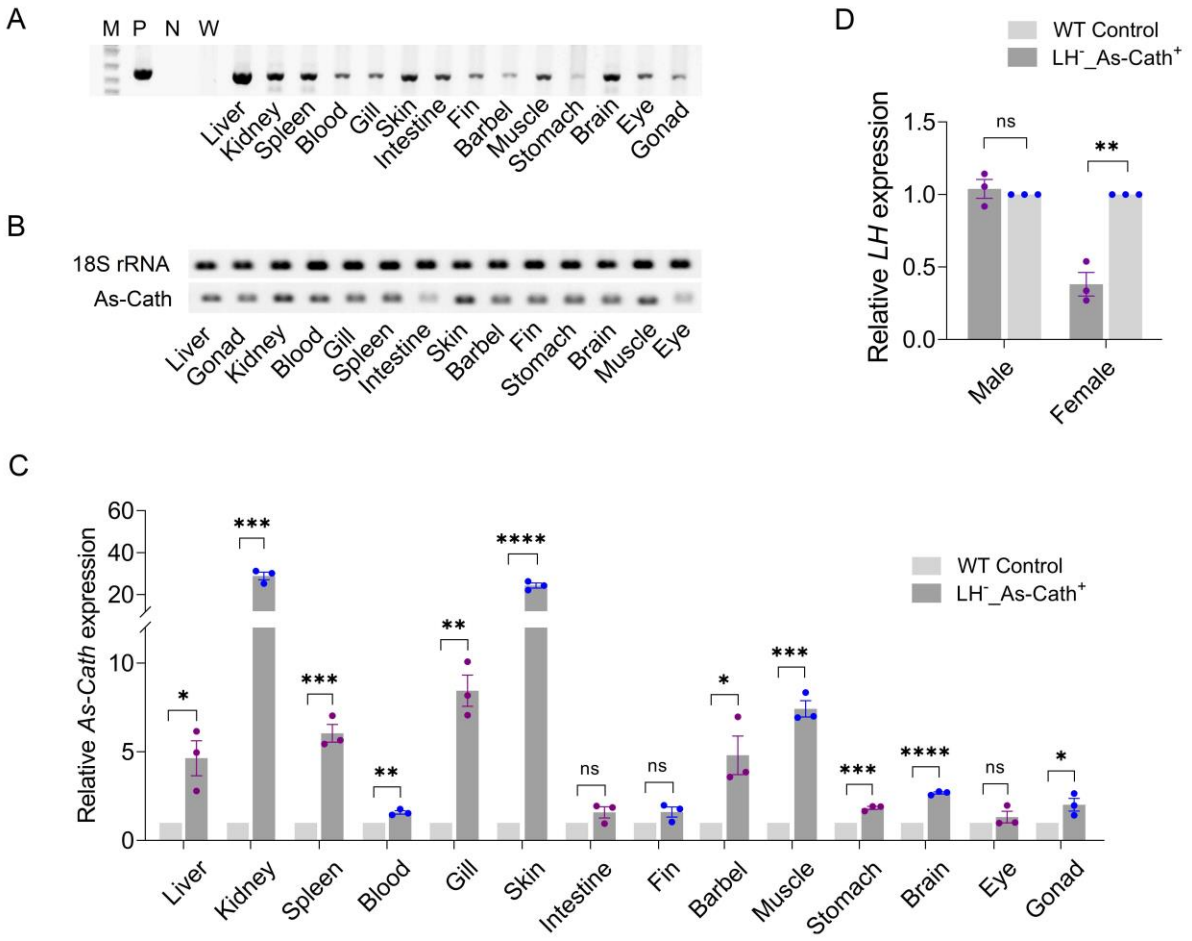


Figure 6. Mosaicism detection and the expression of the cathelicidin gene from *Alligator sinensis* (*As-Cath*) in the LH^{-}_As-Cath^{+} fish line. (A) PCR amplicons show the *As-Cath* region in 14 tissues from one representative LH^{-}_As-Cath^{+} fish. **(B)** The agarose gel electrophoresis showed the *As-Cath* gene expression in various tissues of P_0 transgenic channel catfish, *Ictalurus punctatus*. **(C)** Relative *As-Cath* gene expression of different tissues from RT-PCR analyses. **(D)** Relative *lh* gene expression of gonads from LH^{-}_As-Cath^{+} males and females. Expression levels were calibrated against corresponding tissues from sibling wild-type fish, and three individuals were employed for each genotype. Lane M, DNA marker (1 kb); Lane P, positive (plasmid or dsDNA donor) control; Lane N, water negative control; Lane W, wild-type control (nCT); * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$; **** = $P < 0.0001$; ns = not significant, by unpaired student's *t*-test or one-way ANOVA.

3.4 Reproductive sterility and restoration of reproduction

A three-round mating experiment determined the promise for complete control of channel catfish reproduction (Figure 7A). Our outcomes revealed that three pairs of WT (100%, 7927 eggs/BW) and two pairs of LH^{+}_As-Cath^{+} fish (50%, 8952 eggs/BW) were spawned respectively during the first two-week natural mating, but no spawn was observed in the LH^{-}_As-Cath^{+} pairs (0%).

Compared to the LH⁻_As-Cath⁺ pairs, WT and LH⁺_As-Cath⁺ fish had higher spawnability under natural pairing conditions (P = 0.0148 and P = 0.1743). In addition, the LH⁺_As-Cath⁺ pairs did not show a significant difference in spawnability compared to the WT pairs (P = 0.2143) (Figure 7B).

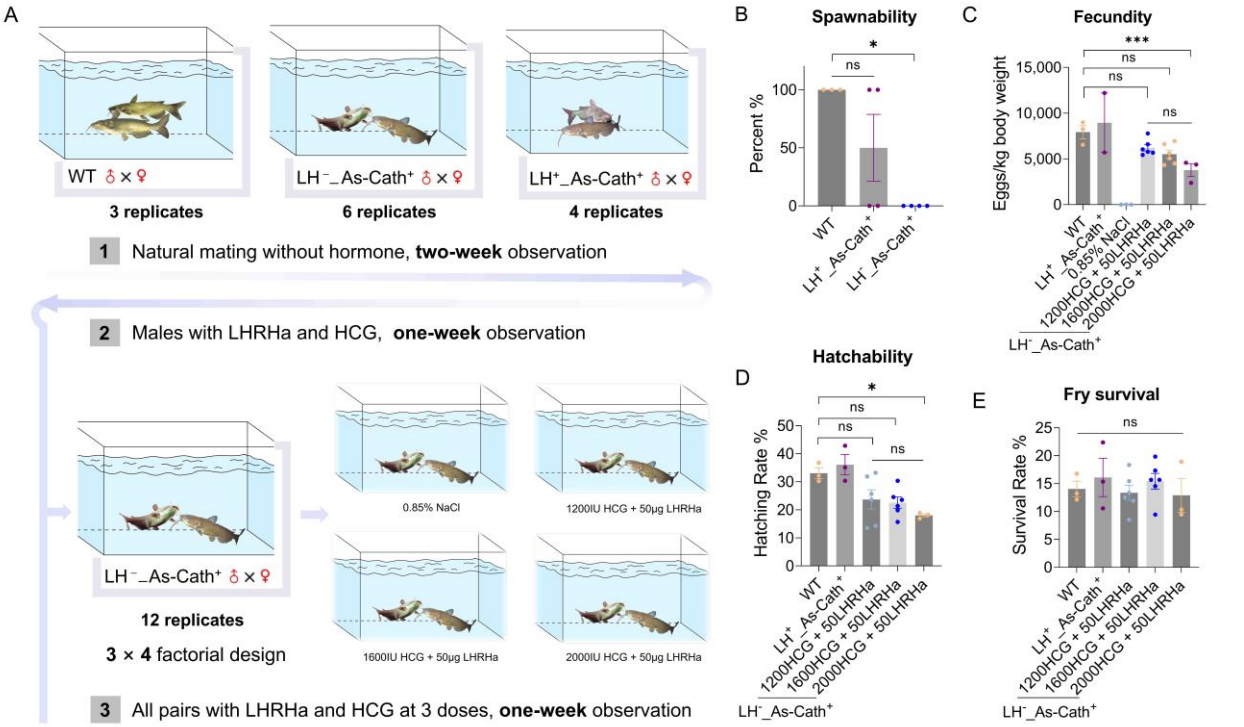


Figure 7. Reproductive determination and restoration of the *As-Cath*-integrated fish lines. (A) A three-round design of the reproduction experiment. Three genotypes of P_0 founders: WT, LH⁻_As-Cath⁺, and LH⁺_As-Cath⁺ fish were involved. First round, 3, 6 and 4 pairs as replicates for each genotype were set up randomly in 13 tanks for mating without hormone treatments, and a two-week observation was adopted. Second round, moved out spawned pairs and primed un-mated males with a 50 μg/kg LHRHa implant and 1600 IU/kg HCG to determine the reproduction of LH⁻_As-Cath⁺ females, observing for one week. Third round, 12 pairs of LH⁻_As-Cath⁺ fish were complemented and re-paired and treated with three doses of LHRHa and HCG in a 3 × 4 factorial design for one week. **(B)** Detection of spawnability for LH⁻_As-Cath⁺ fish during natural mating. **(C, D, E)** Potential effects of different hormone treatments on the fecundity and hatchability of P_0 generation, and fry survival of F_1 generation. LH, luteinizing hormone; LHRHa, luteinizing hormone-releasing hormone analogue; HCG, human chorionic gonadotropin; * = $P < 0.05$; ** = $P < 0.01$; ns = not significant, by unpaired student's *t*-test or one-way ANOVA.

Furthermore, a one-week hormone priming (50 μg/kg LHRHa + 1600 IU/kg HCG) of the males did not stimulate LH⁻_As-Cath⁺ females to give eggs, indicating LH-deficient females blocked oocyte maturation and ovulation. However, our results discovered that a combination of LHRHa

and HCG can effectively induce spawning for the LH^{-}_As-$Cath^{+}$ females when both males and females were primed. Specifically, two, two and one female gave eggs after 24 to 48 hours post-hormone injection from the 1200 IU (6213 eggs/BW), 1600 IU (5514 eggs/BW) and 2000 IU/kg (3778 eggs/BW) HCG group combined with 50 μg/kg LHRHa, respectively. These three treatments significantly improved the fecundity compared to 0.85 % NaCl injection ($P < 0.0001$). Additionally, the fecundity decreased with increasing hormone dosage, but the difference among these three hormone dosages was not significant ($P = 0.0731$). Nevertheless, the fecundity can be restored to a normal level when 1200 ($P = 0.2627$) or 1600 ($P = 0.1983$) IU/kg HCG combined with 50 μg/kg LHRHa was adopted (Figure 7C). Compared with the WT and the other hormonal-therapy groups, the 2000 IU/kg HCG group significantly reduced the fecundity (3778 eggs/BW, $P = 0.0494$) and hatchability (18.01%, $P = 0.0476$) (Figure 7D). Although different hormonal treatments had varying effects on fecundity and hatchability, they had no effects on fry survival at the early stage ($P = 0.1018$) (Figure 7E).

3.5 Growth comparison in P_0 and F_1

As mentioned above, three WT, two LH^{+}_As-$Cath^{+}$, and five LH^{-}_As-$Cath^{+}$ families were generated from our three-round mating experiment. However, genotype analysis determined that only one family in the LH^{+}_As-$Cath^{+}$ line (33.33% [10/30] integrated rate in the F_1 offspring) and two families in the LH^{-}_As-$Cath^{+}$ line (40% [12/30] integrated rate in the F_1 progeny of family 1 and 46.67% [14/30] integrated rate in the F_1 offspring of family 2), respectively, had the *As-Cath* gene detectable in the F_1 generation. These results further confirmed the existence of the mosaic phenomenon in the P_0 founders.

To determine the effects of *lh* disruption and *As-Cath* integration on fish growth, we compared the BW over time of the P_0 founders and the F_1 progeny, respectively. The growth data suggested that the LH^{-}_As-$Cath^{+}$ individuals did not show superiority in terms of growth in the first nine months in the P_0 generation. Nonetheless, P_0 LH^{-}_As-$Cath^{+}$ fish exhibited the largest body gain compared to other genotypes. Furthermore, significantly faster growth was demonstrated in the F_1 generation of LH^{-}_As-$Cath^{+}$ after a three-month culture. Hence, our results indicated more immediate growth potential for the LH^{-}_As-$Cath^{+}$ fish than the WT fish (Table 3).

Table 3. Mean monthly body weight (BW), sample size (N) over time of P_0 and F_1 *As-Cath*-integrated, negative and control channel catfish, *Ictalurus punctatus*. P_0 founders were generated in June 2020, and F_1 progeny were produced in June 2022. For both generations, four genotypes: WT, LH^+_As-$Cath^-$, LH^-_As-$Cath^+$, and LH^+_As-$Cath^+$ were included. Fish were kept separately in 60-L aquaria with the density of 2 fry/L until 4 months post hatch, then they were pit-tagged (10/2/2020) and transferred to a 1,200-L circular tank (~800-L water) with a mix of these 4 genotypes (initial number of fish was 30, 30, 28 and 32) and fed daily to satiation. Differences in BW among these four genotypes were compared using one-way ANOVA followed by Tukey's multiple comparisons test. Means with different letters as superscripts are significantly different ($P < 0.05$).

	Genotype	Mean body weight (g) of fish at different ages (Mean ± SEM)									
		10/2/2020		11/14/2020		12/14/2020		1/25/2021		3/6/2021	
		BW	N	BW	N	BW	N	BW	N	BW	N
P_0	WT	27.20 ± 1.77^a	60	37.15 ± 2.83^a	30	42.45 ± 3.08^{ab}	30	36.75 ± 2.31^a	30	50.75 ± 3.58^a	27
	LH^+_As-$Cath^-$	26.30 ± 2.24^a	60	36.40 ± 2.14^a	30	38.30 ± 3.20^a	29	35.25 ± 3.18^a	29	51.10 ± 2.28^a	29
	LH^-_As-$Cath^+$	23.10 ± 1.72^a	41	41.30 ± 2.60^a	28	49.65 ± 2.35^b	21	43.20 ± 2.75^a	20	58.45 ± 4.21^a	20
	LH^+_As-$Cath^+$	27.75 ± 2.39^a	63	39.95 ± 2.73^a	32	47.25 ± 3.26^{ab}	33	34.50 ± 3.58^a	33	50.85 ± 2.89^a	33
		8/9/2022		9/11/2022		10/12/2022					
		BW	N	BW	N	BW	N				
F_1	WT	2.63 ± 0.16^a	60	15.13 ± 1.00^a	54	22.90 ± 1.23^a	54				
	LH^+_As-$Cath^-$	2.60 ± 0.16^a	60	14.67 ± 0.91^a	56	21.30 ± 1.03^a	54				
	LH^-_As-$Cath^+$	3.03 ± 0.14^a	60	19.57 ± 1.31^b	59	26.03 ± 1.32^b	57				
	LH^+_As-$Cath^+$	2.70 ± 0.12^a	60	13.14 ± 1.05^a	58	22.13 ± 1.09^a	58				

Note: WT, wild-type fish without injection; LH^+_As-$Cath^-$, negative fish without the *As-Cath* insertion or *lh* gene mutation; LH^-_As-$Cath^+$, on-target positive fish with the integration of the *As-Cath* gene at the *lh* locus; LH^+_As-$Cath^+$, off-target positive fish with the *As-Cath* insertion but no *lh* mutation.

3.6 Enhanced resistance against fish pathogens

Enhanced resistance against *F. covae* and *E. ictaluri* of *As-Cath*-integrated fish was observed compared to WT/negative individuals from our challenge experiments in both P_0 and F_1 generations. According to *F. covae* challenge results, there was no significant difference in survival rate between the two types of controls (WT and LH^+_As-$Cath^-$) in both P_0 (13.33% *vs.* 20%, $P = 0.8682$) and F_1 generation (26.67% *vs.* 40%, $P = 0.8955$). However, LH^-_As-$Cath^+$ and LH^+_As-$Cath^+$ fish exhibited significantly improved survival post *F. covae* infection compared to the WT control group in both P_0 founders (LH^-_As-$Cath^+$ *vs.* WT: 73.33% *vs* 13.33%, $P = 0.0016$; LH^+_As-$Cath^+$ *vs.* WT: 66.67% *vs.* 13.33%, $P = 0.0014$) and F_1 progeny (LH^-_As-$Cath^+$ *vs.* WT: 86.67% *vs.* 26.67%, $P = 0.0010$; LH^+_As-$Cath^+$ *vs.* WT: 73.33% *vs.* 26.67%, $P = 0.0127$). Additionally, on-target insertion of the *As-Cath* gene resulted in improved resistance against *F. covae* than in the off-target positives without statistically differing in both generations (73.33% *vs.* 66.67%, $P = 0.7726$ for P_0, and 86.67% *vs.* 73.33%, $P = 0.3613$ for F_1). Furthermore, our findings revealed that the F_1 progeny was more resistant to *F. covae* than its P_0 parents (Figure 8AB).

Increased resistance to *E. ictaluri* was also observed in the P_0 (LH^-_As-$Cath^+$ *vs.* WT: 73.33% *vs.* 33.33%, $P = 0.0125$; LH^+_As-$Cath^+$ *vs.* WT: 60% *vs* 33.33%, $P = 0.0427$) and F_1 generations (LH^-_As-$Cath^+$ *vs.* WT: 66.67% *vs.* 40%, $P = 0.0558$; LH^+_As-$Cath^+$ *vs.* WT: 73.33% *vs.* 40%, $P = 0.0350$), with results that were similar to those of the *F. covae* challenge. Overall, *As-Cath*-integrated individuals showed a significant improvement in the survival rate compared to the WT fish (66.67% *vs.* 33.33%, $P = 0.0381$ for P_0; 70% *vs.* 40%, $P = 0.0335$ for F_1). Nevertheless, there was no significant difference in LH^-_As-$Cath^+$ and LH^+_As-$Cath^+$ fish (73.33% *vs.* 60%, $P = 0.4566$ for P_0; 66.67% *vs.* 73.33%, $P = 0.6851$ for F_1) (Figure 8CD).

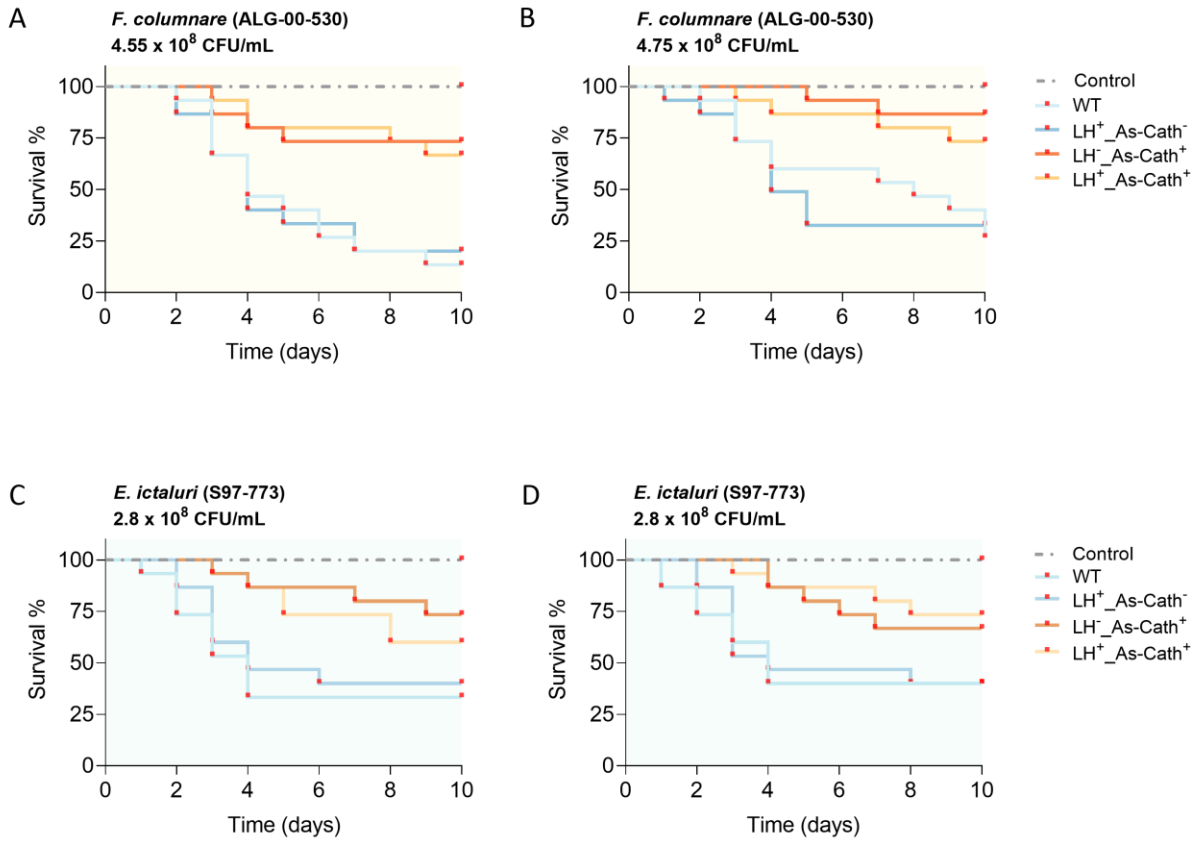


Figure 8. Kaplan-Meier plots of *As-Cath* integrated channel catfish against two fish bacterial pathogens. (A, B) Survival curves of P_0 and F_1 generations for a variety of genotypes infected by *Flavobacterium covae*, respectively. **(C, D)** Survival curves of P_0 founders and F_1 progeny for different genotypes infected by *Edwardsiella ictaluri*, respectively. In addition to these bacterial infection groups, one control group with medium immersion was implemented for each challenge experiment, and the immersion dose was presented in each figure. Comparison of different survival curves was determined by the Log-rank (Mantel-Cox) test. WT, wild-type, non-injected fish line; LH^+_As-$Cath^-$, negative fish line (micro-injected fish without *lh* mutation and *As-Cath* insertion); LH^-_As-$Cath^+$, on-target positive fish (*As-Cath* insertion was detected at *lh* locus); LH^+_As-$Cath^+$, off-target positive fish (*As-Cath* insertion was detected but not at *lh* locus).

4. Discussion

In contrast to the previous gene-editing oriented exclusively to the improvement of the desired traits, the present study took into account ways to lessen the potential impact of transgenic fish on the ecosystems and genetic biodiversity. Specifically, we successfully integrated an AMG into the reproduction-associated locus using different CRISPR/Cas9-mediated systems. We identified a suitable KI system for channel catfish to achieve boosted resistance against fish pathogens and

reproductive control, reducing the reliance on antibiotics and anti-parasitics in aquaculture. The HA-mediated CRISPR/Cas9 system displayed a high integrated rate, low off-target events, and low toxicity. In addition, reproduction is entirely controllable and can only be restored to normal levels of fecundity with hormone therapy in the new fish line. In general, the insertion of the cathelicidin gene at the *lh* locus for enhanced resistance against infectious diseases and reproductive confinement to improve consumer-valued qualities and to promote the environmental friendliness of transgenic fish appears promising.

There have been several obstacles involved in the CRISPR/Cas9-mediated KI system when it is used in the embryos of non-model animals. In the history of genome editing, the initial CRISPR/Cas9 systems were proposed based on mammalian cells or embryos of the model animals. From model to non-model animals, there are several uncertainties, such as embryo size, developmental period, and the sensitivity to Cas9 protein that researchers have to optimize a fitted system when starting a new species' genome editing. Yoshimi et al. (2016) demonstrated that the ssODN-mediated end joining approach induced a high integrated rate of 17.6% (3/17) in rats when a short ssODN template was provided. Conversely, recent works indicated that ssODN-mediated KI could induce a high percentage (17.8%) of indel mutations in sheep (Menchaca et al., 2020). In the current study, we used CRISPR/Cas9 systems mediated by ssODN and HA to create on-target KIs of the *As-Cath* gene at the *lh* locus. Although a high KI efficiency of 22.38% (64/286) was detected in the ssODN-mediated system, it caused a high off-target frequency (> 90%) in the channel catfish. Our results are in agreement with findings in zebrafish, which have illustrated that erroneous ssODN integration occurred when various template lengths were adopted (Boel et al., 2018). These studies suggest that ssODN-mediated KI efficiency in fish models relies heavily on ssODN templates (Kan et al., 2017), and caution is warranted when employing ssODNs to create KI models.

Compared to the ssODN-mediated system, HA-assisted KI can achieve a 20–30% HDR-mediated knockin in human cells with various homogenous sequences (Byrne et al., 2015; Zhang et al., 2017). In addition, Simora et al. (2020) determined that HA-mediated CRISPR/Cas9 provided with a linear dsDNA donor displayed a total integrated rate of 29% at the non-coding region of channel catfish genome, which is drastically higher than that of this work (29% *vs.* 11.16% [51/457]). We believe this difference in integration rate is due to the different sample

sizes, unknown functions in the target regions (non-coding *vs. lh* locus), efficiency of sgRNA and HA, and unpredictable genetic interaction; the larger sample size from our study could give more robust conclusions. These findings reveal that the HA-mediated system is more effective in the catfish species compared to the ssODN. The KI efficiency of HDR-induced CRISPR/Cas9 has been at a low level including in cell lines and model animals (Yoshimi et al., 2016; Zhang et al., 2017; Boel et al., 2018). Fortunately, new CRISPR/Cas-mediated techniques are constantly being developed. For instance, the CRISPR/Cas12i-mediated system shows promise in multiplexed genome editing with high mutation rates in human T cells (McGaw et al., 2022). Additionally, Kelly et al. (2021) established a CRISPR/Cas9 HITI system for the insertion of large DNA donors with high integrated efficiency of 36% in human 293T cells. Recently, a new approach named dCas9-SSAP demonstrated a high on-target KI efficiency (~20%) knocking in long sequences in mammalian cells (Wang et al., 2022). These new tools or systems are encouraging to be applied from model to non-model animals and could improve genome-editing efficiency.

Although we predicted and avoided possible off-target sites using the well-acknowledged software, the actual integration results showed the existence of off-target activities. This is mainly due to the failure of *in silico* prediction to predict *bona-fide* off-target sites *in vivo* (Ran et al., 2013; Heigwer et al., 2014). Furthermore, the frequency of off-target events is higher *in vivo* of animal experiments than that of in cellular experiments *in vitro* (Zhang et al., 2015). The majority of published studies contend that the observed unintended mutations/insertions is one major concern in the application of the CRISPR/Cas9 system, which could confound the interpretation of findings (Pattanayak et al., 2013; Cho et al., 2014; Heigwer et al., 2014). However, although some reports claim that no detectable undesirable mutations/insertions from the genotypes or phenotypes have been revealed in mice and fish (Shen et al., 2013; Iyer et al., 2015; Simora et al., 2020), the following underlying potentials could be noted: 1) Unaltered phenotypes may be observed since the off-target cleavage can occur in a non-coding region (Wang et al., 2014). 2) The researchers tend to focus on the P_0 founders with intended insertions rather than those harboring possible off-target mutations (Li et al., 2013ab). 3) Most published research using animal models does not use genome-wide methodologies for detecting off-target cases, which could conceal some infrequent off-target editing sites (Zhang et al., 2015). In the same case, with the exception of *lh* mutations, we did not conduct a thorough detection on all off-target individuals due to its being time-consuming and expensive. Nevertheless, this does not

preclude us from keeping the non-analyzed off-target individuals as we will eventually genotype them in a genome-wide and unbiased way.

Genetic mosaicisms have been and will still be another obstacle to applying CRISPR/Cas9-mediated genome editing in practical applications. In this study, we failed to effectively obtain 100% of individuals without mosaics. In essence, mosaicism from CRISPR/Cas9-genome-edited organisms is common in the case of fertilized egg-based editing, and mosaic animals have been observed in mice (Oliver et al., 2015; Raveux et al., 2017), rats (Li et al., 2013b) and zebrafish (Jao et al., 2013; Auer et al., 2014) with a variety of frequencies. CRISPR/Cas9 engineered mosaicisms bring undesired consequences, hindering the generation of homozygous positive offspring and prolonging the generation of homozygotes. We evaluated the *As-Cath* gene expression from five on-target positive P_0 founders and found that one individual had no expression in the gonad. In our study, several mosaic events were determined in the germline, resulting in the inability to transfer the *As-Cath* gene to the offspring. Thus, we believe that mosaicism is also common and unavoidable in non-model fish. Although early sperm/testis or egg/ovary genotyping can be effective in avoiding the creation of undesirable offspring, it is challenging to access the germline DNA without sacrificing the parents. Of importance, we still maintain our mosaic populations for genotyping and phenotyping in the further F_2 and F_3 progeny until homozygous individuals are obtained. Future research could reduce mosaicisms by delivering CRISPR/Cas9 components to very early-stage zygotes (Mehravara et al., 2019). Alternatively, the new strategies, i.e., *Easi*-CRISPR, C-CRISPR (Mehravara et al., 2019), CRISPR/Cas9 HITI (Kelly et al., 2021) and dCas9-SSAP (Wang et al., 2022) could be used to prevent the induction of mosaic animals.

Regardless of the type of CRISPR/Cas9-mediated genome editing, microinjection always has irreversible effects on embryos, i.e., increased mortality and decreased hatchability from our current study. High embryonic deaths were observed from shame- and CRISPR/Cas9-mediated-microinjection in our study, revealing that major mortality occurs due to the injection of the yolk, while fewer impacts are from the DNA donors and reagents (Simora et al., 2020). Although a high dosage resulted in a high embryonic mortality and lower hatching rate, it did not significantly reduce the fry survival rate compared to the injected-control group, which is in agreement with the findings from Elaswad et al. (2018). This may be because microinjection

only has a detrimental effect on the yolk of the embryo. Still, this effect no longer affects the fry once the fertilized eggs have successfully hatched. Given the unavoidable physical lethality of embryos, off-target effects and mosaicisms, we recommend microinjection of ~3000 fertilized eggs for non-model fish species in order to afford enough gene-edited fish for subsequent validation experiments.

To assess the pleiotropic effects, we compared the growth performance of the on-target/off-target *As-Cath*-integrated fish line with the WT population. Our findings demonstrated that off-target insertions did not exhibit growth depression or improvement in various families of P_0 founders. Nonetheless, the preliminary data revealed the LH^{-}_As-$Cath^{+}$ fish had a greater gain in body weight compared to the WT individuals after a three-month culture in the tank, indicating that the growth differences are emerging in the F_1 progeny. This variation may be due to heterozygous individuals lacking stable genetic traits, or off-target integrations in other regions concealing growth advantages in the P_0 generation (Zhang et al., 2015). cfGnRH-deficient channel catfish did not show significant effects in growth and survival throughout a four-year culture compared to the WT fish (Qin et al., 2022). However, potential pleiotropic effects could exist when the *lh* gene is replaced by the *As-Cath* in our cases. Therefore, P_0 mosaic founders carrying the *As-Cath* gene should be used to produce F_1, F_2 and F_3 homozygous families, and then the comparisons of the growth, survival rate, seinability and carcass traits could be performed to avail the enhanced performance of LH^{-}_As-$Cath^{+}$ fish line more transparent to farmers and the public in the future.

HDR-mediated KI is rarely applied in aquaculture due to the very low integration efficiency, but most of the traits were achieved by NHEJ-mediated KO (Blix et al., 2021; Yang et al., 2021). In addition, few studies have proved that gene-mutants can induce disease-resistant fish lines via KO to date (Wang et al., 2022). By contrast, the integration of AMG is encouraging to improve resistance against pathogens in fish (Dunham and Su, 2020; Wang et al., 2022). However, consumers generally have relatively little awareness of transgenesis and have more negative attitudes toward genetically modified organisms than genome-edited organisms (Hallerman et al., 2022), hence the public pushback against transgenic/gene-edited animals is hindering them from reaching the market. Here, we reasonably contend that cathelicidin transgenic catfish would not pose a threat to food safety since 1) Meat from artificially grown alligators is edible even when

consumed raw, and the gut will digest most proteins and inactivate them. 2) Eventually, amino acids rather than proteins are absorbed by humans. 3) Even though the gene sequence is ever-changing in various beings, there are only 20 different types of encoded amino acids that are frequently consumed by humans. In this vein, we are raising attention to potential benefits and risks of our *As-Cath* transgenic catfish by making them transparent to the public.

Nonetheless, scientists and breeders need to be aware of the possible damage that genetically modified fish could cause to the environment and ecosystem (Dunham and Su, 2020). On the one hand, reproductive sterility via genome editing has been attracting the attention of researchers and offering opportunities to reduce environmental risks in aquaculture (Hallerman et al., 2022). On the other hand, representative examples have illustrated that reproductive confinement is promising in model and cultured fish by knocking out/disrupting gonadal development-related genes (Su et al., 2014; Su et al., 2015; Qin et al., 2016; Wargelius et al., 2016; Gay et al., 2018). Recently, Qin et al. (2022) demonstrated that the reproduction-blocked channel catfish are sterile, and this reproductive confinement can be lifted through hormone therapy with LHRHa. In this study, the dose of 1600 IU/kg HCG coupled with 50 μg/kg LHRHa can restore fecundity at the highest level in comparison to other hormone treatments, but this improvement is not significant from that of 1200 IU/kg HCG. Therefore, a low dose of 1200 IU/kg HCG is recommended for hormone therapy to restore the reproduction of the sterile fish line to reduce costs. In addition to genetically achieving reproductive sterility, well-confined culture systems should be adopted to avoid the escape of mutant/transgenic individuals, especially in the experimental phase of transgenic fish.

In summary, we established a sterile catfish line that confers enhanced resistance to fish pathogens by expressing the cathelicidin protein. Our study has demonstrated that the insertion of the cathelicidin gene at the *lh* locus by harnessing the HA- or ssODN-mediated CRISPR/Cas9 system can be a robust approach to produce sterilized and environmentally-sound fish lines with enhanced disease resistance. Encouragingly, CRISPR/Cas9-mediated KI of AMGs at the reproduction-related loci coupled with hormone therapy could be applied in other commercial fish to increase profits and lower environmental dangers posed by escaped genetic-modified individuals. Notably, even though the desired traits (on-target insertions) can be quickly achieved through CRISPR/Cas9-mediated genome editing, this does not safeguard that we will be able to yield enough non-mosaic P_0 founders. We contend the genome-editing tool should be used as a

complement to existing breeding techniques, not a replacement for them. Hence, a combination of genome editing, and conventional selective breeding is required to maximize the benefits of CRISPR/Cas9 tools more effectively in aquatic applications and to hasten the breeding process. In conclusion, this study showed the potential of overexpressing a disease-resistant peptide inserted at a reproduction-related gene using CRISPR/Cas9 in channel catfish, which may provide a strategy of decreasing bacterial disease problems in catfish at the same time reducing environmental risks.

5. References

Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. *Cell*. 2006;124:783–801. doi:10.1016/j.cell.2006.02.015.

Armstrong JB, Malacinski GM. Developmental Biology of the Axolotl. New York: Oxford University Press; 1989.

Auer TO, Duroure K, De Cian A, Concordet JP, Del Bene F. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res*. 2014;24:142–53. doi:10.1101/gr.161638.113.

Bae S, Park J, Kim J-S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*. 2014;30:1473–75. doi:10.1093/bioinformatics/btu048.

Blix TB, Dalmo RA, Wargelius A, Myhr AI. Genome editing on finfish: Current status and implications for sustainability. *Rev Aquac*. 2021;13:2344–63. doi:10.1111/raq.12571.

Boel A, De Saffel H, Steyaert W, Callewaert B, De Paepe A, Coucke PJ, et al. CRISPR/Cas9-mediated homology-directed repair by ssODNs in zebrafish induces complex mutational patterns resulting from genomic integration of repair-template fragments. *Dis Model Mech*. 2018;11:dmm035352. doi:10.1242/dmm.035352.

Byrne SM, Ortiz L, Mali P, Aach J, Church GM. Multi-kilobase homozygous targeted gene replacement in human induced pluripotent stem cells. *Nucleic Acids Res*. 2015;43:e21. doi:10.1093/nar/gku1246.

Chen F, Pruett-Miller SM, Huang Y, Gjoka M, Duda K, Taunton J, et al. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat Methods*. 2011;8:753–5. doi:10.1038/nmeth.1653.

Chen Y, Cai S, Qiao X, Wu M, Guo Z, Wang R, et al. As-CATH1-6, novel cathelicidins with potent antimicrobial and immunomodulatory properties from *Alligator sinensis*, play pivotal

roles in host antimicrobial immune responses. *Biochem J*. 2017;474:2861–85. doi:10.1042/BCJ20170334.

Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*. 2014;24:132–41. doi:10.1101/gr.162339.113.

Chu L, Li J, Liu Y, Hu W, Cheng CHK. Targeted gene disruption in zebrafish reveals noncanonical functions of LH signaling in reproduction. *Mol Endocrinol*. 2014;28:1785–95. doi:10.1210/me.2014-1061.

Coogan M, Alston V, Su B, Khalil K, Elaswad A, Khan M, et al. CRISPR/Cas-9 induced knockout of myostatin gene improves growth and disease resistance in channel catfish (*Ictalurus punctatus*). *Aquaculture*. 2022;557:738290. doi:10.1016/j.aquaculture.2022.738290.

Davis KB. Age at puberty of channel catfish, *Ictalurus punctatus*, controlled by thermoperiod. *Aquaculture*. 2009;292:244–47. doi:10.1016/j.aquaculture.2009.04.023.

Doudna JA, Charpentier E. The new frontier of genome engineering with CRISPR-Cas9. *Science*. 2014;346:1258096. doi:10.1126/science.1258096.

Dunham RA, Su B. Genetically Engineered Fish: Potential Impacts on Aquaculture, Biodiversity, and the Environment. In: Chaurasia A, Hawksworth DL, Pessoa de Miranda M, editors. GMOs: Implications for Biodiversity Conservation and Ecological Processes, Cham: Springer International Publishing; 2020, p. 241–75. doi:10.1007/978-3-030-53183-6_11.

Elaswad A, Khalil K, Ye Z, Liu Z, Liu S, Peatman E, et al. Effects of CRISPR/Cas9 dosage on TICAM1 and RBL gene mutation rate, embryonic development, hatchability and fry survival in channel catfish. *Sci Rep*. 2018;8:16499. doi:10.1038/s41598-018-34738-4.

Gay S, Bugeon J, Bouchareb A, Henry L, Delahaye C, Legeai F, et al. MiR-202 controls female fecundity by regulating medaka oogenesis. *PLoS Genet*. 2018;14:e1007593. doi:10.1371/journal.pgen.1007593.

Grier HJ. Cellular organization of the testis and spermatogenesis in fishes. *Am Zool*. 1981;21:345–57. doi:10.1093/icb/21.2.345.

Hallerman EM, Dunham R, Houston RD, Walton M, Wargelius A, Wray-Cahen D. Towards production of genome-edited aquaculture species. *Rev Aquac*. 2022;1–5. doi:10.1111/raq.12739.

Heigwer F, Kerr G, Boutros M. E-CRISP: fast CRISPR target site identification. *Nat Methods*. 2014;11:122–23. doi:10.1038/nmeth.2812.

Hilchie AL, Wuerth K, Hancock REW. Immune modulation by multifaceted cationic host defense (antimicrobial) peptides. *Nat Chem Biol*. 2013;9:761–68. doi:10.1038/nchembio.1393.

Hisano Y, Sakuma T, Nakade S, Ohga R, Ota S, Okamoto H, et al. Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish. *Sci Rep*. 2015;5:8841. doi:10.1038/srep08841.

Iyer V, Shen B, Zhang W, Hodgkins A, Keane T, Huang X, et al. Off-target mutations are rare in Cas9-modified mice. *Nat Methods*. 2015;12:479. doi:10.1038/nmeth.3408.

Jao L-E, Wente SR, Chen W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc Natl Acad Sci* USA. 2013;110:13904–909. doi:10.1073/pnas.1308335110.

Kan Y, Ruis B, Takasugi T, Hendrickson EA. Mechanisms of precise genome editing using oligonucleotide donors. *Genome Res*. 2017;27:1099–111. doi:10.1101/gr.214775.116.

Kelly JJ, Saee-Marand M, Nyström NN, Evans MM, Chen Y, Martinez FM, et al. Safe harbor-targeted CRISPR-Cas9 homology-independent targeted integration for multimodality reporter gene-based cell tracking. *Sci Adv*. 2021;7: eabc3791. doi:10.1126/sciadv.abc3791

Khalil K, Elayat M, Khalifa E, Daghash S, Elaswad A, Miller M, et al. Generation of myostatin gene-edited channel catfish (*Ictalurus punctatus*) via zygote injection of CRISPR/Cas9 system. *Sci Rep*. 2017;7:7301. doi:10.1038/s41598-017-07223-7.

Ledford H. Salmon approval heralds rethink of transgenic animals. *Nature*. 2015;527:417–18. doi:10.1038/527417a.

Li D, Qiu Z, Shao Y, Chen Y, Guan Y, Liu M, et al. Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat Biotechnol*. 2013;31:681–83. doi:10.1038/nbt.2661.

Li W, Teng F, Li T, Zhou Q. Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nat Biotechnol*. 2013;31:684. doi:10.1038/nbt.2652.

Liu Z, Liu S, Yao J, Bao L, Zhang J, Li Y, et al. The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts. *Nat Commun*. 2016;7:11757. doi:10.1038/ncomms11757.

McGaw C, Garrity AJ, Munoz GZ, Haswell JR, Sengupta S, Keston-Smith E, et al. Engineered Cas12i2 is a versatile high-efficiency platform for therapeutic genome editing. *Nat Commun*. 2022;13:2833 doi:10.1038/s41467-022-30465-7.

Mehravara M, Shirazia A, Nazaric M, Banand M. Mosaicism in CRISPR/Cas9-mediated genome editing. *Dev Biol*. 2019;445:156–62. doi:10.1016/j.ydbio.2018.10.008.

Menchaca A, Dos Santos-Neto PC, Souza-Neves M, Cuadro F, Mulet AP, Tesson L, et al. Otoferlin gene editing in sheep via CRISPR-assisted ssODN-mediated homology directed repair. *Sci Rep*. 2020;10:5995. doi:10.1038/s41598-020-62879-y.

Mookherjee N, Anderson MA, Haagsman HP, Davidson DJ. Antimicrobial host defence peptides: functions and clinical potential. *Nat Rev Drug Discov*. 2020;19:311–32. doi:10.1038/s41573-019-0058-8.

Mosimann C, Kaufman CK, Li P, Pugach EK, Tamplin OJ, Zon LI. Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish. *Development*. 2011;138:169–77. doi:10.1242/dev.059345.

Murakami Y, Ansai S, Yonemura A, Kinoshita M. An efficient system for homology-dependent targeted gene integration in medaka (*Oryzias latipes*). *Zoological Lett*. 2017;3:10. doi:10.1186/s40851-017-0071-x.

Oliver D, Yuan S, McSwiggin H, Yan W. Pervasive genotypic mosaicism in founder mice derived from genome editing through pronuclear injection. *PLOS One*. 2015;10: e0129457. doi:10.1371/journal.pone.0129457.

Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol*. 2013;31:839–43. doi:10.1038/nbt.2673.

Qin G, Qin Z, Lu C, Ye Z, Elaswad A, Bangs M, et al. Gene editing of the catfish gonadotropin-releasing hormone gene and hormone therapy to control the reproduction in channel catfish, *Ictalurus punctatus*. *Biology* (Basel). 2022;11:649. doi:10.3390/biology11050649.

Qin Z, Li Y, Su B, Cheng Q, Ye Z, Perera DA, et al. Editing of the luteinizing hormone gene to sterilize channel catfish, *Ictalurus punctatus*, using a modified zinc finger nuclease technology with electroporation. *Mar Biotechnol*. 2016;18:255–63. doi:10.1007/s10126-016-9687-7.

Qiu P, Shandilya H, D'Alessio JM, O'Connor K, Durocher J, Gerard GF. Mutation detection using Surveyor^{TM} nuclease. *Biotechniques*. 2004;36:702–7. doi:10.2144/04364PF01.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013;8:2281–308. doi:10.1038/nprot.2013.143.

Raveux A, Vandormael-Pournin S, Cohen-Tannoudji M. Optimization of the production of knock-in alleles by CRISPR/Cas9 microinjection into the mouse zygote. *Sci Rep*. 2017;7:42661. doi:10.1038/srep42661.

Shen B, Zhang J, Wu H, Wang J, Ma K, Li Z, et al. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res*. 2013;23:720–23. doi:10.1038/cr.2013.46

Simora RMC, Li S, Abass NY, Terhune JS, Dunham RA. Cathelicidins enhance protection of channel catfish, *Ictalurus punctatus*, and channel catfish ♀ × blue catfish, *Ictalurus furcatus* ♂ hybrid catfish against *Edwarsiella ictaluri* infection. *J Fish Dis*. 2020;43:1553–62. doi:10.1111/jfd.13257.

Simora RMC, Wang W, Coogan M, El Husseini N, Terhune JS, Dunham RA. Effectiveness of cathelicidin antimicrobial peptide against Ictalurid catfish bacterial pathogens. *J Aquat Anim Health*. 2021;33:178–89. doi:10.1002/aah.10131.

Simora RMC, Xing D, Bangs MR, Wang W, Ma X, Su B, et al. CRISPR/Cas9-mediated knock-in of alligator cathelicidin gene in a non-coding region of channel catfish genome. *Sci Rep*. 2020;10:22271. doi:10.1038/s41598-020-79409-5.

Storici F, Snipe JR, Chan GK, Gordenin DA, Resnick MA. Conservative repair of a chromosomal double-strand break by single-strand DNA through two steps of annealing. *Mol Cell Biol*. 2006;26:7645–57. doi:10.1128/mcb.00672-06.

Su B, Peatman E, Shang M, Thresher R, Grewe P, Patil JG, et al. Expression and knockdown of primordial germ cell genes, *vasa*, *nanos* and *dead end* in common carp (*Cyprinus carpio*) embryos for transgenic sterilization and reduced sexual maturity. *Aquaculture*. 2014;S72–S84:420–21. doi:10.1016/j.aquaculture.2013.07.008.

Su B, Shang M, Grewe PM, Patil JG, Peatman E, Perera DA, et al. Suppression and restoration of primordial germ cell marker gene expression in channel catfish, *Ictalurus punctatus*, using knockdown constructs regulated by copper transport protein gene promoters: Potential for reversible transgenic sterilization. *Theriogenology*. 2015;84:1499–512. doi:10.1016/j.theriogenology.2015.07.037.

Waltz E. First genetically engineered salmon sold in Canada. *Nature*. 2017;548:148–48. doi:10.1038/nature.2017.22116.

Wang C, Qu Y, Cheng JKW, Hughes NW, Zhang Q, Wang M, et al. dCas9-based gene editing for cleavage-free genomic knock-in of long sequences. *Nat Cell Biol*. 2022;24:268–78. doi:10.1038/s41556-021-00836-1.

Wang G, Li X, Wang Z. APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res*. 2016;44:D1087–93. doi:10.1093/nar/gkv1278.

Wang J, Su B, Dunham RA. Genome-wide identification of catfish antimicrobial peptides: A new perspective to enhance fish disease resistance. *Rev Aquac*. 2022;14:2002–22. doi:10.1111/raq.12684.

Wang J, Wilson AE, Su B, Dunham RA. Functionality of dietary antimicrobial peptides in aquatic animal health: Multiple meta-analyses. *Anim Nutr*. 2023;12:200–214. doi:10.1016/j.aninu.2022.10.001.

Wang T, Wei JJ, Sabatini DM, Lander, ES. Genetic screens in human cells using the CRISPR-Cas9 system. *Science*. 2014;343:80–84. doi:10.1126/science.1246981.

Wargelius A, Leininger S, Skaftnesmo KO, Kleppe L, Andersson E, Taranger GL, et al. *Dnd* knockout ablates germ cells and demonstrates germ cell independent sex differentiation in Atlantic salmon. *Sci Rep*. 2016;6:21284. doi:10.1038/srep21284.

Wefers B, Meyer M, Ortiz O, Hrabé de Angelis M, Hansen J, Wurst W, et al. Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. *Proc Natl Acad Sci* USA. 2013;110:3782–87. doi:10.1073/pnas.1218721110.

Xing D, Su B, Bangs M, Li S, Wang J, Bern L, et al. CRISPR/Cas9-mediate knock-in method can improve the expression and effect of transgene in P1 generation of channel catfish (*Ictalurus punctatus*). *Aquaculture*. 2022a;560:738531. doi:10.1016/j.aquaculture.2022.738531.

Xing D, Su B, Li S, Bangs M, Creamer D, Coogan M, et al. CRISPR/Cas9-mediated transgenesis of the Masu salmon (*Oncorhynchus masou*) *elovl2* gene improves n-3 fatty acid content in channel catfish (*Ictalurus punctatus*). *Mar Biotechnol*. 2022b;24:513–23. doi:10.1007/s10126-022-10110-6.

Yamaguchi Y, Nagata J, Nishimiya O, Kawasaki T, Hiramatsu N, Todo T. Molecular characterization of *fshb* and *lhb* subunits and their expression profiles in captive white-edged rockfish, *Sebastes taczanowskii*. *Comp Biochem Physiol A: Mol Integr Physiol*. 2021;261:111055. doi:10.1016/j.cbpa.2021.111055.

Yang Z, Yu Y, Tay YX, Yue GH. Genome editing and its applications in genetic improvement in aquaculture. *Rev Aquac*. 2022;14:178–91. doi:10.1111/raq.12591.

Yoshimi K, Kunihiro Y, Kaneko T, Nagahora H, Voigt B, Mashimo T. ssODN-mediated knock-in with CRISPR-Cas for large genomic regions in zygotes. *Nat Commun*. 2016;7:10431. doi:10.1038/ncomms10431.

Zhang J-P, Li X-L, Li G-H, Chen W, Arakaki C, Botimer GD, et al. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biology*. 2017;18. doi:10.1186/s13059-017-1164-8.

Zhang X-H, Tee LY, Wang X-G, Huang Q-S, Yang S-H. Off-target effects in CRISPR/Cas9-mediated genome engineering. *Mol Ther Nucleic Acids*. 2015;17:e264. doi:10.1038/mtna.2015.37.

CHAPTER THREE

Integration of alligator cathelicidin gene via two CRISPR/Cas9-assisted systems enhances bacterial resistance in blue catfish, *Ictalurus furcatus*

Abstract

CRISPR/Cas9-mediated genome editing has paved new avenues for improving production-valued traits in aquaculture by knocking out or disrupting functional genes. However, utilizing CRISPR/Cas9-based knock-in (KI) of exogenous genes can expedite genetic improvement of traits such as disease resistance, which remains problematic in farmed fish. In this study, we successfully generated transgenic blue catfish (*Ictalurus furcatus*) of primarily Rio Grande strain ancestry with site-specific KI of the alligator (*Alligator sinensis*) cathelicidin (*As-Cath*) gene into the luteinizing hormone (*lh*) locus via two CRISPR/Cas9-mediated KI systems, assisted by the linear double-stranded DNA (dsDNA) and double-cut plasmid, respectively. High integration rates were observed with linear dsDNA (16.67%, [13/78]) and double-cut plasmid strategies (24.53%, [26/106]). In addition, the on-target KI efficiency of the double-cut plasmid strategy (16.04%, [17/106]) was 1.67 times higher than that of the linear dsDNA strategy (10.26%, [8/78]) based on the odds ratio. The relative expression of the *As-Cath* transgene of P_1 founders was detected in nine tissues, dominated by the kidney, skin, and muscle (14.30-, 7.71- and 6.92-fold change, $P < 0.05$). Moreover, the *As-Cath* transgenic blue catfish showed a higher cumulative survival rate than that of wild-type controls (80% *vs*. 30%, $P < 0.05$) following *Flavobacterium covae* infection. Survival during culture supports the challenge data as survival of *As-Cath* transgenic individuals was 97.1% while that of pooled non-transgenic individuals was observed to be less 87.0% ($P = 0.15$). The growth rates and external morphology of the transgenic and wild-type siblings were not different ($P > 0.05$), indicating no pleiotropic effects of the *As-Cath* transgene integration at the *lh* locus in the P_1 founders for this trait. Taken together, our findings demonstrate that CRISPR/Cas9-assisted KI of an antimicrobial peptide gene can be achieved in blue catfish with high integration efficiency, and *As-Cath* transgenic blue catfish have improved disease resistance, which is a promising strategy for disease reduction in aquaculture.

Keywords: Genome editing, cathelicidin, antimicrobial peptide, disease resistance, blue catfish

1. Introduction

Fish disease is one of the major obstacles to aquaculture development, as large-scale disease outbreaks can result in a series of losses, such as reduced yields, increased costs, and reduced income for farmers (Dunham and Elaswad, 2018; Naylor et al., 2021). Traditionally, farmers prefer to use antibiotic-like drugs to prevent or treat fish from diseases, thereby reducing financial losses (Cabello, 2006). For many years, antibiotic misuse has polluted the water environment, causing pathogens to develop drug resistance, and even endangering human health (Bondad-Reantaso et al., 2020; Karunasagar et al., 2020). In view of this, the WHO and FAO advocate the development and search for alternatives that do not cause antimicrobial resistance or produce antibiotic residues (WHO, 2017; FAO, 2021), and the U.S. government also strictly limits the use of antibiotics for preventive and therapeutic purposes in aquaculture (FDA, 2022). Although some vaccines have been developed, they are often pathogen- or fish species-specific under experimental conditions. Their successful applications in commercial pond environments are still dependent on vaccine efficacy, ease of administration, and availability due to the approval process and regulations (Sommerset et al., 2005; Dunham and Elaswad, 2018; Priya and Kappalli, 2022). As a promising substitute, antimicrobial peptides (AMPs) have broad-spectrum antimicrobial properties and do not cause drug resistance (Hancock, 2001; Wang et al., 2016). In recent years, AMPs have been used not only in the development of human medicine, such as anticancer drugs, but also in the prevention and treatment of livestock and aquatic animal diseases (Mookherjee et al., 2020; Wang et al., 2022a; Wang et al., 2023a).

Numerous studies have demonstrated the benefits of dietary supplementation with AMPs on innate immunity and disease resistance in aquatic animals (Welker et al., 2010; Lin et al., 2015; Abdel-Wahab et al., 2021; Mi et al., 2022). On the one hand, as a feed additive, high demand is required, and relying only on natural AMPs will not be sufficient to fill that market need. On the other hand, synthesizing bulk AMPs artificially is time-consuming, labor-intensive, and expensive (Wang et al., 2023a). The administration of AMPs as feed additives for disease reduction usually provides short-term disease control. Alternatively, genetic strategy via the integration of antimicrobial peptide genes (AMGs) into the genome of target fish can encode functional AMPs and enable their expression, leading to disease-resistant genetic lines with multigenerational inheritance (Wang et al., 2022a; Wang et al., 2023b; Wang et al., 2023c). The

use of engineered AMGs as transgenes to strengthen fish resistance to invading pathogens has yielded promising results in recent years. One of the most important developments in disease-resistant enhancement was Dunham's pioneering research on a cecropin-transgenic channel catfish (*Ictalurus punctatus*) line (Dunham et al., 2002). Since then, an increasing number of studies documented that AMG-transgenic fish exhibited an elevated resistance to various pathogens in medaka (*Oryzias latipes*) (Sarmasik et al., 2002), grass carp (*Ctenopharyngodon idella*) (Zhong et al., 2002; Mao et al., 2004), zebrafish (*Danio rerio*) (Yazawa et al., 2006; Hsieh et al., 2010), rainbow trout (*Oncorhynchus mykiss*) (Chiou et al., 2014) and channel catfish (*Ictalurus punctatus*) (Abass et al., 2022). Therefore, AMGs as transgenes are robust and promising for improving disease resistance in fish species, which is beneficial to the aquaculture sector.

Conventional transgenesis of an AMG in a random integration manner confers enhanced immunity to fish. However, the random insertion approach may impair normal gene function and render tracing the specific integration sites impossible. As one of the most effective genetic improvement tools, CRISPR/Cas9-based genome editing is vital in enhancing producer-valued traits in aquaculture (Hallerman et al., 2022). A considerable number of studies documented that CRISPR/Cas9-mediated knock-in (KI) can create on-target insertions of foreign genes in zebrafish (Auer et al., 2014; Morita et al., 2017) and medaka (Watakabe et al., 2018) via the homology-directed repair (HDR) or non-homologous end joining pathways. Although effective KI systems for model fish have been reported, only a few such approaches have been applied in farmed fish. Several CRISPR/Cas9-assisted KI strategies including linear double-stranded DNA (dsDNA)-, double-cut plasmid-, and ssODN (single-stranded oligodeoxynucleotide)-mediated systems were adopted to generate enhanced disease-resistant channel catfish (Simora et al., 2020; Wang et al., 2023b) and docosahexaenoic acid/eicosapentaenoic acid- (DHA/EPA)-enriched channel catfish (Xing et al., 2022) with relatively high efficiencies.

Blue catfish (*Ictalurus furcatus*) is a valuable food fish due to its large size and flavorful flesh (Pflieger, 1997). However, blue catfish tend to contract some bacterial diseases more than channel catfish, especially after handling or hauling (Graham, 1999). Dunham et al. (2008) discovered that blue catfish were more susceptible to *Flavobacterium columnare* bacteria than channel catfish. Therefore, improving the disease resistance of blue catfish is critical to

achieving commercial farming as this species is hybridized with channel catfish to produce an F_1 hybrid exhibiting heterobeltiosis for several traits (Gosh et al., 2022). Cathelicidins are an important AMP family with broad-spectrum antibacterial, antifungal, antiviral, and antiparasitic properties (Mookherjee et al., 2020). Alligator (*Alligator sinensis*) cathelicidin (*As-Cath*) demonstrated potent antimicrobial activity against a variety of pathogens *in vitro* and *in vivo* (Chen et al., 2017). A blast search of the cathelicidin sequence against the blue catfish genome prior to the initiation of the current experiment did not reveal the cathelicidin or cathelicidin-like gene.

CRISPR/Cas9-mediated genome editing has already been widely used in channel catfish to improve growth, control reproduction, increase omega-3 fatty acid content and enhance disease resistance (Khalil et al., 2017; Qin, 2019; Simora et al., 2020; Coogan et al., 2022; Xing et al., 2022). We recently generated *As-Cath* transgenic channel catfish lines via different CRISPR/Cas9-mediated KI systems, and the transgenic fish displayed increased resistance against fish bacteria (Wang et al., 2023b). However, no research has been conducted to create transgenic germlines in the blue catfish.

Luteinizing hormone (LH) acts as a key regulator throughout the reproductive cycle and significantly contributes to the final maturation of female fish (Gen et al., 2003; Chu et al., 2014). Studies conducted earlier have shown that oocyte maturation and ovulation were disrupted in *lh*-mutant zebrafish (Chu et al., 2014). In addition, *lh*-deficient channel catfish are sterile, and the sterilization can be temporarily reversed by hormone therapy (Qin et al., 2016; Qin 2019). In this case, the gonads and gametes are fully developed, but cannot attain final maturation. To establish disease-resistant blue catfish germlines and reversible transgenic sterilization system, we replaced the *lh* gene with the *As-Cath* transgene by harnessing CRISPR/Cas9-mediated KI systems coupled with a linear dsDNA (dsDNA_As-Cath) or double-cut plasmid (pUC57_As-Cath). The efficiencies of total KI and on-target KI of the two systems were assessed, as well as the potential effects of CRISPR/Cas9-mediated microinjection on the hatchability, fry survival and growth in blue catfish. The relative expression of the *As-Cath* transgene compared to the wild-type (WT) individuals was determined for various tissues.

Furthermore, we compared the resistance of *As-Cath*-transgenic and WT individuals against *Flavobacterium covae* bacteria.

2. Materials and methods

2.1 Ethical approval

The experimental blue catfish were of an unknown percentage or complete Rio Grande strain ancestry based on each individual being heavily spotted, and were reared at the Fish Genetics Research Unit, E.W. Shell Fisheries Research Center, Auburn University, AL. The Institutional Animal Care and Use Committee at Auburn University (AU-IACUC) approved the experimental techniques used in this study. All experiments were carried out following the procedures and standards established by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC).

2.2 Target locus for transgene insertion

As the target integration site, we chose the blue catfish *lh* locus expecting to create disease-resistant and sterile genetic lines. Based on the published genome of blue catfish (Wang et al., 2022b), we targeted the middle of exon 2 of *lh* site using a single guide RNA (sgRNA) targeting is in (Figure 9A). The inserted coding sequence (CDS) of cathelicidin gene is derived from *Alligator sinensis* (*As-Cath*, GeneBank accession number: XM_006037211.3) (Chen et al., 2017).

2.3 Design of donor DNA, sgRNA, and CRISPR/Cas9 system

Gene-targeted KI can be achieved via the HDR pathway using the linear dsDNA or engineered plasmid as donor templates. The two types of donors harboring the *As-Cath* gene targeting the *lh* locus were adopted and installed by CRISPR/Cas9-mediated system to conduct on-target KI. The first system was built with a dsDNA as the donor (dsDNA_As-Cath). The main *As-Cath* cassette (promoter_As-Cath_polyA) was flanked by two 300-bp homology arms (HA1 and HA2) derived from the blue catfish *lh* gene to create the dsDNA donor. In the dsDNA_As-Cath system, the sgRNA only cleaves the blue catfish genomic DNA (gDNA) at the *lh* locus, and then the linear dsDNA donor is pasted to the cleavage, repairing the damaged DNA. The donor in the second system was an engineered double-cut plasmid (pUC57_As-Cath). To create the double-cut

plasmid donor, the same linear dsDNA flanked by the sgRNA recognition sequences (sgRNA-PAM, 23 bp) on each side was cloned into the pUC57 vector at the *Eco*RV enzyme digestion site (Appendix 5). In the pUC57_As-Cath system, sgRNAs simultaneously cut the *lh* locus and the plasmid donor, and then the cleaved *lh* and the linear dsDNA generated from the plasmid can induce DNA repair via the HDR pathway. In detail, the main *As-Cath* cassette includes the zebrafish ubiquitin (UBI) promoter driving the expression of the *As-Cath* transgene (Mosimann et al., 2011), the CDS of the *As-Cath* gene, and a poly-A tail to enhance translation (Figure 9A, Appendix 5). The linear dsDNA and circular plasmid were synthesized by Genewiz (South Plainfield, NJ).

The sgRNA was designed using the CRISPRscan online tool (https://www.crisprscan.org/), and potential off-target events were excluded using the online tool Cas-OFFinder (http://www.rgenome.net/cas-offinder/) (Bae et al., 2014). According to the manufacturer's instructions, the sgRNAs were generated using the Maxiscript T7 kit (Thermo Fisher Scientific, Waltham, MA) *in vitro*. The RNA Clean and Concentrator Kit (Zymo Research, Irvine, CA) purified sgRNAs. The concentrations and qualities of the sgRNAs were determined using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and a 1 % agarose gel with the 1× tris-borate-EDTA (TBE) buffer. The synthetic sgRNAs were diluted to 300 ng/μL, aliquoted (2 μL/tube) into PCR tubes and stored at −80 ℃ until use. The Cas9 protein powder was purchased from PNA BIO Inc. (Newbury Park, CA) and was diluted to 50 ng/μL with DNase/RNase-free water and then stored at −20 ℃ until use. The sgRNA and universal primer used in this study are listed in Table 4.

2.4 Microinjection, transgenic fish production and rearing

Mature blue catfish females and males were paired for artificial spawning in aquaria (60 cm × 45 cm × 30 cm). The general artificial spawning procedures followed those of Su et al. (2013). Briefly, we selected individuals weighing more than two kilograms for pairing. Blue catfish pairs (body weight difference < 1 kg) were anesthetized in a solution of 100 mg/L tricaine methanesulfonate (MS222; Syndel, Ferndale, WA) and the water was buffered using100 mg/L sodium bicarbonate. Females were then administered intraperitoneal injections of luteinizing hormone-releasing hormone analogue (LHRHa) at 20 μL/kg body weight (BW) followed by a

Table 4. Oligonucleotide sequences for single guide RNA (sgRNA) synthesis and PCR. Primers used for the detection of the alligator (*Alligator sinensis*) cathelicidin (*As-Cath*) transgene/ubiquitin promoter region, luteinizing hormone (*lh*) mutation, and the relative expression of the *As-Cath* gene in blue catfish (*Ictalurus furcatus*). HA, homology arm.

Oligo name	Nucleotide sequence (5' → 3')	Product Size (bp)	Purpose
sgRNA synthesis			
sgRNA	TTCAAACCGCCATCTGCAGC	—	sgRNA synthesis
Universal Primer	TTTTGCACCGACTCGGTGCCACTTTTT CAAGTTGATAACGGACTAGCCTTATTT TAACTTGCTATTTCTAGCTCTAAAAC	—	Scaffold of the sgRNA synthesis
PCR Primers			
Cath-F	TTCAGGAGCCGTACTGTTCC	597	Determine the Cath-polyA region of the *As-Cath* transgenic fish
Cath-R	GCATTCTAGTTGTGGTTTGTCCA		
Prom-F	ACCCTTTGCCACAGTTCTCC	542	Determine the Prom-Cath region of the *As-Cath* transgenic fish
Prom-R	GGCCCTTGGTTGTAGACG		
HA1-F	TAAGGCCACGTTTCGATTCT	573	Determine the junction of the HA1
HA1-R	TCATTTTGCCGTCTGTTGTT		
HA2-F	TGAGTTTGGACAAACCACAAC	598	Determine the junction of the HA2
HA2-R	TTGATTGAAAATGTTTCCCTGTT		
LH-F	TGAGCGATCACAGCAAAATC	594	Determine the mutation of the *lh* gene
LH-R	GCAGCTTAGTGCGACAGGAT		
qRT-PCR Primers			
Cath_qPCR-F	GCAGGGGTCTATTCAAGAAGC	125	For quantitative real-time PCR (qRT-PCR) of *As-Cath* transgene
Cath_qPCR-R	GTCTGGATCTCACCGCCTTC		
18s-F	GAGAAACGG CTACCACATCC	128	Internal control for quantitative real-time PCR
18s-R	GATACGCTCATT CCGATTACAG		

100 μL/kg of resolving dose 12 h later. At the same time, males were injected with one dose of 50 μL/kg LHRHa. Injected pairs were gently placed into an aquarium with flow-through pond water and aerated with compressed air (dissolved oxygen above 6 mg/L; water temperature 24—26℃). The females were checked for eggs every 4 h, at 36 h after the hormone injection. The males were removed and euthanized to collect testes once a small egg layer was observed at the bottom of the tanks, and females were anesthetized to hand-strip eggs in a 20-cm greased spawning pan. Fresh testes were rinsed, weighed, and crushed. Sperm were prepared in 0.9 % saline solution (g:v = 1:10). Two milliliters of sperm solution were added to approximately 300 eggs and gently mixed. After a one-minute mixing, sufficient pond water was added to the eggs to activate the sperm, then the sperm/egg mixture was gently swirled for 30 s. The embryos were kept in a single layer in the pan with additional water supplied, and the embryos were allowed to harden for 15 min before microinjection.

The CRISPR/Cas9-mediated microinjection solution was composed of Cas9 protein, sgRNA, donor template and phenol-red indicator in the ratio of 2:1:1:0.5. For each system, 4 μL of Cas9

protein (50 ng/μL), 2 μL of sgRNA (300 ng/μL), 2 μL of donor template (linear dsDNA or double-cut plasmid, 50 ng/μL), and 1 μL of phenol red solution were mixed for microinjection. The Cas9 protein and sgRNA were first mixed and incubated on ice for 10 min to form a Cas9-sgRNA complex, then donor templates were added. The mixed solution for each system was microinjected into one-cell stage embryos as previously described (Khalil et al., 2017). Every 6 μL of the mixture was loaded into a 1.0 mm OD borosilicate-glass needle that was pulled by a micropipette puller (Sutter Instruments, Model P-97, VWR, GA), then injected into 600 embryos. We microinjected 1,000 embryos, dividing them into 5 random replicates for each KI system, and another 1,000 embryos divided into 5 replicates were prepared as a non-injected control group (nCT). In addition to CRISPR/Cas9-based microinjection and nCT groups, we also microinjected 1,000 embryos (divided into 5 replicates) with phenol red (diluted with 0.9 % saline) without donors as an injected group (iCT). All these embryos were from the same parents, and the microinjection was terminated after 90 min post-fertilization.

After microinjection, all injected and control embryos were transferred into sterilized 10-L tubs filled with 7-L Holtfreter's solution (59 mmol NaCl, 2.4 mmol NaHCO3, 1.67 mmol MgSO4, 0.76 mmol CaCl2, 0.67 mmol KCl) (Armstrong and Malacinski, 1989) for hatching. All tubs were randomly placed in two flow-through hatching troughs (300 cm × 60 cm × 45 cm), and each tub received one airstone for continuous oxygen (> 5 mg/L). A heater was utilized upstream near the water inlet to keep the water temperature at 26—28 ℃ in each trough. Holtfreter's solution was completely replaced twice per day, and dead embryos/fry were collected and recorded daily to determine the hatchability and early fry survival.

All hatched sac-fry were moved to new tubs filled with pond water and fed ad libitum four times per day with live *Artemia nauplii*. After one week of culture in tubs, all fry were stocked into 60 L recirculating aquaria by treatment for growth. The feed pellet size was adjusted to match the size of the fish's mouth as the fish grew. In detail, fry in tanks were fed ad libitum with Purina® AquaMax® powdered feed (50 % crude protein, 17 % crude fat, 3 % crude fiber, and 12 % ash) four times per day for three months. Then fingerlings were fed with Cargill Aquaxcel WW Fish Starter 4512 (45 % crude protein, 12 % crude fat, 3 % crude fiber, and 1 % phosphorus) twice a day for three months to apparent satiation. Juvenile fish were fed with WW 4010 Transition feed

(40 % crude protein, 10 % crude fat, 4 % crude fiber, and 1 % phosphorus) once a day to apparent satiation.

2.5 Integration analysis and mutation detection

All fingerlings were pit-tagged (Biomark Inc., Boise, Idaho, USA) after a 5-month culture to individual identification. The pelvic fin clip and barbel were mixed and collected from each anesthetized fish for DNA extraction and genotypic identification. The marked fish from different groups were then combined and randomly distributed into two flow-through tanks (1,200 L volume filled with ~800 L of water) with the same density (150 fish/tank) for comparison of growth. During this phase, all fish were provided WW 4010 Transition feed once a day for satiation. The same genotyping strategy was adopted for these two CRISPR/Cas9-mediated systems: the CDS region of the *As-Cath* gene was amplified to confirm transgene insertion using primers Cath-F/R (forward and reverse), and the promoter region was amplified using primers Prom-F/R. With respect to junctions, HA regions were amplified using primers HA1-F/R and HA2-F/R to determine an on-target insertion. Primers were designed using the online software Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) and are listed in Table 4. PCR was performed in a 10-μL system, and PCR products were resolved and visualized by running 1.0% agarose gel with 1 × tris-acetate-EDTA (TAE) buffer, and a bright band of each region with the corresponding length indicated an on-target positive individual (LH^{-}_As-$Cath^{+}$). Some individuals had the *As-Cath* transgene inserted but no detectable junction regions; we then classified them as potential off-target positives (LH^{+}_As-$Cath^{+}$).

Concerning the non-*As-Cath* inserted fish, we needed to detect the potential *lh*-mutant individuals (LH^{-}_As-$Cath^{-}$). In this case, PCR was performed in a 20 μL-volume system using the Expand High FidelityPLUS PCR System (Roche Diagnostics, Indianapolis, IN, USA) according to Elaswad et al. (2018), and LH-F/R primers were used. Then, the surveyor mutation detection assay was carried out using the Surveyor Mutation Detection Kit (Integrated DNA Technologies, IDT, Coralville, Iowa, USA) per the instructions (Qiu et al., 2004). A negative control reaction was included in the assay by using gDNA from the nCT group. Surveyor-digested DNA samples were electrophoresed in a 2% agarose gel with 1 × TBE buffer for 1 hour

and compared to WT samples. Additionally, we identified the fish as negative (LH^+_As-$Cath^-$) when no *As-Cath* insertions and no *lh* mutations were detected.

To confirm the inserted transgene or mutations, we cloned the PCR products of putative positive or mutant individuals using the TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA) for sequencing following the instructions. The PCR products were transformed into One Shot TOP10F chemically competent *Escherichia coli* (Invitrogen, Carlsbad, CA). Then three colonies from each transgenic fish were randomly picked to perform Colony PCR, and liquid *E. coli* cultures were prepared for rolling circle amplification (RCA) sequencing by Sequetech (Mountain View, CA) (using the M13 forward primer). As for the potential *lh*-mutant individuals, we selected ten colonies from each fish to perform Colony PCR and RCA sequencing. Finally, the sequencing results were blasted with transgenes/mutations using MAFFT (version 7, https://mafft.cbrc.jp/alignment/server/) to identify inserted DNA fragments or mutant sequences.

2.6 Determination of mosaicism and transgene expression

Three 12-month-old on-target positive fish and three sham-injected control fish were chosen and sacrificed. The fin, barbel, skin, muscle, intestine, head kidney, stomach, liver, and gonad of a single individual were collected in 1.5 mL tubes and immediately transferred to liquid nitrogen for DNA and RNA isolation. To determine the potential mosaicism and the relative expression of the *As-Cath* transgene, PCR, and quantitative real-time PCR (qRT-PCR) were conducted. Total RNA was isolated from various tissues using the TRIzol reagent (Thermo Fisher Scientific) and was reverse transcribed to cDNA using the iScript™ Synthesis Kit (Bio-Rad, Hercules, CA) following the manufacturer's protocols.

qRT-PCR was performed on a C1000 Thermal Cycler using SsoFast™ EvaGreen Supermix kit (Bio-Rad, Hercules, CA) according to the instructions. The cDNA products were diluted to 250 ng/μL, and 1 μL template was used in a 10 μL reaction volume. The mRNA level of 18S rRNA was used as an internal control, and the detailed qRT-PCR procedure was set up according to Coogan et al. (2022). The raw crossing-point (C_t) values were collected using CFX Manager Software (version 1.6, Bio-Rad). The expression level of the *As-Cath* transgene to the 18S rRNA gene of transgenic fish against that of non-transgenic sibling fish was converted to determine

fold differences. Each sample was analyzed in triplicate using the formula $2^{(-\Delta\Delta CT)}$ (Livak and Schmittgen, 2001), which sets the zero expression of the non-transgenic full-siblings to 1× for comparison. The primers used for qRT-PCR are listed in Table 4.

2.7 *Flavobacterium covae* challenge

Healthy P_1 blue catfish fingerlings with BW 6.25 ± 1.03 g (mean ± SEM), including four genotypes (10 fish/genotype): LH^-_As-$Cath^+$, LH^+_As-$Cath^+$, LH^+_As-$Cath^-$ and WT were separated into four net baskets (15cm × 15cm × 25cm), and acclimated in a hatching trough for three days before being transferred to a challenging tank (3m × 0.6m × 1.2m) with aeration (dissolved oxygen > 5 mg/L). The challenge tank was supplied with flow-through pond water to maintain a depth of 20 cm. A revived *F. covae* isolate (strain ALG-00-530) on modified Shieh agar (MSA) was inoculated into a culture of 12 mL of modified Shieh broth (MSB) in a 50-mL sterile flask and grown for 12 h at 28°C in a shaker incubator at 150 rpm. The culture was then subcultured into 200 mL MSB (5 mL additions) within 500 mL flasks and grown for another 12 h. The optical density was adjusted to OD_{540} = 0.731 and then spread plate dilutions were performed to determine the final inoculum concentration. One hundred microliters of the inoculum were serially diluted and spread onto MSA agar plates and incubated at 28 °C for 48 h to quantify the concentration of the inoculum. One combined flask containing 325 mL of inocula (4.75×10^8 CFU/mL) was immediately dispersed into the challenge tank; the fish were then immersed statically for 1 h at ~28 °C (immersion dose: 2.57×10^6 CFU/mL) without flow water. Meanwhile, a mock-challenged tank was used as the control but incorporated another 20 fish (5 fish/genotype) in a 20-cm depth of pond water for 1 h with sterile 325 mL of MSB instead of the bacterial culture. After the 1 h bacterial infection, water flow in the tanks was resumed.

During the first 72 h of the experiment, tanks were observed for mortality every four hours, then three times daily. Infected fish were continuously monitored for clinical signs of *F. covae* for 10 days by isolating bacteria from the liver to determine the cause of death, and dead individuals were recorded over time.

2.8 Pleiotropic effect

To evaluate any pleiotropic effects caused by the two CRISPR/Cas9 mediated *As-Cath* into the *lh* locus, the individual BW was collected at 5-, 7-, 20- and 31-months post-hatch (mph) to compare growth. Then the fish were photographed individually at 31 mph. Body shape parameters, including the total length (TL), standard length (SL), head length (HL), eye diameter (ED), body depth (BD), and caudal depth (CD), were measured/analyzed individually using ImageJ software (https://imagej.nih.gov/ij/) to evaluate the potential deformity (Figure S7 in Appendix 2).

2.9 Statistical analysis

Hatchability, fry survival rate, and growth data from different systems/genotypes were analyzed using one-way ANOVA/Tukey's multiple comparisons test. The body shape parameters from different groups were compared using multivariate ANOVA followed by Hotelling's T-squared test. To compare the KI efficiency of different systems, Fisher's exact test and odds ratio (OR) were adopted. The survival curves of challenge experiments from four genotypes were compared by the Kaplan-Meier plots followed by Log-rank (Mantel-Cox) test. Gene expression between transgenic and non-transgenic fish was analyzed with an unpaired Student's *t*-test. All statistical analysis was achieved using GraphPad Prism 9.4.1 (GraphPad Software, LLC). The statistical significance level was set as $P < 0.05$, and all data were presented as the mean ± standard deviation (SD).

3. Results

3.1 Targeted KI of the *As-Cath* into the *lh* locus

In both systems, low hatching rates were observed following the CRISPR/Cas9-mediated microinjection. All microinjected groups had lower hatching rates (8.50 ± 0.44% for iCT, $P = 0.004$; 8.40 ± 1.06% for dsDNA_As-Cath, $P = 0.003$; 12.40 ± 0.79% for pUC57_As-Cath, $P = 0.624$) compared to the nCT group (14.00 ± 1.21%). However, the pUC57_As-Cath system had a significantly higher hatching rate than the dsDNA_As-Cath ($P = 0.035$) (Figure 9B). In the present study, we obtained 65, 78 and 106 alive fingerlings after microinjecting 1,000 embryos in the iCT, dsDNA_As-Cath and pUC57_As-Cath groups, respectively. In addition, 92 fry were generated from the nCT group without microinjection. Our statistical results indicated that the

microinjection did not affect early fry survival compared to the nCT group (P = 0.062) (Figure 9C).

Both linear dsDNA and double-cut-plasmid strategies integrated the *As-Cath* transgenic germline in blue catfish. Total KI efficiency in the pUC57_As-Cath system was significantly higher than that in the dsDNA_As-Cath system [24.53% (26/106) *vs*. 16.67% (13/78), P = 0.019] (Figure 9D). Moreover, the pUC57_As-Cath system had a significantly higher on-target KI efficiency than the dsDNA_As-Cath [16.04% (17/106) *vs*. 10.26% (8/78), P = 0.007] (Figure 9E). According to the OR, the total and on-target KI rates of the pUC57_As-Cath system were 1.63 and 1.67 times higher than those of the dsDNA_As-Cath system, respectively (Table 5). These findings suggested that the pUC57_As-Cath system effectively created the *As-Cath* transgenic genetic lines by increasing KI efficiency and reducing the off-target events in blue catfish. All the positive individuals were confirmed by gel electrophoresis and sequencing (Figure 9F-G, Figure S8 in Appendix 2).

Table 5. The summary of total knock-in (KI) and on-target KI efficiency from two CRISPR/Cas9-mediated systems in blue catfish (*Ictalurus furcatus*). KO, knock out; OR, odds ratio; #, the odds ratio of total KI efficiency of the pUC57_As-Cath system relative to dsDNA_As-Cath, OR = (26×65)/(13×80) = 1.63; ##, the odds ratio of on-target KI efficiency of the pUC57_As-Cath system relative to dsDNA_As-Cath, OR = (17×70)/(8×89) = 1.67; dsDNA_As-Cath, CRISPR/Cas9-mediate KI system coupled with a dsDNA as the donor; pUC57_As-Cath, CRISPR/Cas9-mediate KI system coupled with a double-cut plasmid as the donor.

System	Efficiency (number)			OR
	Total KI	On-target KI	Only KO	
dsDNA_As-Cath	16.67% (13/78)	10.26% (8/78)	5.88% (3/51)	1.63[#]
pUC57_As-Cath	24.53% (26/106)	16.04% (17/106)	8.33% (6/72)	1.67[##]

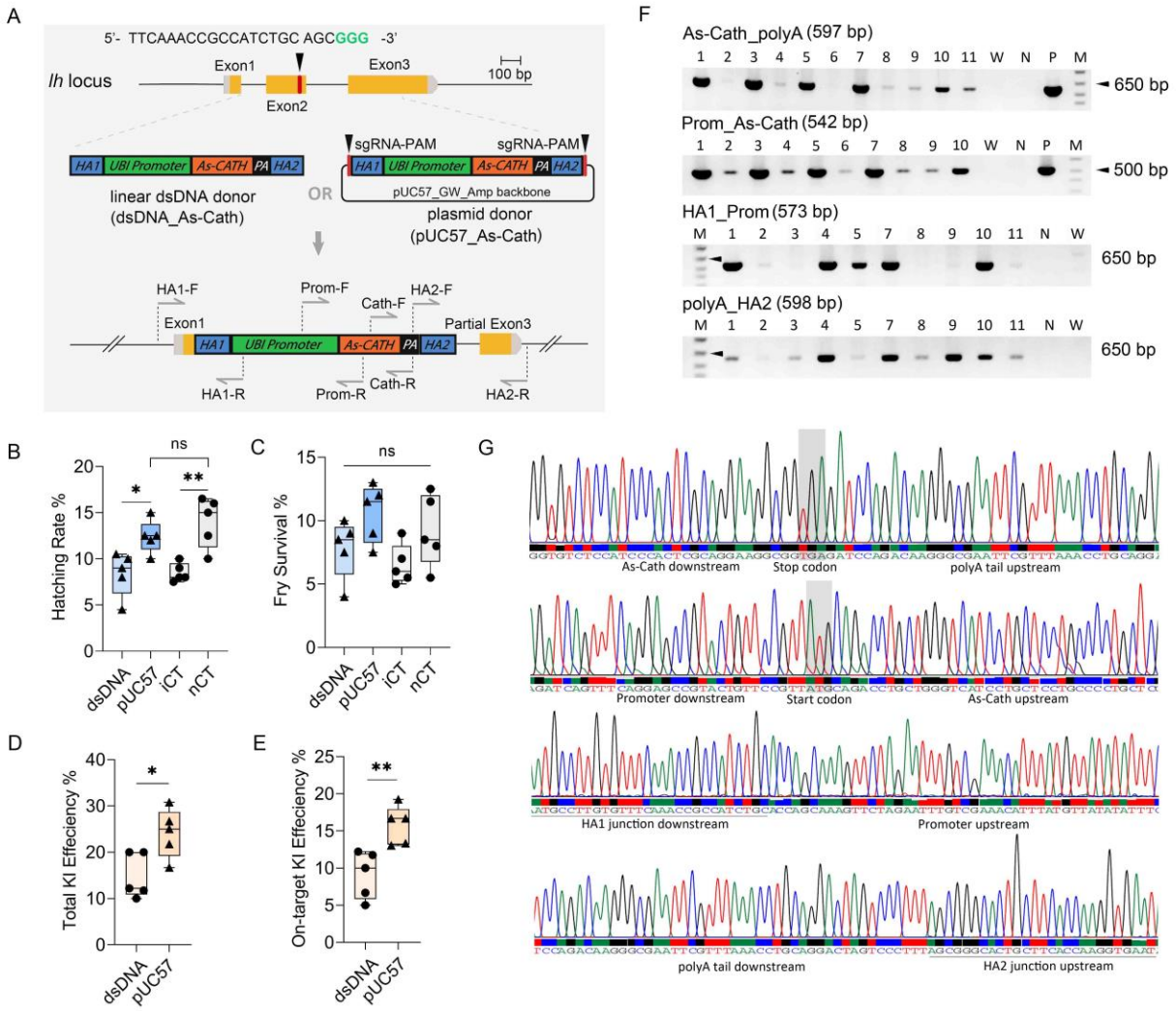


Figure 9. The knock-in (KI) efficiency of the *As-Cath* transgene at the *lh* locus using two different CRISPR/Cas9-mediated systems in blue catfish (*Ictalurus furcatus*). (A) Schematic illustration of two different CRISPR/Cas9-mediated systems (dsDNA_As-Cath *vs.* pUC57_As-Cath). The sgRNA protospacer adjacent motif (PAM) was highlighted in green, and the sgRNA target sites were indicated with black arrows. The detection strategy for positive individuals was illustrated using specific primer sets. **(B, C)** The effects of the microinjection on hatching rate and fry survival. **(D, E)** Total and on-target KI efficiency of dsDNA_As-Cath and pUC57_As-Cath systems. **(F)** On-target positive fish were identified by detecting the As-Cath region (As-Cath_polyA, 597 bp), promoter region (prom_As-Cath, 542 bp), and junctions (HA1_prom, 573 bp; polyA_HA2, 598 bp) using 1% TAE gel electrophoresis. The number above each lane represented the ID of the fish. Full gel electrophoresis is attached in Figure S8 in Appendix 2. **(G)** Sequencing results corresponding to Figure 9F. *As-Cath*, alligator (*Alligator sinensis*) cathelicidin; *lh*, luteinizing hormone; HA, homologous arm; UBI, zebrafish ubiquitin; PA, poly-A tail; Lane N, negative control using water as a template; Lane W, wild-type control; Lane P, positive control using a plasmid donor as a template; Lane M, DNA marker (1 kb), 300, 500 and 650-bp bands were highlighted with black triangles; LH^{-}_As-$Cath^{+}$, on-target positive fish, the *As-Cath* transgene was integrated at the *lh* locus; ns, not significant; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

3.2 Determination of *lh* mutations

In addition to the *As-Cath* transgenic blue catfish, we screened 5.88% (3/51) and 8.33% (6/72) *lh*-mutant fish from the dsDNA_As-Cath and pUC57_As-Cath systems, respectively. Compared to the WT control, *lh*-mutant individuals showed multiple bands after the Surveyor® digestion (Figure 10A, Figure S9 in Appendix 2). The Sanger sequencing results confirmed that the *lh* gene was mutated within or outside the Cas9 cleavage site. These sequence mutations included substitutions, deletions, and insertions that caused or did not cause amino acid (AA) alterations. The same type of substitution (C to G) was observed in individuals #9', #10', and #11', resulting in a mutant AA (Cysteine to Tryptophan). Two types of deletions were detected in #8' fish (−6 bp, −2 bp), leading to large amounts of mutated AA sequences or even premature termination of translation. In addition, insertions were observed in individuals #2', #5', and #12' (+4 bp, +2 bp, +4 bp). These insertions altered AA sequences at positions 52 or 53 and beyond. Two types of mutations were observed in #1' (C to T, C to G) and #8' (−6 bp, −2 bp/+1 bp) individuals, indicating that these two fish were mosaic (different mutations in barbel and fin) (Figure 10B).

The cysteine residues in the LH serve a primary physiological purpose by forming disulfide bonds between the two peptide chains. These disulfide bonds can stabilize the tertiary structure of LH and prevent it from denaturing under extreme conditions. In this study, in all *lh*-mutant blue catfish, the mutation of the *lh* gene resulted in the loss of cysteine residues. In particular, one cysteine residue was mutated in individuals #1', #9', #10' and #11'. And individuals #2', #5', #7', #8' and #12' had two cysteine residues mutated. Therefore, the biological function of LH may be diminished or destroyed due to the weakened and damaged structure.

3.3 Mosaicism and the *As-Cath* transgene expression

The potential on-target mosaic transgenic blue catfish were discovered in this study. Fish #1, #5 and #7 were determined as on-target positives by sequencing (Figure 9F). Our results indicated that two of these three fish (#1 and #5) had the *As-Cath* transgene detected in all nine tissues, including the fin, barbel, skin, muscle, intestine, kidney, stomach, liver, and gonad. However, one of them (#7) had the transgene found in six tissues (barbel, skin, muscle, intestine, kidney, and stomach) (Figure 11A, Figure S10 in Appendix 2), suggesting that #7 fish was an on-target

positive individual with mosaicism. In addition, #1' and #8' individuals were mosaic for the *lh* mutation (Figure 10).

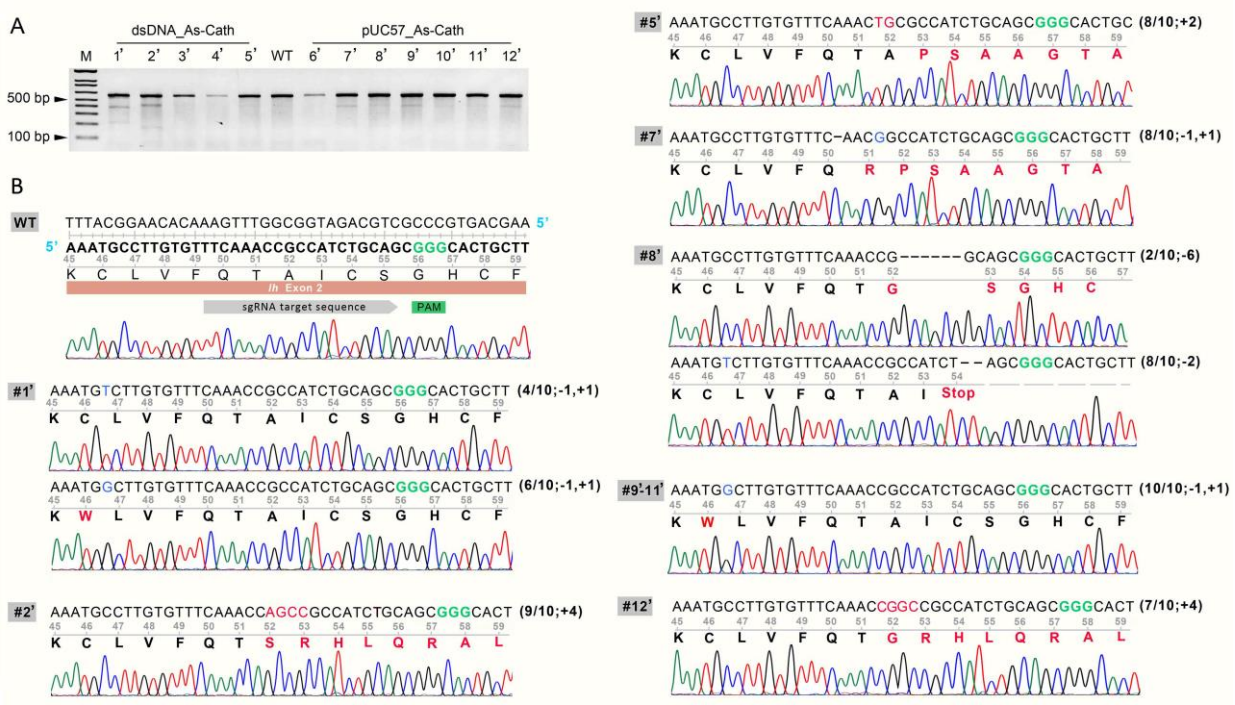


Figure 10. Identification of *lh*-mutant blue catfish (*Ictalurus furcatus*) and mutation analysis. (A) Identification of edited *lh* gene in P_1 blue catfish using the Surveyor mutation assay. Full gel electrophoresis is provided in Figure S9 in Appendix 2. Lane WT, wide-type control; Lane M, DNA marker (1 kb); 100- and 650-bp bands were highlighted with black triangles. **(B)** The CRISPR/Cas9 induced insertion-deletion (indel) mutations in the *lh* gene of blue catfish. Single guide RNA (sgRNA) target sequence and PAM are highlighted. Dashes and red letters indicate the indel of nucleotides, respectively. Blue letters show the substitution of nucleotides. Altered amino acid sequences are highlighted in red. Numbers on the right side of each sequence represent the number of nucleotides that have been deleted (−) or inserted (+). Each mutant allele was detected 4 times in 10 sequencing reactions (4/10). *lh*, luteinizing hormone; PAM, protospacer adjacent motif.

We found that the expression of the *As-Cath* gene was detected even in the absence of pathogenic infection in three LH^+_As-$Cath^+$ individuals. The three highest mRNA levels compared to the non-transgenic fish were determined in the kidney (14.30-fold change), skin (7.71-fold), and muscle (6.92-fold), followed by the liver (6.61-fold), intestine (5.78-fold) and gonad (3.81-fold). However, the barbel (3.24-fold), fin (1.97-fold), and stomach (1.59-fold) had the lowest expression compared to other tissues (Figure 11B).

3.4 Enhanced resistance against *F. covae*

The P_1 generation of *As-Cath*-integrated blue catfish exhibited enhanced resistance against *F. covae* bacteria compared to WT individuals from our challenge experiment. Most of the infected WT blue catfish showed visible clinical signs: saddleback, fin erosion and yellow skin coloration on the body. However, some *As-Cath* transgenic fish had mild clinical signs, with frayed fins but survived the infection. In comparison to the WT individuals, the LH^-_As-$Cath^+$ and LH^+_As-$Cath^+$ fish lines showed significantly improved survival rates post-*F. covae* infection in P_1 founders (LH^-_As-$Cath^+$ *vs.* WT: 80% *vs.* 30%, $P = 0.010$; LH^+_As-$Cath^+$ *vs.* WT: 80% *vs.* 30%, $P = 0.024$). Although on-target and off-target positives showed the same resistant enhancement against *F. covae* (LH^-_As-$Cath^+$ *vs.* LH^+_As-$Cath^+$: 80% *vs.* 80%), the on-target positive fish had a slower death rate (death on the 5th day post-infection) compared to the off-target positive individuals (death on the 3rd day post-infection) ($P = 0.5012$) (Figure 11C). In addition, 4 (4/30), 3 (3/30), 1 (1/23), 0 (0/12) and 2(2/9) fish were dead/lost during the first 31 months culture from WT, LH^+_As-$Cath^-$, LH^-_As-$Cath^+$, LH^+_As-$Cath^+$ and LH^-_As-$Cath^-$, respectively. This survival during culture supports the challenge data as survival of *As-Cath* transgenic individuals was 97.1% while that of pooled non-transgenic individuals was observed to be less 87.0% ($P = 0.15$).

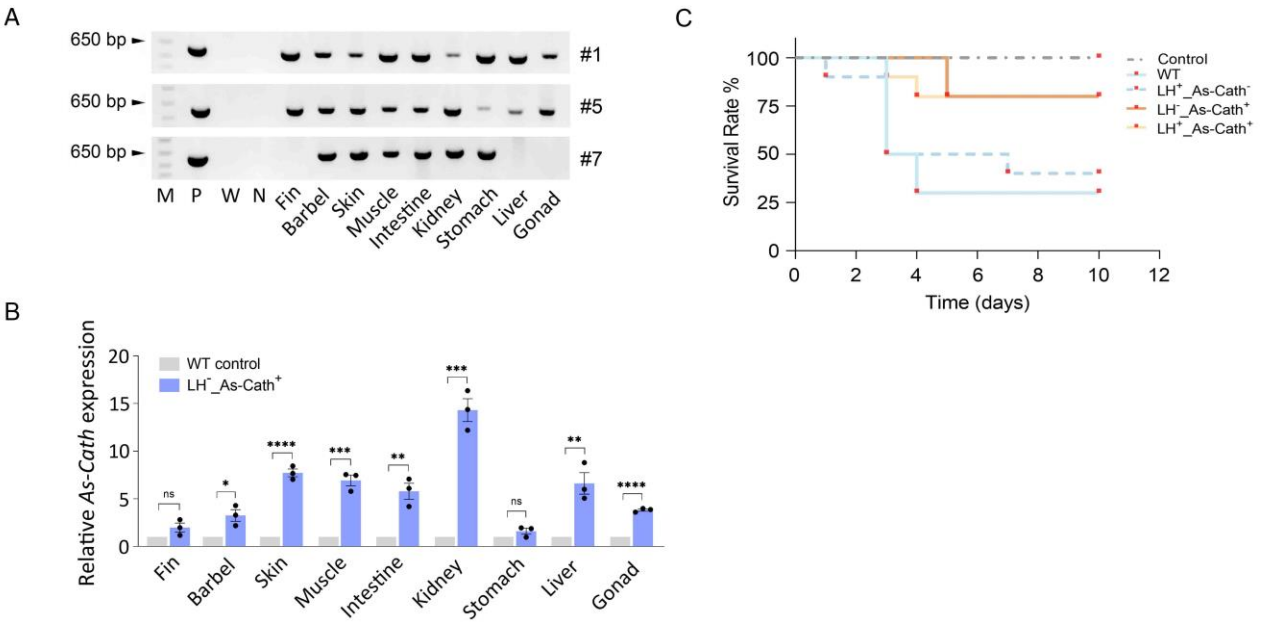


Figure 11. Identification of mosaic individuals and relative expression of the *As-Cath* transgene in blue catfish (*Ictalurus furcatus*). (A) The *As-Cath* transgene was detected in nine, nine, and six tissues from three positive fish (#1, #5, and #7), respectively, and the #7 individual was mosaicism. Full gel electrophoresis is provided in Figure S10 in Appendix 2. **(B)** Relative expression of the *As-Cath* transgene in various tissues from the LH^{-}_As-$Cath^{+}$ genetic line compared to WT fish (N = 3). **(C)** Kaplan-Meier plots of *As-Cath* integrated blue catfish against *Flavobacterium covae*. Four genotypes (10 fish/genotype) were immersed in 2.57×10^6 CFU/mL bacteria for 1 h in one challenge tank then flow-through water was resumed. In addition to the bacterial infection tank, one control tank (Control) with medium immersion was implemented. The Log-rank (Mantel-Cox) test to compare survival curves. *As-Cath*, alligator (*Alligator sinensis*) cathelicidin; *lh*, luteinizing hormone; WT, wild-type fish; LH^{+}_As-$Cath^{–}$, negative fish line (microinjected fish without *lh* mutation or *As-Cath* insertion); $LH^{–}$_As-$Cath^{+}$, on-target positive fish, the *As-Cath* transgene was inserted at the *lh* locus; LH^{+}_As-$Cath^{+}$, off-target positive fish, the *As-Cath* transgene was detected but not at the *lh* locus; ns, not significant; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

3.5 External morphology and body weight

Multivariate ANOVA results showed that there were no apparent differences in body shape parameters (TL, SL, HL, ED, BD, and CD) between WT and transgenic/gene-edited blue catfish ($P = 0.667$), indicating that *As-Cath*-transgenic blue catfish did not display the deformity or pleiotropic changes in body shape compared to the WT fish (Figure 12A, Table 6). The growth and survival rates at 5, 7, 20, and 31 mph are displayed in Table 7. The initial observed BW (5 mph) of the WT group (6.77 ± 1.72 g) was lower than that of LH^{+}_As-$Cath^{–}$ (7.30 ± 2.59 g), $LH^{–}$_As-$Cath^{+}$ (7.73 ± 2.81 g), LH^{+}_As-$Cath^{+}$ (8.14 ± 2.60 g) and $LH^{–}$_As-$Cath^{–}$ fish (8.22 ± 1.94 g), but no significant differences were observed in statistics among these groups ($P = 0.158$). Similarly, there were no significant differences in BW among these genotypes at 7, 20 and 31 mph (all $P > 0.05$) (Figure 12B). These data suggested that the P_1 *As-Cath*-transgenic and *lh*-deficient individuals did not show superiority in growth compared to the WT fish in the recirculating culture system.

Table 6. Morphological measurement of total length (TL), standard length (SL), head length (HL), eye diameter (ED), body depth (BD), and caudal depth (CD) in P_1 blue catfish (*Ictalurus furcatus*) at 31 months post-hatch. SD, standard deviation; CV, coefficient of variation; N, sample size. The body shape parameters from different genotypes were compared using multivariate ANOVA. SL, HL, BD, and CD were standardized by dividing by TL. No statistical differences were observed (P = 0.667) between different genotypes.

Trait	WT (N = 26)		LH^+_As-$Cath^-$ (N = 27)		LH^-_As-$Cath^+$ (N = 22)		LH^+_As-$Cath^+$ (N = 12)		LH^-_As-$Cath^-$ (N = 7)	
	Mean ± SD	CV	Mean ± SD	CV	Mean ± SD	CV	Mean ± SD	CV	Mean ± SD	CV
TL (cm)	13.33 ± 2.45	18.38	12.52 ± 3.03	24.20	14.68 ± 1.46	9.9	12.91 ± 2.78	21.53	15.21 ± 1.23	8.09
ED (cm)	0.43 ± 0.09	20.37	0.47 ± 0.08	16.37	0.42 ± 0.08	18.33	0.41 ± 0.08	18.16	0.45 ± 0.08	17.77
SL/TL	0.87 ± 0.04	4.60	0.94 ± 0.06	6.38	0.87 ± 0.02	2.23	0.89 ± 0.07	7.86	0.91 ± 0.03	3.30
HL/TL	0.13 ± 0.01	7.69	0.15 ± 0.02	13.33	0.13 ± 0.01	7.69	0.14 ± 0.01	7.14	0.12 ± 0.01	8.33
BD/TL	0.19 ± 0.01	5.26	0.23 ± 0.01	4.35	0.18 ± 0.01	5.55	0.24 ± 0.01	4.17	0.20 ± 0.01	5.00
CD/TL	0.07 ± 0.005	7.14	0.08 ± 0.006	7.50	0.07 ± 0.005	7.14	0.08 ± 0.005	6.25	0.07 ± 0.005	7.14

Note: WT, wild-type fish without microinjection; LH^+_As-$Cath^-$, negative fish without the *As-Cath* insertion or *lh* mutation; LH^-_As-$Cath^+$, on-target positive fish with the integration of the *As-Cath* transgene at the *lh* locus; LH^+_As-$Cath^+$, off-target positive fish with the *As-Cath* insertion but no *lh* mutation; LH^-_As-$Cath^-$, only *lh*-mutant without *As-Cath* insertion; LH/*lh*, luteinizing hormone; *As-Cath*, alligator (*Alligator sinensis*) cathelicidin.

Table 7. Mean body weight (BW) ± standard deviation (SD), coefficient of variation (CV), sample size (N) over time of P_1 blue catfish (*Ictalurus furcatus*), including wild type (WT), negative (LH^+_As-$Cath^-$), *As-Cath* on-target integrated (LH^-_As-$Cath^+$), *As-Cath* off-target integrated (LH^+_As-$Cath^+$), and only *lh*-mutant (LH^-_As-$Cath^-$). P_1 founders were generated on 7th June 2020. Fish were kept separately in 60-L aquaria with the density of 2 fry/L until 5 months post hatch (mph), then they were pit-tagged (11/7/2020) and transferred to a flow-through tank (~500-L water) with a mix of these five genotypes (initial number of fish was 30, 30, 25, 14, and 9) and fed daily to satiation. At the age of 20 mph, all fish were transferred into a larger circular tank (~800-L water) for growth. Differences in BW among these five genotypes were compared using one-way ANOVA. No statistical differences were observed ($P > 0.05$) between different genotypes at 5, 7, 20 and 31 mph.

Genotype	Mean BW (g) of blue catfish at different ages (Mean ± SD)											
	11/7/2020 (5 mph)			1/4/2021 (7 mph)			2/12/2022 (20 mph)			1/5/2023 (31 mph)		
	BW	CV	N	BW	CV	N	BW	CV	N	BW	CV	N
WT	6.77 ± 1.72	25.36	30	6.85 ± 2.47	36.07	29	122.23 ± 43.90	35.92	28	462.43 ± 169.01	36.55	26
LH^+_As-$Cath^-$	7.30 ± 2.59	35.46	30	8.42 ± 2.73	32.47	30	136.62 ± 40.93	29.96	29	501.02 ± 173.33	34.60	27
LH^-_As-$Cath^+$	7.73 ± 2.81	36.32	23	8.40 ± 2.69	32.00	23	135.18 ± 32.23	23.84	22	467.90 ± 103.66	22.15	22
LH^+_As-$Cath^+$	8.14 ± 2.60	31.90	12	7.89 ± 3.15	39.92	12	141.25 ± 49.77	35.24	12	447.96 ± 164.21	36.66	12
LH^-_As-$Cath^-$	8.22 ± 1.94	23.57	9	8.17 ± 1.98	24.30	9	131.88 ± 39.63	30.05	8	531.71 ± 98.29	18.49	7

Note: WT, wild-type fish without microinjection; LH^+_As-$Cath^-$, negative fish without the *As-Cath* insertion or *lh* mutation; LH^-_As-$Cath^+$, on-target positive fish with the integration of the *As-Cath* transgene at the *lh* locus; LH^+_As-$Cath^+$, off-target positive fish with the *As-Cath* insertion but no *lh* mutation; LH^-_As-$Cath^-$, only *lh*-mutant without *As-Cath* insertion; LH/*lh*, Luteinizing hormone; *As-Cath*, alligator (*Alligator sinensis*) cathelicidin.

4. Discussion

Genome editing tools have been extensively employed to enhance producer-favorable traits in fish by disrupting or silencing functional genes. Nevertheless, CRISPR-mediated genome editing for KI transgenes is still rarely achieved due to poor integration efficiency. This study used CRISPR/Cas9 systems coupled with linear dsDNA and double-cut plasmid donors to confer immunity to columnaris disease for blue catfish via a transgenically encoded alligator cathelicidin (*As-Cath*) gene inserted and replacing a reproduction-related locus (*lh*). The double-cut plasmid strategy resulted in a high integration rate with low off-target events. In addition, a high mRNA level of the *As-Cath* transgene was determined in the kidney, and mucosal tissues (skin and muscle) in the LH^-_$As\text{-}Cath^+$ genetic line. In contrast to WT blue catfish, knocking in the *As-Cath* transgene at the *lh* locus increased survival up to 2.5-fold post-bacterial infection. Our findings demonstrate that it is feasible to create *As-Cath* transgenic blue catfish utilizing precise CRISPR/Cas9 tools, which is a promising approach to disease reduction in aquaculture for this and other species.

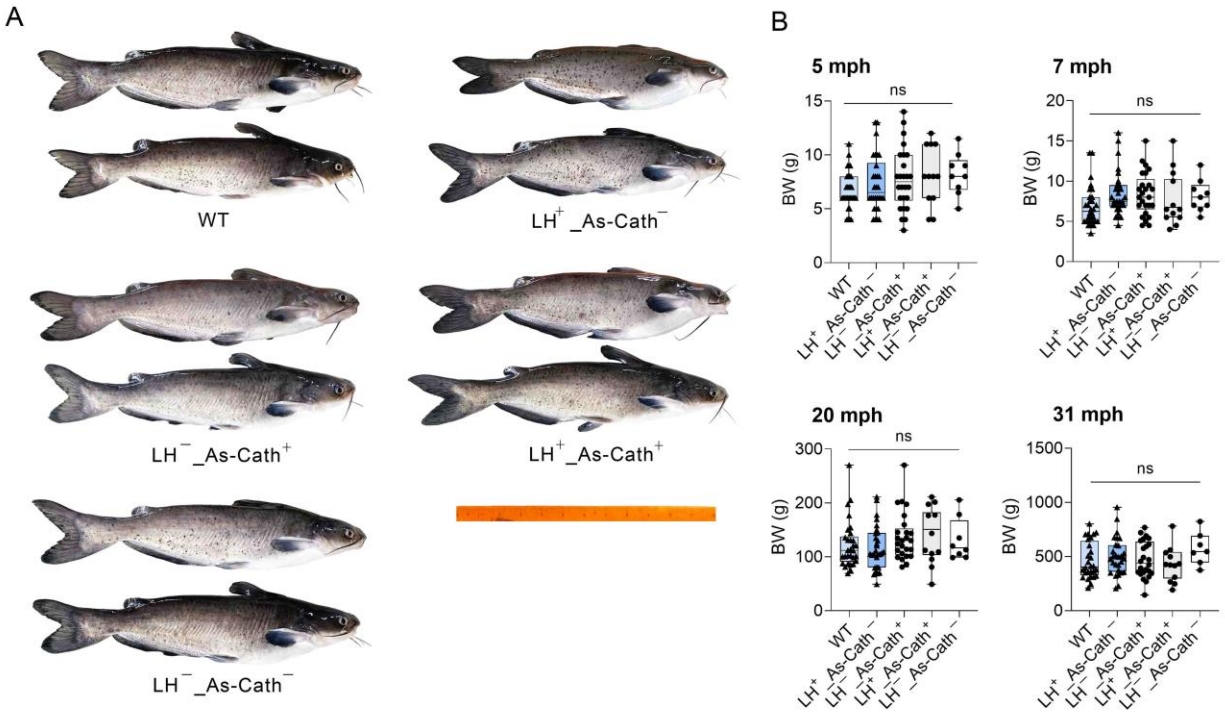


Figure 12. Body shape and body weight (BW) determination of P_1 WT, negative, *As-Cath*-transgenic, and *lh*-mutant blue catfish, *Ictalurus furcatus*. (A) A representative body shape from each genotype was photographed at 31 mph, scale bar = 15 cm. **(B)** A comparison of the BW at various ages.

P_1 founders were generated on 7th June 2020. Fish were kept separately in 60-L aquaria at a density of 2 fry/L until 5 months post hatch (5 mph), then pit-tagged and transferred to a flow-through tank (~500-L water) with a mix of these five genotypes (initial numbers of fish were 30, 30, 23, 12 and 9) and fed daily to satiation. The deformity was evaluated by multivariate ANOVA followed by the Hotelling test. Differences in BW among these four genotypes were compared using one-way ANOVA. *As-Cath*, alligator (*Alligator sinensis*) cathelicidin; *lh*, luteinizing hormone; WT, wild-type fish; LH^+_As-$Cath^-$, negative fish line, microinjected fish without *lh* mutation or *As-Cath* insertion; LH^-_As-$Cath^+$, on-target positive fish, the *As-Cath* transgene was inserted at the *lh* locus; LH^+_As-$Cath^+$, off-target positive fish, the *As-Cath* transgene was detected but not at the *lh* locus; ns, not significant (all $P > 0.05$).

Over the past few years, our team has been generating genetically engineered catfish with improved traits via CRISPR/Cas9-mediated genome editing, but the low hatching rate of embryos has made it challenging to obtain transgenic/gene-edited blue catfish. The reduction in embryo survival, partially due to microinjection and the majority due to artificial spawning and poor gamete quality, was more pronounced in blue catfish than in channel catfish, as reflected by higher mortality (> 85% *vs*. ~55%), severely limiting microinjection success. These findings suggest that blue catfish embryos are more fragile and sensitive to the physical damage caused by microinjection, and probably poor-quality gametes are used for the microinjection, resulting in heavy mortality. In this case, we would need to build a more robust microinjection procedure (Figure 13A). For the blue catfish brood stock, fortification of brood stock with protein-enriched aquafeed and forage fish (e.g., sunfish, minnows) for at least six months prior to spawning (Dunham and Elaswad, 2018) may be beneficial. Then, the optimal hormone type and dose to induce spawning should be determined using various ovulation-inducing hormones. Because of the erratic spawning behavior of blue catfish, artificial fertilization procedures have been adopted from those for channel catfish. Timing induction of ovulation, knowledge of the spawning, and optimization of the artificial spawning procedure for blue catfish are needed. Lastly, the concentration, motility, and velocity of fresh spermatozoa should be evaluated to detect the sperm quality to increase artificial fertilization rates. In addition to sperm quality, egg viability could be assessed by sampling using a trypan blue solution (Strober, 2001). Taking all these factors into consideration, the quality of fertilized eggs used for microinjection could be increased.

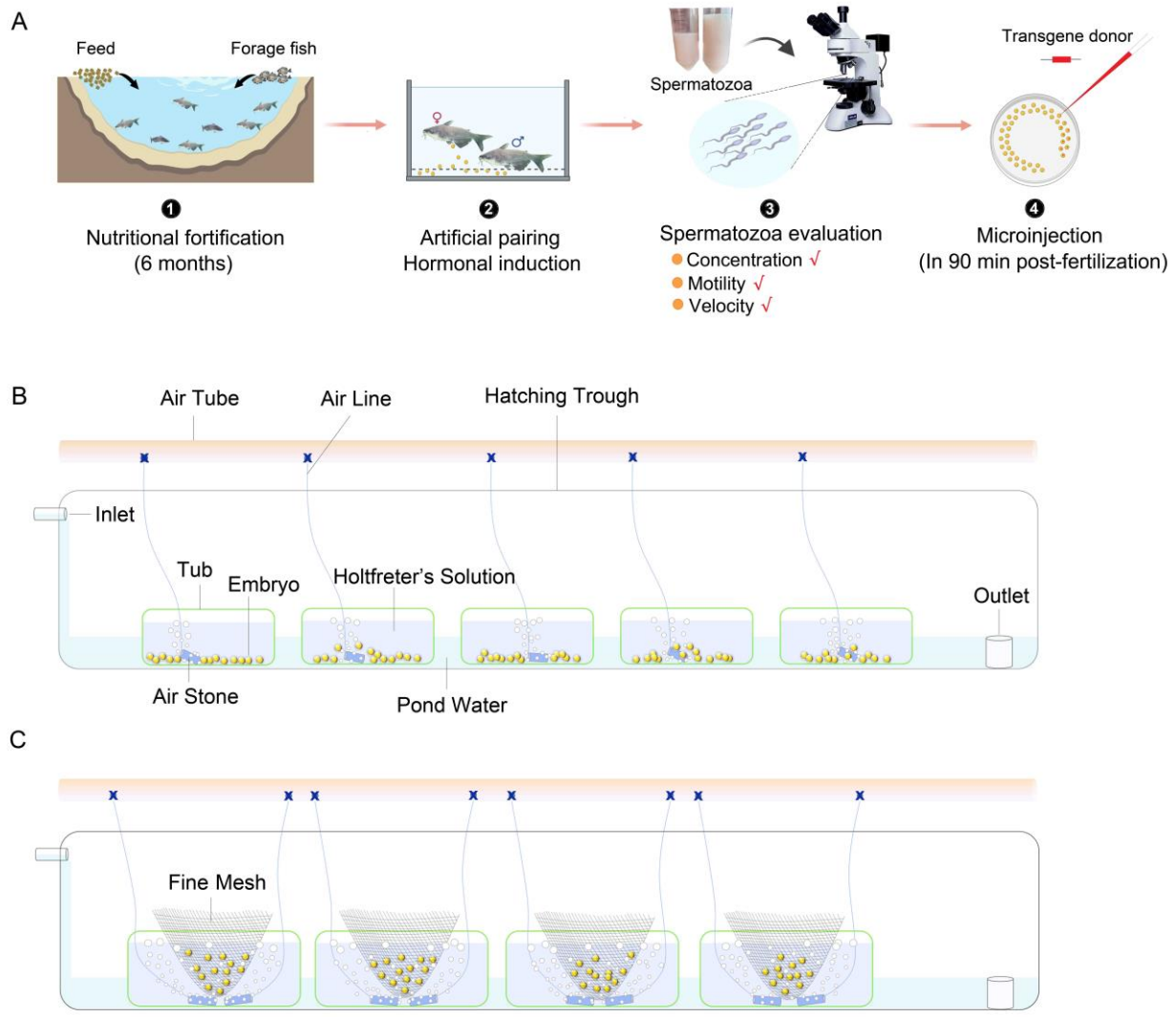


Figure 13. Tailored microinjection procedure and hatching apparatus for blue catfish (*Ictalurus furcatus*). (A) A procedure of the CRISPR/Cas9-based genome editing using microinjection in blue catfish. Step1: For six months prior to spawning, blue catfish brood would be fed with protein-enriched aquafeed and forage fish (e.g., sunfish, minnows) daily for nutrient enrichment. Step2: Various ovulation-inducing hormones would be applied to determine the optimal hormone type and dose. Step3: The concentration, motility and velocity of sperm should be evaluated to determine spermatozoa quality prior to fertilization. Step4: Microinjection with transgene donors should be accomplished within 90 min post-fertilization. **(B)** Schematic diagram of conventional hatching setup for blue catfish in this study. Ten 10-L tubs (200 embryos/tub, filled with 7 L of Holtfreter's solution) are placed at the bottom of the hatching trough in this setup. An air stone is provided for each tub to keep dissolved oxygen over 5 mg/L , and the trough is supplied with continuous flow-through pond water to keep all tubs at the same hatching temperature. Embryos usually remain stationary at the bottom of the tub. **(C)** Schematic illustration of the tailored hatching apparatus for blue catfish. In this setup, ten 15-L tubs (100 embryos/tub, filled with 10 L of Holtfreter's solution) would be placed at the bottom of a hatching trough. A fine metal mesh will be used to keep the embryos from touching the bottom of tub, and two air stones will be provided to gently roll the eggs. The trough will be supplied with continuous flow-through pond water to keep all tubs at the same hatching temperature.

Intriguingly, even in the non-injected group, a large number of decayed embryos were observed after 12 hours of fertilization. Therefore, another possibility is that inappropriate hatching methods lead to high embryo mortality. The hatching of the blue catfish embryos was conducted utilizing the conventional hatching method as that used for separated and manipulated channel catfish embryos (Figure 13B). For the conventional setup, ten 10-L tubs (200 embryos/tub, filled with 7 L of Holtfreter's solution) were placed at the bottom of a hatching trough. An air stone was provided for each tub to keep dissolved oxygen above 5 mg/L , and the trough was supplied with continuous flow-through pond water to keep all tubs at the same hatching temperature. Previously, a 45% hatching rate was obtained in channel catfish for this method (Simora et al., 2020; Wang et al., 2023b), whereas hatchability for similarly treated blue catfish embryos has typically not reached 15%. The absence of water flushing in the conventional method leads to the blue catfish embryos remaining stationary at the bottom of the tub, and they are more likely to decompose. Thus, a hypothetical tailored hatching methodology for the blue catfish was designed (Figure 13C). In the tailored configuration, ten 15-L tubs (100 embryos/tub, filled with 10 L of Holtfreter's solution) were placed at the bottom of a hatching trough. A fine metal mesh was used to keep the embryos from touching the bottom of the tub, and two air stones are provided to gently agitate and flush the embryos.

In addition, the slow maturation rate is another potential obstacle to the generation of transgenic blue catfish. Blue catfish have a sexual maturity period of 4-6 years (Dunham et al., 1994; Graham, 1999), which hinders the frequent acquisition of artificially fertilized eggs for microinjection. In addition to providing enough mature founders for spawning, an alternative strategy is adopting xenogenesis to shorten the reproductive cycle of blue catfish. Previous studies have shown that sterile fish as recipients can produce donor-derived gametes via the transplantation of undifferentiated germ cells derived from diploid fish (donors) (Wong et al., 2011; Perera et al., 2017; Shang et al., 2018; Hettiarachchi et al., 2022). Hypothetically, triploid channel catfish parents (reproductively sterile fish) are produced using a high-pressure method. Then blue catfish stem cells are transplanted into these triploid channel catfish, and once these triploid channel catfish mature, they have the capability to produce donor-derived gametes. If xenogenic triploid channel catfish ♀ (transplanted with blue catfish stem cells) are mated with xenogenic triploid channel catfish ♂ (transplanted with blue stem cells), blue catfish

sperm/eggs should be obtained from these xenogenic males/females in two-three years. This method takes advantage of the short reproductive cycle of channel catfish by allowing blue catfish, which have a long reproductive cycle, to reach sexual maturity and gametogenesis in a short amount of time. The abovementioned strategies could be applied to improve transgenic embryo survival, providing new approaches to achieve the microinjection of blue catfish embryos on a larger scale.

Prior to the introduction of ZFN, TALEN, and CRISPR tools, transgenic fish were created by randomly integrating vector-engineered plasmids containing targeted genes via microinjection or electroporation (Zhu et al., 1985; Chen and Powers, 1990; Dunham et al., 2002; Mao et al., 2004). Although effective in integrating foreign genes into fish genomes, these strategies risked untraceable inserted DNA fragments and disruption of functional genes (Song and Stieger, 2017). The CRISPR system has the ability to precisely drive transgene integration at specific loci. This greatly reduces the incidence of unwanted insertions. The use of linear dsDNA as a donor template usually results in a lower integration efficiency compared to a plasmid or a double-cut plasmid (Zhang et al., 2017; Simora et al., 2020). We generated knock-ins using a linear dsDNA donor with an insertion rate of 16.7% in surviving fingerlings. This result was comparable to the precise insertion rate of 10–28.6% in channel catfish and mice (Gu et al., 2018; Simora et al., 2020; Wang et al., 2023b). In the present study, a higher integration rate of 24.53% was observed using the double-cut plasmid donor compared to the linear dsDNA, which is consistent with the efficiency in medaka (25–27%) (Murakami, et al., 2017). To improve the efficiency of the transgene, a cocktail design of donor plasmid was used in channel catfish, combining HDR, NHEJ, and microhomology-mediated end joining (MMEJ) pathways by using multiple homology arms of different lengths. This new approach has achieved a high integration rate of 28.1% (Xing et al., 2023), which could be applied in blue catfish and other fish species to improve the efficiency of transgenesis.

Although online off-target tools were used in this study to screen for off-target-free sgRNAs, we detected the off-target insertions of the *As-Cath* transgene and *lh* mutations in both CRISPR/Cas9-mediated systems. This is mainly due to *in silico* predictions failing to predict true off-target cleavages *in vivo* (Ran et al., 2013; Heigwer et al., 2014). Moreover, off-target events are more frequent within *in vivo* animal experiments than within *in vitro* cellular experiments

(Zhang et al., 2015). Nevertheless, our results displayed that HDR-mediated KI was dominant compared to NHEJ-mediated knockout (16.67% *vs*. 5.88% and 24.53% *vs*. 8.33%) in the presence of donor templates, which was found at different frequencies in channel catfish using either dsDNA or ssODN donors (Xing et al., 2022; Wang et al., 2023b). The ssODN-assisted end joining approach resulted in a high integration rate of 17.6% (3/17) in the rat (Yoshimi et al., 2016), but this efficiency was variable in different species. For example, ssODN-mediated KI showed a low integration rate of 1–4% when different ssODNs were used in zebrafish (Boel et al., 2018). Recently, Xing et al. (2022) demonstrated that KI efficiency could reach as high as 37.5% (81/216) in channel catfish using two ssODNs as DNA repair templates. However, this high integration efficiency was also accompanied by a high off-target effect. These findings indicated that the CRISPR/Cas9-mediated system coupled with a double-cut plasmid donor was more efficient than the one using a dsDNA or ssODN donor for catfish. Despite the inevitability of off-target cleavage, it is promising that newly developed CRISPR/Cas-based tools with high on-target efficiency in model animals or human cells, such as *Easi*-CRISPR, C-CRISPR (Quadros et al., 2017), CRISPR/Cas9-HITI (Kelly et al., 2021), and dCas9-SSAP (Wang et al., 2022c) can be applied to non-model fish, aquaculturally relevant fish in the future to reduce potential off-target effects.

Genetic mosaicism has been and will continue to be another obstacle to the practical application of CRISPR/Cas9-mediated genome editing. To date, there is no effective strategy to completely eliminate mosaicism (Mehravara et al., 2019). In this study, we were unable to generate 100 % mosaic-free individuals, demonstrating that mosaicism is common and unavoidable in fish and other animals. Essentially, mosaicism of transgenic/gene-edited organisms is common in the case of embryo-based editing. Mosaic animals have been observed in mice, rats, zebrafish, and channel catfish with varying frequencies (Auer et al., 2014; Mehravara et al., 2019; Simora et al., 2020). CRISPR/Cas9-induced mosaicism has greatly alleviated this problem as mosaicism is much less compared to the original random gene transfer procedures (Dunham, 2023). We evaluated the expression of the *As-Cath* transgene in nine tissues from three on-target positive P_1 founders, finding that one female had no transgene expression in the gonad. Thus, 67% of the brood stock should be able to produce offspring. Much lower transmission rates were obtained in the past with traditional approaches.

Antimicrobial peptide genes (AMGs) as transgenes can significantly improve the survival rate of channel catfish (Dunham et al., 2002; Abass et al., 2022), grass carp (Zhong et al., 2002; Mao et al., 2004), rainbow trout (Chiou et al., 2013) after pathogen infection in comparison to WT fish. The survival rate can be increased up to 100%. For example, cecropin transgenic channel catfish had a 100% survival rate, whereas only 27% of WT fish survived after infection with *Flavobacterium columnare* (Dunham et al., 2002). Our results showed that *As-Cath*-transgenic blue catfish exhibited boosted resistance to *F. covae* bacteria with a twofold higher survival rate compared to WT individuals (80% *vs*. 30%). Earlier studies have illustrated AMG transgenes contribute to the modulation of innate immunity in fish by improving lysozyme activity, activating immune/anti-inflammatory factors, and facilitating the synergistic expression of transgenic AMGs and immune-related genes (Wang and Cheng, 2023). Additionally, transgenic AMGs can inhibit the proliferation of pathogens to increase survival (Mao et al., 2004; Wang and Cheng, 2023).

Genetically modified (GM) animals, including transgenic/gene-edited fish, are not universally embraced by consumers due primarily to concerns regarding food safety and environmental risk. The current study contributes not only to the improvement of disease resistance of blue catfish through transgenesis but also to the genetic achievement of reproductive confinement of these GM fish to avoid introgression caused by escapees mating with wild populations. The available data indicates that transgenic fish would likely not be fit in the natural environment. However, the conservative and ethical approach is to prevent the establishment of transgenic fish in the wild to prevent breeding and introgression that could genetically contaminate native/natural wild populations. The best strategy to prevent establishment of or introgression of transgenic genotypes in the wild is genetic sterilization of transgenic fish with the intervention of man the only option for restoring fertility. The production of LH^{-}_As-$Cath^{+}$ genetic lines deficient for *lh* gene is a good strategy for achieving this goal. Previous studies have shown that *lh*-mutant female zebrafish and channel catfish were sterile (Chu et al., 2014; Qin, 2019). More recently, our work showed that replacing the *As-Cath* at the *lh* locus of the channel catfish genome can produce sterilized fish lines, and fertility can be restored by hormone therapy (Wang et al., 2023b). Theoretically, LH^{-}_As-$Cath^{+}$ blue catfish are potentially reproductively confined

preventing any potential genetic contamination of wild fish populations. Future research should evaluate the effect of *lh*-knockout on the gonadal development via performing tissue sections, measuring the LH hormone changes and *lh* gene expression, evaluating the feasibility of fertility restoration of P_1 mutants via hormone therapy, thus, allowing reproductive confinement of these gene-edited transgenic fish and eliminating the chance of gene flow while allowing the production of disease resistant fingerlings for culture application. We hypothesize that this reversible sterilization will be successful with blue catfish as *lh* knockout followed by hormone therapy sterilized and then restored fertility in the closely related species, channel catfish (Qin, 2019).

Transgenic and gene edited fish now appear on the table as growth hormone gene-transgenic Atlantic salmon (*Salmo salar*; AquAdvantage salmon) (Ledford, 2015; Waltz, 2017), leptin-edited tiger puffer (*Takifugu rubripes*), and myostatin-deficient red sea bream (*Pagrus major*; https://doi.org/10.1038/s41587-021-01197-8) have been approved for consumption and are on the market. Creating transgenic and gene edited fish with characteristics valued by producers and consumers is only the first step to impact. The larger challenge for commercializing GM fish is public education (Dunham, 2023). Now cathelicidin transgenic, *lh* knockout blue catfish and cathelicidin transgenic, *lh* knockout channel catfish (Wang et al., 2023b) have been produced. Future research should evaluate the performance and sterility of the hybrid between LH^{-}_As-$Cath^{+}$ channel catfish females and LH^{-}_As-$Cath^{+}$ blue catfish males. The WT hybrid is the best genotype for pond culture of ictalurid catfish in the US catfish industry, and now accounts for 60–70% of catfish production (Torrans and Ott, 2018). The WT hybrid exhibits heterobeltiosis for disease resistance (Dunham et al., 2008). However, they do not have total disease resistance and significant mortality can occur. Combining interspecific hybridization with cathelicidin transgenesis and knockout of reproductive genes such as *lh* may result in a catfish with near total disease resistance that is sterile with no possible impact on native gene pools.

Our work provides the aquaculture community with another new disease-resistant genetic line of catfish, demonstrating that transgenic fish possessing AMG transgenes hold great promise for disease reduction in aquaculture. Finally, the use of AMG transgenic fish offers a promising

approach for ecologically sustainable aquaculture that can effectively address the growing challenges posed by antimicrobial resistance.

5. References

Abass NY, Simora RMC, Wang J, Li S, Xing D, Coogan M, et al. Response of cecropin transgenesis to challenge with *Edwardsiella ictaluri* in channel catfish *Ictalurus punctatus*. *Fish Shellfish Immunol*. 2022;126:311–317. doi:10.1016/j.fsi.2022.05.050

Abdel-Wahab MM, Taha NM, Lebda MA, Elfeky MS, Abdel-Latif HMR. Effects of bovine lactoferrin and chitosan nanoparticles on serum biochemical indices, antioxidative enzymes, transcriptomic responses, and resistance of Nile tilapia against *Aeromonas hydrophila*. *Fish Shellfish Immunol*. 2021;111:160–169. doi:10.1016/j.fsi.2021.01.017

Armstrong JB, Malacinski GM. Developmental biology of the axolotl. Oxford University Press, New York, 1989.

Auer TO, Duroure K, De Cian A, Concordet J-P, Del Bene F. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res*. 2014;24:142–153. doi:10.1101/gr.161638.113

Bae S, Park J, Kim J-S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*. 2014;30:1473–1475. doi:10.1093/bioinformatics/btu048

Boel A, De Saffel H, Steyaert W, Callewaert B, De Paepe A, Coucke PJ, et al. CRISPR/Cas9-mediated homology-directed repair by ssODNs in zebrafish induces complex mutational patterns resulting from genomic integration of repair-template fragments. *Dis Model Mech.* 2018;11:dmm035352. doi:10.1242/dmm.035352

Bondad-Reantaso MG, Lavilla-Pitogo C, Lopez MML. Guidance in development of aquaculture component of a national action plan on antimicrobial resistance. *Asian Fish Sci.* 2020;33:119–124. doi:10.33997/j.afs.2020.33.S1.017

Cabello FC. Heavy use of prophylactic antibiotics in aquaculture: a growing problem for human and animal health and for the environment. *Environ Microbiol*. 2006;8:1137–1144. doi:10.1111/j.1462-2920.2006.01054.x

Chen TT, Powers DA. Transgenic fish. *Trends Biotechnol*. 1990;8:209–215. doi:10.1016/0167-7799(90)90178-Z

Chen Y, Cai S, Qiao X, Wu M, Guo Z, Wang R. As-CATH1-6, novel cathelicidins with potent antimicrobial and immunomodulatory properties from *Alligator sinensis*, play pivotal roles in host antimicrobial immune responses. *Biochem J*. 2017;474:2861–2885.

doi:10.1042/BCJ20170334

Chiou PP, Chen MJ, Lin C-M, Khoo J, Larson J, Holt R, et al. Production of homozygous transgenic rainbow trout with enhanced disease resistance. *Mar Biotechnol*. 2014;16:299–308. doi:10.1007/s10126-013-9550-z

Chu L, Li J, Liu Y, Hu W, Cheng CHK. Targeted gene disruption in zebrafish reveals noncanonical functions of LH signaling in reproduction. *Mol Endocrinol*. 2014;28:1785–1795. doi:10.1210/me.2014-1061

Coogan M, Alston V, Su B, Khalil K, Elaswad A, et al. CRISPR/Cas-9 induced knockout of myostatin gene improves growth and disease resistance in channel catfish (*Ictalurus punctatus*). *Aquaculture*. 2022;557:738290. doi:10.1016/j.aquaculture.2022.738290

Dunham RA, Hyde C, Masser M, Plumb JA, Smitherman RO, Perez R, et al. Comparison of culture traits of channel catfish, *Ictalurus punctatus*, and blue catfish *I. furcatus*. *J Appl Aquac*.1994;3:257–268. doi:10.1300/J028v03n03_04

Dunham RA, Elaswad, A. Catfish Biology and Farming. *Annu Rev Anim Biosci*. 2018;6:305–325. doi:10.1146/annurev-animal-030117-014646

Dunham RA. Aquaculture and Fisheries Biotechnology: Genetic Approaches. 3rd Edition. CABI Publishing, Wallingford, UK, 2023;613.

Dunham RA, Umali GM, Beam R, Kristanto AH, Trask M. Comparison of production traits of NWAC103 channel catfish, NWAC103 channel catfish × blue catfish hybrids, Kansas Select 21 channel catfish, and blue catfish grown at commercial densities and exposed to natural bacterial epizootics. *North Am J Aquacult*. 2008;70:98–106. doi:10.1577/A07-006.1

Dunham RA, Warr GW, Nichols A, Duncan PL, Argue B, Middleton D, et al. Enhanced bacterial disease resistance of transgenic channel catfish *Ictalurus punctatus* possessing cecropin genes. *Mar Biotechnol*. 2002;4:338–344. doi:10.1007/s10126-002-0024-y

Elaswad A, Khalil K, Ye Z, Liu Z, Liu S, Peatman E. Effects of CRISPR/Cas9 dosage on TICAM1 and RBL gene mutation rate, embryonic development, hatchability and fry survival in channel catfish. *Sci Rep*. 2018;8:16499. doi:10.1038/s41598-018-34738-4

FAO. The FAO action plan on antimicrobial resistance 2021–2025. 2021. doi:10.4060/cb5545en

FDA. Approved Aquaculture Drugs. 2022. https://www.fda.gov/animal-veterinary/aquaculture/approved-aquaculture-drugs

Gen K, Yamaguchi S, Okuzawa K, Kumakura N, Tanaka H, Kagawa H. Physiological roles of FSH and LH in red seabream, *Pagrus major*. *Fish Physiol Biochem*. 2003;28:77–80. doi:10.1023/B:FISH.0000030480.97947.ba

Gosh K, Hanson TR, Drescher D, Robinson D, Bugg W, Chatakondi N, et al. Economic effect of hybrid catfish (channel catfish ♀ × blue catfish ♂) growth variability on traditional and intensive production systems. *North Am J Aquacult*. 2022;84:25–41. doi:10.1002/naaq.10211

Graham K. A review of the biology and management of blue catfish. American Fisheries Society. 1999;24:37–49. https://fisheries.org/docs/pub_sympsample.pdf

Gu B, Posfai E, Rossant J. Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. *Nat Biotechnol*. 2018;36:632–637. doi:10.1038/nbt.4166

Hancock RE. Cationic peptides: effectors in innate immunity and novel antimicrobials. *Lancet Infect Dis*. 2001;1:156–164. doi:10.1016/S1473-3099(01)00092-5

Heigwer F, Kerr G, Boutros M. E-CRISP: fast CRISPR target site identification. *Nat Methods*. 2014;11:122–123. doi:10.1038/nmeth.2812

Hettiarachchi DU, Alston VN, Bern L, Shang M, Wang J, Xing D, et al. Producing xenogenic channel catfish, *Ictalurus punctatus* with cryopreserved testes and ovarian tissues of blue catfish, *I. furcatus*. *Aquaculture*. 2022;561:738691. doi:10.1016/j.aquaculture.2022.738691

Hsieh J-C, Pan C-Y, Chen J-Y. Tilapia hepcidin (TH)2-3 as a transgene in transgenic fish enhances resistance to *Vibrio vulnificus* infection and causes variations in immune-related genes after infection by different bacterial species. *Fish Shellfish Immunol*. 2010;29:430–439. doi:10.1016/j.fsi.2010.05.001

Karunasagar I. Complexities involved in source attribution of antimicrobial resistance genes found in aquaculture products. *Asian Fish Sci*. 2020;33:16–21. doi:10.33997/j.afs.2020.33.S1.003

Kelly JJ, Saee-Marand M, Nyström NN, Evans MM, Chen Y, Martinez FM, et al. Safe harbor-targeted CRISPR-Cas9 homology-independent targeted integration for multimodality reporter gene-based cell tracking. *Sci Adv*. 2021;7:eabc3791. doi:10.1126/sciadv.abc3791

Khalil K, Elayat M, Khalifa E, Daghash S, Elaswad A, Miller M, et al. Generation of myostatin gene-edited channel catfish (*Ictalurus punctatus*) via zygote injection of CRISPR/Cas9 system. *Sci Rep*. 2017;7:7301. doi:10.1038/s41598-017-07223-7

Ledford H. Salmon approval heralds rethink of transgenic animals. *Nature*. 2015;527, 417–418. doi:10.1038/527417a

Lin X, Chen W, Lin S, Luo L. Effects of dietary cecropin on growth, non-specific immunity and disease resistance of tilapia (*Oreochromis niloticus* × *O. aureus*). *Aquac Res*. 2015;46:2999–3007. doi:10.1111/are.12457

Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*. 2001;25:402–408. doi:10.1006/meth.2001.1262

Mao W, Wang Y, Wang W, Wu B, Feng J, Zhu Z. Enhanced resistance to *Aeromonas hydrophila* infection and enhanced phagocytic activities in human lactoferrin-transgenic grass carp (*Ctenopharyngodon idellus*). *Aquaculture*. 2004;242:93–103. doi:10.1016/j.aquaculture.2004.07.020

Mehravara M, Shirazia A, Nazaric M, Banand M. Mosaicism in CRISPR/Cas9-mediated genome editing. *Dev Biol*. 2019;445:156–162. doi:10.1016/j.ydbio.2018.10.008

Mi R, Li X, Sun Y, Wang Q, Tian B, Ma S, et al. Effects of microbial community and disease resistance against *Vibrio splendidus* of Yesso scallop (*Patinopecten yessoensis*) fed supplementary diets of tussah immunoreactive substances and antimicrobial peptides. *Fish Shellfish Immunol*. 2022;121:446–455. doi:10.1016/j.fsi.2021.10.006

Mookherjee N, Anderson MA, Haagsman HP, Davidson DJ. Antimicrobial host defence peptides: functions and clinical potential. *Nat Rev Drug Discov*. 2020;19:311–332. doi:10.1038/s41573-019-0058-8

Morita H, Taimatsu K, Yanagi K, Kawahara A. Exogenous gene integration mediated by genome editing technologies in zebrafish. *Bioengineered*. 2017;8:287–295. doi:10.1080/21655979.2017.1300727

Mosimann C, Kaufman CK, Li P, Pugach EK, Tamplin OJ, Zon LI. Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish. *Development*. 2011;138:169–177. doi:10.1242/dev.059345

Murakami Y, Ansai S, Yonemura A, Kinoshita M. An efficient system for homology-dependent targeted gene integration in medaka (*Oryzias latipes*). *Zool Lett*. 2017;3:10. doi:10.1186/s40851-017-0071-x

Naylor RL, Hardy RW, Buschmann AH, Bush SR, Cao L, Klinger DH, et al. A 20-year retrospective review of global aquaculture. *Nature*. 2021;591:551–563. doi:10.1038/s41586-021-03308-6

Perera DA, Alsaqufi A, Shang M, Wade DC, Su B, Elaswad A, et al. Xenogenesis-production of channel catfish × blue catfish hybrid progeny by fertilization of channel catfish eggs with sperm from triploid channel catfish males with transplanted blue catfish germ cells. *North Am J Aquacult*. 2017;79:61–74. doi:10.1080/15222055.2016.1221008

Pflieger WL. The fishes of Missouri. Missouri Department of Conservation, Jefferson City. 1997.

Priya TJ, Kappalli S. Modern biotechnological strategies for vaccine development in aquaculture–Prospects and challenges. *Vaccine*. 2022;40:5873–5881. doi:10.1016/j.vaccine.2022.08.075

Qin G. Gene Editing and Hormone Therapy to Control Reproduction in Channel Catfish, *Ictalurus punctatus*. Doctoral dissertation. 2019. Auburn University, AL, USA. http://hdl.handle.net/10415/6690

Qin Z, Li Y, Su B, Cheng Q, Ye Z, Perera DA, et al. Editing of the luteinizing hormone gene to sterilize channel catfish, *Ictalurus punctatus*, using a modified Zinc Finger Nuclease technology with electroporation. *Mar Biotechnol*. 2016;18:255–263. doi:10.1007/s10126-016-9687-7

Qiu P, Shandilya H, D'Alessio JM, O'Connor K, Durocher J, Gerard GF. Mutation detection using Surveyor™ nuclease. *BioTechniques*. 2004;36;702–707. doi:10.2144/04364PF01

Quadros RM, Miura H, Harms DW, Akatsuka H, Sato T, Aida T, et al. Easi-CRISPR: a robust method for one-step generation of mice carrying conditional and insertion alleles using long ssDNA donors and CRISPR ribonucleoproteins. *Genome Biol*. 2017;18:92–106. doi:10.1186/s13059-017-1220-4

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013;8:2281–308. doi:10.1038/nprot.2013.143

Sarmasik A, Warr G, Chen TT. Production of transgenic medaka with increased resistance to bacterial pathogens. *Mar Biotechnol*. 2002;4:310–322. doi:10.1007/s10126-002-0023-z

Shang M, Su B, Perera DA, Alsaqufi A, Lipke EA, Cek S, et al. Testicular germ line cell identification, isolation, and transplantation in two North American catfish species. *Fish Physiol Biochem*. 2018;44:717–733. doi:10.1007/s10695-018-0467-3

Simora RMC, Xing D, Bangs MR, Wang W, Ma X, Su B, et al. CRISPR/Cas9-mediated knock-in of alligator cathelicidin gene in a non-coding region of channel catfish genome. *Sci Rep*. 2020;10:22271. doi:10.1038/s41598-020-79409-5

Sommerset I, Krossøy B, Biering E, Frost P. Vaccines for fish in aquaculture. *Expert Rev Vaccines.* 2005;4:89–101. doi:10.1586/14760584.4.1.89

Song F, Stieger K. Optimizing the DNA donor template for homology-directed repair of double-strand breaks. *Mol Ther Nucleic Acids*. 2017;7:53–60. doi:10.1016/j.omtn.2017.02.006

Strober W. Trypan blue exclusion test of cell viability. *Curr Protoc Immunol*. 2001;21:1–2. doi:10.1002/0471142735.ima03bs21

Su B, Perera DA, Zohar Y, Abraham E, Stubblefield J, Fobes M, et al. Relative effectiveness of carp pituitary extract, luteininzing hormone releasing hormone analog (LHRHa) injections and LHRHa implants for producing hybrid catfish fry. *Aquaculture*. 2013;372–375:133–136. doi:10.1016/j.aquaculture.2012.10.001

Torrans L, Ott B. Effect of grading fingerling hybrid catfish (♀ channel catfish × ♂ blue catfish) on growth, production, feed conversion, and food fish size distribution. *North Am J Aquacult*.

2018;80:187–192. doi:10.1002/naaq.10024

Waltz E. First genetically engineered salmon sold in Canada. *Nature*. 2017;548:148. doi:10.1038/nature.2017.22116

Wang G, Li X, Wang Z. APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res*. 2016;44:D1087–1093. doi:10.1093/nar/gkv1278

Wang J, Su B, Dunham RA. Genome-wide identification of catfish antimicrobial peptides: A new perspective to enhance fish disease resistance. *Rev Aquac*. 2022a;14: 2002–2022. doi:10.1111/raq.12684

Wang H, Su B, Butts IAE, Dunham RA, Wang X. Chromosome-level assembly and annotation of the blue catfish *Ictalurus furcatus*, an aquaculture species for hybrid catfish reproduction, epigenetics, and heterosis studies. *Gigascience*. 2022b;11:giac070. doi:10.1093/gigascience/giac070

Wang J, Cheng Y. Harnessing antimicrobial peptide genes to expedite disease-resistant enhancement in aquaculture: Transgenesis and CRISPR/Cas9-mediated genome editing. *Rev Aquac*. doi:10.1101/2023.01.05.522886

Wang C, Qu Y, Cheng JKW, Hughes NW, Zhang Q, Wang M, et al. dCas9-based gene editing for cleavage-free genomic knock-in of long sequences. *Nat Cell Biol*. 2022c;24:268–278. doi:10.1038/s41556-021-00836-1

Wang J, Wilson AE, Su B, Dunham RA. Functionality of dietary antimicrobial peptides in aquatic animal health: Multiple meta-analyses. *Anim Nutr*. 2023a;12:200–214. doi:10.1016/j.aninu.2022.10.001

Wang J, Su B, Xing D, Bruce TJ, Li S, Bern L, et al. Generation of eco-friendly channel catfish, *Ictalurus punctatus*, harboring alligator cathelicidin gene with robust disease resistance by harnessing different CRISPR/Cas9-mediated systems. 2023b. doi:10.1101/2023.01.05.522889

Wang J, Su B, Bruce TJ, Wise AL, Zeng P, Cao G, et al. CRISPR/Cas9 microinjection of transgenic embryos enhances the dual-gene integration efficiency of antimicrobial peptide genes for bacterial resistance in channel catfish, *Ictalurus punctatus*. *Aquaculture*. 2023c. doi:10.1016/j.aquaculture.2023.739725

Watakabe I, Hashimoto H, Kimura Y, Yokoi S, Naruse K, Higashijima S. Highly efficient generation of knock-in transgenic medaka by CRISPR/Cas9-mediated genome engineering. *Zool Lett*. 2018;4:3. doi:10.1186/s40851-017-0086-3

Welker TL, Lim C, Yildirim-Aksoy M, Klesius PH. Dietary bovine lactoferrin increases resistance of juvenile channel catfish, *Ictalurus punctatus*, to enteric septicemia. *J World Aquac Soc*. 2010;41:28–39. doi:10.1111/j.1749-7345.2009.00330.x

WHO. Guidelines on use of medically important antimicrobials in food-producing animals. Geneva: World Health Organization 2017;1–67. https://www.who.int/publications-detail-redirect/9789241550130

Wong T-T, Saito T, Crodian J, Collodi P. Zebrafish germline chimeras produced by transplantation of ovarian germ cells into sterile host larvae. *Biol Reprod*. 2011;84:1190–1197. doi:10.1095/biolreprod.110.088427

Xing D, Su B, Li S, Bangs M, Creamer D, Coogan M, et al. CRISPR/Cas9-Mediated transgenesis of the masu salmon (*Oncorhynchus masou*) elovl2 gene improves n-3 fatty acid content in channel catfish (*Ictalurus punctatus*). *Mar Biotechnol*. 2022;24:513–523. doi:10.1007/s10126-022-10110-6

Xing D, Shang M, Li S, Wang W, Hasin T, Hettiarachchi D, et al. CRISPR/Cas9-mediated precision integration of *fat-1* and *fat-2* from *Caenorhabditis elegans* at long repeated sequence in channel catfish (*Ictalurus punctatus*) and the impact on n-3 fatty acid level. *Aquaculture*. 2023;567:739229. doi:10.1016/j.aquaculture.2023.739229

Yazawa R, Hirono I, Aoki T. Transgenic zebrafish expressing chicken lysozyme show resistance against bacterial diseases. *Transgenic Res*. 2006;15:385–391. doi:10.1007/s11248-006-0009-0

Yoshimi K, Kunihiro Y, Kaneko T, Nagahora H, Voigt B, et al. ssODN-mediated knock-in with CRISPR-Cas for large genomic regions in zygotes. *Nat Commun*. 2016;7:10431. doi:10.1038/ncomms10431.

Zhang J-P, Li X-L, Li G-H, Chen W, Arakaki C, Botimer GD, et al. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol*. 2017;18:35. doi:10.1186/s13059-017-1164-8

Zhang X-H, Tee LY, Wang X-G, Huang Q-S, Yang S-H. Off-target effects in CRISPR/Cas9-mediated genome engineering. *Mol Ther Nucleic Acids* 2015;17:e264. doi:10.1038/mtna.2015.37

Zhong J, Wang Y, Zhu Z. Introduction of the human lactoferrin gene into grass carp (*Ctenopharyngodon idellus*) to increase resistance against GCH virus. *Aquaculture*. 2002;214:93–101. doi:10.1016/S0044-8486(02)00395-2

Zhu Z, Li G, He L, Chen S. Novel gene transfer into the fertilized eggs of goldfish (*Carassius auratus* L. 1758). *J Appl Ichthyol*. 1985;1:31–34. doi:10.1111/j.1439-0426.1985.tb00408.x

CHAPTER FOUR

CRISPR/Cas9 microinjection of transgenic embryos enhances the dual-gene integration efficiency of antimicrobial peptide genes for bacterial resistance in channel catfish, *Ictalurus punctatus*

Abstract

Integrating a vector-engineered antimicrobial peptide gene (AMG) into the fish genome effectively modulated the innate immune system and increased resistance to infectious disease in channel catfish (*Ictalurus punctatus*). CRISPR/Cas9-assisted microinjection of cecropin (*Cec*) and cathelicidin (*Cath*) was employed to create dual-AMG integrated (*_Cec^{+}/*_$Cath^{+}$) transgenic embryos with high integration rates. Additionally, a univariate-multiple logit regression model was fitted to determine the synergistic expression of transgenes and endogenous AMGs in the head kidney post-bacterial infection. Transgenic-embryo-based genome editing significantly increased the efficiency of dual-AMG integration from 17.6% to 37.3%. The survival rate of single-AMG (50% *vs.* 20%, $P = 0.023$) and dual-AMG (70% *vs.* 20%, $P = 0.005$) integrated fish was dramatically higher than that of wild-type fish (20%) following *Edwardsiella ictaluri* challenge. More dual-AMG fry survived than expected based on integration and inheritance rates of single-AMG transgenics compared to other genotypes. Logistic regression (LR) analysis indicated that individual body weight and gender did not affect survival, while the transgenes *Cec* and *Cath* contributed directly to the survival during the bacterial infection. Furthermore, transgenes enhanced fish disease resistance by regulating the expression of *TCP* and *NK-lysin* genes. This study demonstrates that it is promising to generate dual-gene integrated genetic lines with a high integration efficiency by adopting transgenic-embryo-based CRISPR/Cas9-mediated genome editing, and an LR model is feasible for assessing the synergistic effects of gene expression.

Keywords: Genome editing; knock in; antimicrobial peptide; disease resistance; logistic regression; aquaculture

1. Introduction

Over the last decade, versatile applications of CRISPR/Cas9-mediated genome editing have sparked a great interest in the life science community in full swing to improve the favorable traits due to its high on-target efficiency (Gratacap et al., 2019; Blix et al., 2021). In aquaculture, some consumer-focused characteristics of fish, including muscle mass (Coogan et al., 2022ab; Shahi et al., 2022), reproduction (Jiang et al., 2017), coloration (Cal et al., 2019; Mandal et al., 2020), disease resistance (Ma et al., 2018) and omega-3 fatty acids (Xing et al., 2022ab) can be altered by disrupting or introducing a gene of interest through genome editing or transgenesis.

The high level of industrialization and intensification of aquaculture is bringing fish disease outbreaks to a peak, and an increasing number of alternative mitigation strategies are being investigated to control the potential environmental impacts, and antimicrobial resistance brought on by antibiotic abuse (FAO, 2021). As promising substitutes for antibiotics and anti-parasitics, antimicrobial peptides (AMPs) are well recognized by immunologists for their highly effective spectrum bactericidal and anti-inflammatory capabilities (Hancock et al., 2016). AMPs are used not only in developing human anti-cancer drugs (Mookherjee et al., 2020) but are also effective as feed additives in the prevention and pharmacotherapy of animal diseases (Rodrigues et al., 2021; Silveira et al., 2021; Wang et al., 2022). For instance, cecropin and lactoferrin as feed supplements have positive effects on the improvement of immune responses in livestock and fish (Dai et al., 2020; Abdel-Wahab et al., 2021; Wang et al., 2023). Nevertheless, the marketing of AMPs as additives is severely constrained by their high demand/cost and brief permanence.

Alternatively, the researchers used genetic engineering to introduce an antimicrobial peptide gene (AMG), which can encode an AMP, into the host to allay these drawbacks (Wang et al., 2022). Integrating foreign AMGs into the genome of targeted species can effectively combat various pathogens through the overexpression of AMPs. A typical pioneering study comes from Dunham et al. (2002), who successfully transferred one copy of the cecropin gene into the genome of channel catfish (*Ictalurus punctatus*) to establish a highly disease-resistant catfish germline, and its viability was supported by subsequent studies (Sarmasik et al., 2002; Chiou et al., 2014; Elaswad et al., 2019; Abass et al., 2022). Previous studies have revealed that AMG-integrated fish can reduce pathogen load (Yazawa et al., 2006; Hsieh et al., 2010). However,

since these AMGs are randomly integrated into the target genome, it is challenging to track down precise mutations. This uncertainty poses potential concerns for the regulatory assessment and future commercialization of gene-edited/transgenic fish. CRISPR/Cas9-mediated genome editing can produce multiple mutations/integrations through one system (Yang et al., 2013; Cong et al., 2013; Ota et al., 2014), allowing for combinations of different trait enhancements or double enhancements of the same trait. In this vein, inducing multiple AMGs at different loci could theoretically achieve more robust improvement in disease resistance compared to a single AMG integration.

A potential complication of the effect of AMG transgene insertion is the epistatic interaction with native immune-related genes. AMG transgenesis can increase lysozyme activity (Mao et al., 2004) and induce the expression of immune-related genes (Hsieh et al., 2010; Peng et al., 2010) to increase survival after pathogen infections. Foreign AMG has induced the expression of immune-related genes or endogenous AMGs, and there may be potential synergistic effects. Published reports demonstrated that AMG-transgenic fish tend to induce a more robust immune response than non-transgenic fish (Wang et al., 2022). Although the expression profiles of immune-associated genes are frequently studied and enumerated independently of the integrated-AMGs, this is not conducive to exploring and revealing the potential synergistic expression of these two types of genes. In addition, the mRNA levels of detected genes are tissue- and time-specific (Chiou et al., 2014), making this approach tedious and time-consuming for investigating the co-expression of exogenous AMGs and immune-related genes. In pathogen-challenge experiments, AMG-integrated individuals also show decreased mortality and large fluctuations in mRNA levels (Lo et al., 2014). If a list of immune responsive genes are perturbed and interrogated, it has not been possible to determine the extent to which exogenous AMGs or immune-related genes contribute to the survival of AMG-transgenic fish.

Logistic regression (LR) is a machine learning method for solving binary/multiple classification problems to estimate the likelihood of observed events of interest (Sperandei, 2014) or exploring the association between explanatory factors and categorical outcomes (Lee, 1986), making it useful in clinical research and epidemiological studies (Ostir and Uchida, 2000). An LR model might include both continuous and categorical explanatory variables, allowing researcher to

interpret raw data (Bender and Grouven, 1997; Ostir et al., 2000). Recently, we used LR models in fish challenge experiments to determine the effect of transgenes on survival and gene expression. LR model helps examine the association of all variables, including binary outcome (die or live), transgenes, body weight, gender and mRNA expression level of immune-related genes, avoiding any confounding effects.

In the current study, cecropin (*Cec*)- and cathelicidin (*Cath*)-transgenic embryos were subjected to CRISPR/Cas9-mediated genome editing to increase the integration efficiency and create fish lines possessing two AMGs in the channel catfish. Inheritance of transgenes and the integration rates of dual-AMG in the same family were investigated. Then the hatchability and fry survival of microinjected individuals were determined compared to non-injected embryos. Additionally, we evaluated the bacterial resistance of single- and dual-AMG transgenic fish from a diversity of genotypes. Finally, a univariate LR model was employed to analyze the contribution of endogenous AMGs to the cumulative survival rate (CSR) and their synergistic effect with transgenic AMGs.

2. Materials and methods

2.1 Ethical approval

All experiments were conducted at the Fish Genetics Research Unit, E.W. Shell Fisheries Research Center, Auburn University, AL. The handling and use of animals were certified by the relevant guidelines from professional training programs. The Institutional Animal Care and Use Committee at Auburn University (AU-IACUC) approved the experimental techniques used in the current study. All fish studies were carried out in compliance with the procedures and standards established by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC).

2.2 Preparation of sgRNA, CRISPR/Cas9 system, and donors

As the target integration sites, we choose the luteinizing hormone (*lh*, accession number: NM_001200080.1) and myostatin (*mstn*, accession number: AF396747.1) genes, which regulate gonad development and growth, respectively. Previous works have revealed that *lh*-mutant fish are temporarily sterilized, and *mstn*-deficient individuals grow faster than the wild type (Qin,

2019; Coogan et al., 2022a). Therefore, the *mstn* gene of channel catfish was replaced by the coding sequence (CDS) of the cecropin (*Cec*) gene (accession number: NM_079851.3) (Hoskins et al., 2007). The *lh* gene was replaced by the CDS of the cathelicidin (*Cath*) gene (accession number: XM_006037211.3) (Chen et al., 2017). In this context, two customized single guide RNAs (sgRNA1 and sgRNA2) were designed to target the exon1 of *mstn* and the exon2 of the *lh* locus of the channel catfish genome, respectively, with high efficiency and success reported (Khalil et al., 2017; Xing et al., 2022c). The sgRNAs were selected using the CRISPRscan online tool (https://www.crisprscan.org/), and candidate sgRNA sequences were aligned with the channel catfish genome via the Basic Local Alignment Search Tool to avoid off-target cleavages. Additionally, putative off-target sites were excluded using the online tool Cas-OFFinder (http://www.rgenome.net/cas-offinder/) (Bae et al., 2014). The Maxiscript T7 kit (Thermo Fisher Scientific, Waltham, MA) synthesized sgRNA1 and sgRNA2 *in vitro* according to manufacturer instructions. Then purified sgRNAs were prepared using the RNA Clean and Concentrator Kit (Zymo Research, Irvine, CA). The concentration and quality of sgRNAs were evaluated with Nanodrop 2000 spectrophotometer (Thermo Scientific, USA) and 2% agarose gel with 1 × TBE buffer. The synthetic sgRNAs were diluted to a concentration of ~300 ng/μL and divided into PCR tubes (2 μl/tube), then stored at -80 ℃ until use. The Cas9 protein powder was purchased from PNA BIO Inc. (Newbury Park, CA), and was diluted with DNase/RNase-free water to 50 ng/μL, keeping at − 80 ℃ until use. Two gene-specific oligonucleotides for sgRNAs recognition and universal primer used in this study are listed in Table 8.

To increase the on-target knock-in (KI) efficiency, double-stranded DNA (dsDNA) flanked by a targeting cassette and two sgRNA recognition sequences were cloned into the pUC57_mini vector at the EcoRV enzyme digestion site to create a dsDNA construct as the double-cut plasmid donor (Appendix 6). In detail, the pUC57_Cec donor was constructed using the CDS of the *Cec* gene flanked by two 300-bp homology arms (HA1 and HA2, derived from the exon1 of the *mstn* gene) and sgRNA1-PAM sequences (23 bp each) (Figure 14A). Similarly, the pUC57-_Cath donor was established using the CDS of the *Cath* gene flanked by two 300-bp homology arms (HA3 and HA4, derived from the exon2 of the *lh* gene) and sgRNA2-PAM sequences (23 bp each) (Figure 14B). Expression of both the *Cec* and *Cath* genes was driven by the zebrafish ubiquitin (UBI) promoter (Mosimann et al., 2011). In addition to the UBI promoter and CDS, a

Table 8. Oligonucleotide sequences for single guide RNA (sgRNA) synthesis, transgene detection, and quantification of transgene/immune-related antimicrobial peptides. PCR primers were used to detect cathelicidin (*Cath*) and cecropin (*Cec*) transgenes in putative transgenic channel catfish, *Ictalurus punctatus*. qRT-PCR primers were used to determine the expression of transgenes and innate antimicrobial peptide genes. *Catpd*, cathepsin D; *UBI*, ubiquitin; *LEAP2*, liver-expressed antimicrobial peptide-2; *NK-lysin*, natural killer lysin; *BPI*, bactericidal permeability-increasing protein; *TCP*, thrombin-derived C-terminal peptides; *H2A*, histone H2A; *CCL3*, C-C motif chemokine 3.

Oligo name	Nucleotide sequence (5' → 3')	Product Size (bp)	Purpose
sgRNA synthesis			
sgRNA1	TTCAAACCGCCATCTGCAGC	—	sgRNA1 synthesis
sgRNA2	GCGGACAGGTATCCGGTAAG	—	sgRNA2 synthesis
Universal Primer	TTTTGCACCGACTCGGTGCCACTTTTT CAAGTTGATAACGGACTAGCCTTATTT TAACTTGCTATTTCTAGCTCTAAAAC	—	Scaffold of the sgRNA synthesis
PCR Primers			
Chr1_Cath-F	GCAGCCAATCACTGCTTGTA	591	Determine the Cath-polyA region of the Chr1_Cath transgenic fish
Chr1_Cath-R	GTGGTTTGTCCAAACTCATCAA		
IR_Cec-F1	CCAATAGGGACTTTCCATTGAC	448	Determine the Cec-polyA region of the IR_Cec transgenic fish (1st PCR)
IR_Cec-R1	CCAGTTAAGCAGTGGGTTCTCT		
IR_Cec-F2	CCCACTTGGCAGTACATCAA	305	Determine the Cec-polyA region of the IR_Cec transgenic fish (2nd PCR)
IR_Cec-R2	GGCGGAGTTGTTACGACATT		
Cec-F	GGAGCCGTACTGTTCCGTTA	352	Determine the Cec-polyA region of the Mstn_Cec transgenic fish
Cec-R	CCCATATGTCCTTCCGAGTG		
Prom1-F	GCAGCCAATCACTGCTTGTA	462	Determine the Prom-Cec region of the Mstn_Cec transgenic fish
Prom1-R	ATTCCGAGGACCTGGATTG		
HA1-F	TGGAGAAAGTTGTGGGTCTG	636	Determine the junction of the HA1
HA1-R	CAGGTCACTGATCCCTCCAT		
HA2-F	CCCCTTGAGCATCTGACTTC	664	Determine the junction of the HA2
HA2-R	AAGCAGTAGTAAAGGGACTCACG		
Cath-F	TTCAGGAGCCGTACTGTTCC	597	Determine the Cath-polyA region of the LH_Cath transgenic fish
Cath-R	GCATTCTAGTTGTGGTTTGTCCA		
Prom2-F	ACCCTTTGCCACAGTTCTCC	542	Determine the Prom-Cath region of the LH_Cath transgenic fish
Prom2-R	GGCCCTTGGTTGTAGACG		
HA3-F	TAAGGCCACGTTTCGATTCT	573	Determine the junction of the HA3
HA3-R	TCATTTTGCCGTCTGTTGTT		
HA4-F	TGAGTTTGGACAAACCACAAC	598	Determine the junction of the HA4
HA4-R	TTGATTGAAAATGTTTCCCTGTT		
qRT-PCR Primers			
Cec_qPCR-F	CTTCGTCTTTGTGCACTCA	166	For quantitative real-time PCR (qRT-PCR) of transgenes (*Cec* and *Cath*) and innate AMGs (*Catpd*, *UBI*, *LEAP2*, *NKl-lysin*, *BPI*, *TCP*, *H2A* and *CCL3*)
Cec_qPCR-R	AGCGGTGGCTGCAACATT		
Cath_qPCR-F	GCAGGGGTCTATTCAAGAAGC	125	
Cath_qPCR-R	GTCTGGATCTCACCGCCTTC		
Catpd-F	CCTCTGATCATGGGGGAGTA	132	
Catpd-R	GCCCATCTTTGTGGACTTGT		
UBI-F	CGCACCCTGTCTGACTACAA	106	
UBI-R	TGGGGGTGGTGTAAGACTTC		
LEAP2-F	GTACTCGCCCAACAGGTAGC	120	
LEAP2-R	GATTCTCCAAAGGGGTGTCA		
NKl-F	GGGCCATGAAGAAAGTGAAA	111	
NKl-R	TGCAGAGACCTCGAAGGAAT		
BPI-F	CGACATGATCCCTTCCAGTT	113	
BPI-R	CCTTCACCAGGAGCTTCATC		
TCP-F	AAGGGAAGAGGGCTTCTCAG	121	
TCP-R	GCCTCCTCATGGTCACAGAT		
H2A-F	GACGTGTGCACAGGCTTCTA	115	
H2A-R	GCCAACTCCAGAATCTCAGC		

CCL3-F	TCTCGTTCTCTCCTGCTGGT	117	
CCL3-R	AGGGATTGGATGTGTCTGGA		
18s-F	GAGAAACGG CTACCACATCC	128	Internal control for quantitative real-time PCR
18s-R	GATACGCTCATT CCGATTACAG		

poly-A tail was provided at the right of CDS to enhance translation in both plasmid donors. Donors pUC57_Cec and pUC57_Cath were integrated respectively into the *mstn* and *lh* loci in a double-cut manner (Zhang et al., 2017). The plasmid donors were synthesized by Genewiz LLC (South Plainfield, NJ).

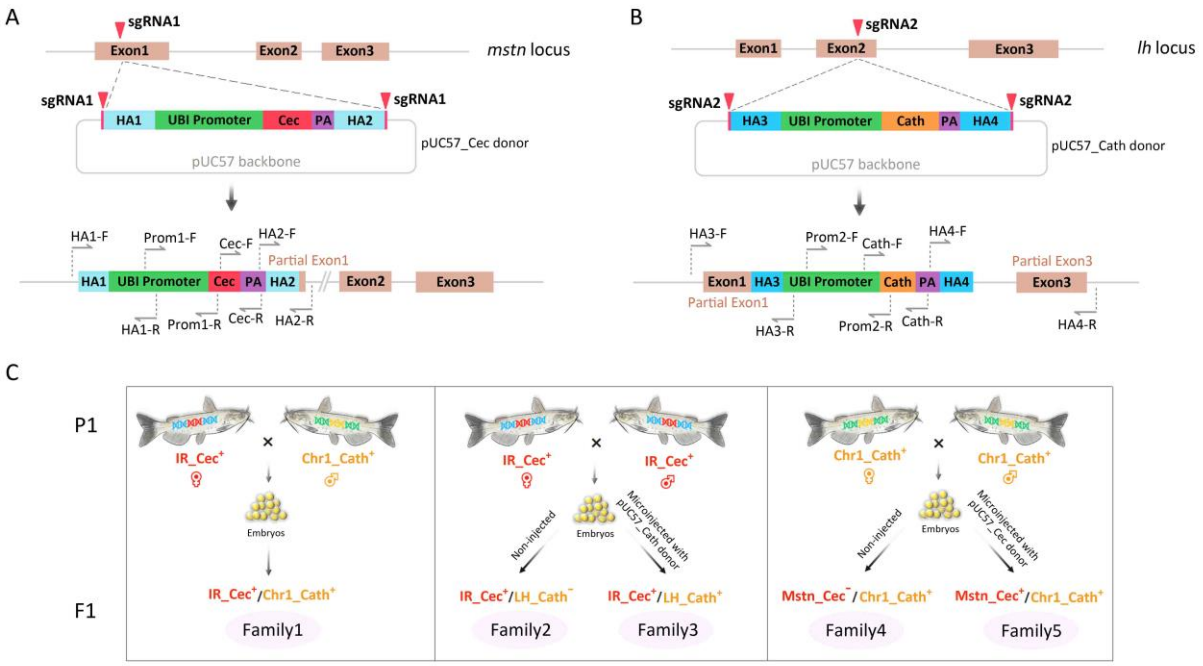


Figure 14. Construction of the plasmid donors and the experimental breeding design for knock-in and knock-out in channel catfish (*Ictalurus punctatus*). (A) Schematic representation of the HA-mediated CRISPR/Cas9 system knocking in the cecropin gene from moth (*Hyalophora cecropia*; *Cec*) at the *mstn* locus. Double-cut plasmid donor (pUC57_Cec) was constructed using the coding sequence (CDS) of the *Cec* gene flanked by two 300-bp homology arms (HA1 and HA2) derived from exon1 of the *mstn* gene and sgRNA1-PAM sequences (23 bp each, highlighted with red lines). **(B)** Schematic illustration of the alligator (*Alligator sinensis*) cathelicidin (*Cath*) gene knocking in via the HA-mediated CRISPR/Cas9 system at the *lh* locus. pUC57_Cath plasmid was constructed using the CDS of the *Cath* gene flanked by two 300-bp homology arms (HA3 and HA4) derived from exon2 of the *lh* gene and sgRNA2-PAM sequences (23 bp each). The structure of the target loci's exons was constructed by brown bars, and sgRNA-targeted sites were indicated by red triangles. Primer sets were illustrated and showed the strategy to detect junctions, the UBI promoter region, and the insert-specific region of transgenes using PCR amplifications. **(C)** Crossbreeding combined with CRISPR/Cas9-aided microinjection to generate progeny harboring two transgenes (*_Cec^{+}/*_Cath^{+}). We did not perform the microinjection for

embryos from IR_Cec^{+}♀ × Chr1_Cath^{+}♂family. In addition to the non-injected control family, the microinjection was conducted in IR_Cec^{+}♀ × IR_Cec^{+}♂ and Chr1_Cath^{+}♀ × Chr1_Cath^{+}♂ using pUC57_Cath and pUC57_Cec plasmids, respectively. IR_Cec^{+}, F_3 cecropin-transgenic channel catfish with the integration of the *Cec* gene in a random manner. Chr1_Cath^{+}, P_1 cathelicidin-transgenic channel catfish produced by integrating the *Cath* gene at a non-coding region of chromosome 1. IR_Cec^{+}/Chr1_Cath^{+}, transgenic fish possessed both *Cec* and *Cath* transgenes (*Cec* was integrated randomly and *Cath* was integrated at the non-coding region of chromosome 1); IR_Cec^{+}/ LH_Cath^{+}, transgenic fish possessed both *Cec* and *Cath* transgenes (*Cec* was integrated randomly and *Cath* was integrated at the *lh* locus); Mstn_Cec^{+}/ Chr1_Cath^{+}, transgenic fish possessed both *Cec* and *Cath* transgenes (*Cec* was integrated at the *mstn* locus and *Cath* was integrated at the non-coding region of chromosome 1). IR, integrated randomly; Chr1, chromosome1; UBI, zebrafish ubiquitin; PA, poly-A tail.

2.3 Brood stock selection and experimental design

Healthy and mature F_3 cecropin- and P_1 cathelicidin-integrated channel catfish parents were utilized. Specifically, the F_3 cecropin-transgenic (IR_Cec) fish line was established by Dunham et al. (2002) with the integration of the moth (*Hyalophora cecropia*) cecropin gene in a random manner. P_1 cathelicidin-transgenic (Chr1_Cath) founders were produced by knocking in the alligator (*Alligator mississippiensis*) cathelicidin gene at a non-coding region in chromosome 1 (Simora et al., 2020). The mating design follows here and is in Figure 14C. Here, three positive pairs (IR_Cec^{+}♀, ♂ and Chr1_Cath^{+}♀, ♂) from each fish line were identified by running the nested PCR and normal PCR to detect the correspondingly inserted DNA fragments. Primers IR_Cec-F_1/R_1 (and IR_Cec-F_2/R_2), and Chr1_Cath-F/R were used for the determination of these two genotypes (Table 8).

To generate the channel catfish line possessing both *Cec* and *Cath* transgenes (*_Cec^{+}/*_Cath^{+}), three pairings (IR_Cec^{+}♀× IR_Cec^{+}♂, IR_Cec^{+}♀× Chr1_Cath^{+}♂ and Chr1_Cath^{+}♀× Chr1_Cath^{+}♂) combined with a CRISPR/Cas9-mediated microinjection were conducted (Figure 14C). The embryos from IR_Cec^{+}♀× Chr1_Cath^{+}♂ were hatched without a second microinjection to produce *_Cec^{+}/*_Cath^{+} individuals (family 1). The embryos from IR_Cec^{+}♀× IR_Cec^{+}♂ were divided into two parts, one-third of them was hatched without microinjection (family 2), while the remaining two-thirds were subjected to the CRISPR/Cas9-based microinjection using the pUC57_Cath plasmid as a donor template (family 3). Similarly, the embryos from Chr1_Cath^{+}♀× Chr1_Cath^{+}♂ were separated into non-injected (family 4) and a group microinjected with the pUC57_Cec donor (family 5). Based on this trail, our objectives

were: 1) determining the inheritance of transgenes *Cec* and *Cath*; 2) comparing the efficiency of dual-AMG integration; and 3) assessing resistance against bacteria in progeny with single- and dual-AMG integrated fish.

2.4 Fertilization, transgenic fish production and rearing

Three mature IR_Cec^{+} and Chr1_Cath^{+} transgenic females each were implanted with 75 μg/kg of luteinizing hormone-releasing hormone analog (LHRHa) and were placed in spawning bags in a flow-through tank (800 L of water) at 26 − 28 ℃ with aeration (> 5 ppm dissolved oxygen) for ovulation (Khalil et al., 2017). Spawning bags were checked for eggs every 4 hours at 36 hours post hormone injection. Females were gently stripped if several eggs were observed in the spawning bag, and the embryos were divided into two groups, as mentioned above. After a female of either genotype spawned, a male of that same genotype and the other genotype were euthanized and sacrificed to prepare sperm in 0.9% saline solution (g: v = 1: 10) for fertilization. Two milliliters of sperm solution were added to ~300 eggs and mixed gently in a 20-cm greased pan. After one-minute quiescence, sufficient pond water was supplemented to activate the sperm, then the sperm/egg mixture was swirled for 30 s. More water was added and kept the embryos in a single layer in the pan, then the embryos were allowed to harden for 15 min before microinjection.

The dsDNA-mediated CRISPR/Cas9 system for microinjection consists of Cas9 protein, sgRNA, plasmid donor, and phenol-red indicator in a 2:1:1:0.5 ratio. Embryos from family 3 and family 5 were subjected to a second microinjection to produce the *_Cec^{+}/*_Cath^{+} fish line. Specifically, 4 μL of Cas9 protein (50 ng/μL), 2 μL of sgRNA1 (~300 ng/μL), 2 μL of pUC57_Cath plasmid (50 ng/μL), and 1 μL of phenol-red solution were mixed and microinjected for family 3. With respect to family5, the same components were adopted for microinjection, but the sgRNA and donor template were replaced with sgRNA$_2$ and pUC57_Cec plasmid, respectively. For each mixture of the CRISPR/Cas9 system, we first mixed Cas9 protein and sgRNA on ice for 10 mins, then the donor plasmids were added. To reduce potential mosaicism and off-target effects, the mixed solution was microinjected into one-cell stage embryos as previously described (Khalil et al., 2017). Every 6 μL of the mixture was loaded into a 1.0-mm OD borosilicate glass capillary that was pulled into a needle by a vertical needle puller (David Kopf Instruments, Tujunga, CA) and injected into 600 embryos during the first 90 min post-fertilization.

To calculate the hatchability, fry survival rate, and inheritance/integration rate, the microinjected (family 3 and 5) and non-injected embryos (family 1, 2, and 4) from each family were separated into five 10-L tubs refilled with 7-L Holtfreter's solution (59 mmol NaCl, 2.4 mmol $NaHCO_3$, 1.67 mmol $MgSO_4$, 0.76 mmol $CaCl_2$, 0.67 mmol KCl) (Armstrong and Malacinski, 1989) with continuous oxygen (> 5 ppm). A total of 5 families × 5 replicates = 25 tubs (200 eggs/tub) were separately and randomly placed into two flow-through hatching troughs, and the rest of the embryos were incubated in another flow-through trough with paddles for oxygenation (> 5 ppm). Drop-in heaters were utilized upstream near the water inlet to keep the water temperature at 26 − 28 ℃ in each trough. Holtfreter's solution was completely replaced twice daily, and dead eggs/fry were recorded and removed.

All sac fry were transferred into new tubs refilled with pond water at six days post-fertilization and fed alive artemia four times per day. After a two-week culture in tubs, fry from each family were placed into separate 60-L tanks in a recirculating system for growth out (120 fry/tank). During this phase, the fry (< 1 g) received Purina® AquaMax® powdered feed (50% crude protein, 17% crude fat, 3% crude fiber, and 12% ash) four times per day for two months. Then fingerlings (5 − 10 g) were fed with Aquaxcel WW Fish Starter 4512 (45% crude protein, 12% crude fat, 3% crude fiber, and 1% phosphorus) twice a day for two months to apparent satiation. Juvenile fish (20 − 50 g) were fed with WW 4010 Transition feed (40% crude protein, 10% crude fat, 4% crude fiber, and 1% phosphorus) once a day (Coogan et al., 2022a).

2.5 Sample collection and integration analysis

After a 4-month culture, all fingerlings (20 − 50 g) were pit-tagged (Biomark Inc., Boise, Idaho, USA) and weighed individually. Mixed samples of fin and barbel clip were collected from each family and stored at − 20℃ until use. DNA extraction was performed using proteinase K followed by protein precipitation as described by Kurita et al. (2004). Different genotyping strategies were employed for different families: Primers IR_Cec-F_1/R_1, IR_Cec-F_2/R_2, and Chr1_Cath-F/R were used for the detection of IR_Cec and Chr1_Cath regions in family 1, 2 and 4; Given family 3, primers IR_Cec-F_1/R_1 and IR_Cec-F_2/R_2 were used for amplifying the IR_Cec

region, and Cath-F/R, Prom2-F/R, HA3-F/R, and HA4-F/R were used to detect the insertion of the *Cath* at the *lh* locus. With respect to family 5, primers Chr1_Cath-F/R were used to determine the Chr1_Cath region, and Cec-F/R, Prom1-F/R, HA1-F/R and HA2-F/R were used to detect the integration of the *Cec* at the *mstn* locus (Figure 14AB). All primers were designed via the online software Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) and listed in Table 8. PCR was performed in a 10-μL system, and PCR products were electrophoresed on a 1.0% agarose gel with 1 × TAE buffer. A bright band of each region with the corresponding length indicated an on-target positive. PCR products were inserted into the pCR^{TM}4-TOPO vector using TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions and transformed into One Shot TOP10F chemically competent *Escherichia coli* (Invitrogen, Carlsbad, CA) as previously described (Elaswad et al., 2018). Then three colonies were randomly picked to perform Colony PCR, and liquid *E. coli* cultures were prepared for rolling circle amplification (RCA) sequencing by Sequetech (Mountain View, CA). Finally, the sequencing results were blasted with transgenes using MAFFT (version 7, https://mafft.cbrc.jp/alignment/server/) to identify inserted DNA sequences.

2.6 *Edwardsiella ictaluri* challenge

A total of 80 channel catfish (mean ± SD weight = 16.49 ± 3.69 g) from three major groups containing eight genotypes (10 fish/genotype, off-target individuals excluded) were used for the *Edwardsiella ictaluri* challenge experiment: wild-type (WT, from family 1), one-AMG integrated (IR_Cec^{+}/*_Cath^{-} [from family 2], IR_Cec^{-}/LH_Cath^{+} [from family 3], *_Cec^{-}/Chr1_Cath^{+} [from family 4], Mstn_Cec^{+}/Chr1_Cath^{-} [from family 5]) and two-AMG integrated (IR_Cec^{+}/Chr1_Cath^{+} [from family 1], IR_Cec^{+}/LH_Cath^{+} [from family 3], Mstn_Cec^{+}/Chr1_Cath^{+} [from family 5]). All catfish were reared and acclimated for five days (temperature ~26℃, DO > 5.0 ppm) in a 200-L tank with flow-through water from a watershed retention pond. The system was contained within a biosecure challenge room at the E.W. Shell Fisheries Research Center, Auburn University, AL. Cryo-stock of *E. ictaluri* (S97-773) was streaked for isolation onto brain-heart infusion agar (BHIA; Hardy Diagnostics) and incubated for 48 h at 28 °C. A single colony was harvested and inoculated into 20-mL brain–heart infusion

broth (BHIB; Hardy Diagnostics). The culture was incubated (28 °C; 175 rpm) for 18 h. Bacteria were subcultured into 250 mL of BHIB for another 18 h at the same growth conditions. The culture was then adjusted to an OD_{600} of 1.11. Before starting *E. ictaluri* infection, the water was lowered to a total of 30 L, then 50 mL of *E. ictaluri* suspension containing 4.0 $\times 10^8$ CFU/mL cells was added to the tank resulting in a final dose of 6.76×10^5 CFU/mL. Fish were immersed statically for one hour with aeration provided from air stones, and then water was restored to 1/2 volume of the tank (DO > 6 ppm, and water temperature was 26 − 28°C). In addition to infected groups, one mock-challenged tank containing 30 fish received only sterile BHIB in lieu of the bacteria.

During the first 72 hours of the experiment, tanks were observed every 4 hours for mortality and then three times daily. Challenged fish were continuously monitored for a total of 10 days for clinical signs of *E. ictaluri* by isolating bacteria on BHIA plates from the liver to discern the cause of death from the pathogen. The moribund fish were considered as dead and necropsied, and the fresh head kidney was transferred directly into liquid nitrogen (within cryovials). After a 10-day observation, all surviving individuals were euthanized with 200 ppm tricaine methanesulfonate (MS222, Ferndale, WA), and head kidneys were quickly obtained for storage in liquid nitrogen. All samples from moribund and sacrificed individuals were then stored at –80 ℃ for RNA isolation.

2.7 RNA isolation and gene expression

Total RNAs were isolated from the kidney using TRIzol reagent (Thermo Fisher Scientific) and were reverse transcribed to cDNA using iScript™ Synthesis Kit (Bio-Rad, Hercules, CA) according to the manufacturer's instructions. To determine the contribution of transgenes to fish survival and the co-expression with endogenous antimicrobial peptide genes (AMGs) after the *E. ictaluri* challenge, the mRNA level of channel catfish AMGs was evaluated, including cathepsin D (*Catpd*, GU588646), ubiquitin (*UBI*, NM_001200293), liver-expressed antimicrobial peptide-2 (*LEAP2*, AY845141), natural killer lysin (*NK-lysin*, AY934592), bactericidal permeability-increasing protein (*BPI*, AY816351), thrombin-derived C-terminal peptides (*TCP*, XM_017458593), histone H2A (*H2A*, XM_017462467) and C-C motif chemokine3 (*CCL3*, XM_017471106), and two transgenes (*Cath* and *Cec*) in the kidney of both dead and alive fish

since those AMGs exhibited upregulation after the pathogen infection in our previous study (Wang et al., 2022).

Real-time quantitative reverse transcription PCR (qRT-PCR) was performed on a C1000 Thermal Cycler using SsoFast™ EvaGreen Supermix kit (Bio-Rad, Hercules, CA). Concentrations of the cDNA products were diluted to 220 ng/μL, and 1 μL was used in a 10 μL qRT-PCR reaction volume. The mRNA level of 18S rRNA was used as an internal control, and the detailed qRT-PCR procedure was set up according to Coogan et al. (2022a). All raw data were expressed relative to the expression levels of 18S rRNA in each sample using the CFX Manager Software version 1.6 (Bio-Rad, Hercules, CA), and crossing-point (CT) values were converted to fold differences using the relative quantification method (Livak and Schmittgen, 2001). The primers used for qRT-PCR are listed in Table 8.

2.8 Logistic regression model construction

In our study, two outcomes (die and live) appeared from the *E. ictaluri* challenge experiment, and we defined a binary response to fit a univariate multiple LR model:

$$Y = \begin{cases} 1, \ if\ a\ fish\ survived\ after\ the\ E.ictaluri\ challenge \\ 0, \ if\ a\ fish\ died\ after\ the\ E.ictaluri\ challenge \end{cases}$$

Fish gender (Sex), body weight (BW), and the mRNA level of transgenes (*Cath* and *Cec*) and innate AMGs (*Catpd*, *H2A*, *UBI*, *LEAP2*, *NK-lysin*, *BPI*, *TCP* and *CCL3*) could contribute to fish survival. Here, Sex was a binary predictor, and we redefined it as a dummy variable:

$$Sex = \begin{cases} 1, \ if\ the\ fish\ was\ a\ female \\ 0, \ if\ the\ fish\ was\ a\ male \end{cases}$$

Therefore, the LR model was used to determine the effects of these 12 predictors on our binary outcomes:

$$\pi(x) = P(Y = 1) = \frac{exp(\beta_0 + \beta_1 Sex + \beta_2 BW + \beta_3 Cath + \ldots + \beta_{12} CCL3)}{1 + exp(\beta_0 + \beta_1 Sex + \beta_2 BW + \beta_3 Cath + \ldots + \beta_{12} CCL3)}$$

Then, $\text{Logit}[\pi(x)] = \beta_0 + \beta_1 Sex + \beta_2 BW + \beta_3 Cath + \ldots + \beta_{12} CCL3$

Where $\pi(x)$ denotes the probability of a fish's survival from the challenge experiment, *x* indicates different predictors. The parameter β_i refers to the effect of x_i on the log odds that Y = 1 (fish is alive after *E. ictaluri* challenge), controlling the other *x*. Concerning the interpretation,

larger β_i is an indicative of greater value of π_i (probability of survival after *E. ictaluri* challenge). For instance, we can conclude that the *Cath* transgene contributes more than the *CCL3* gene in improving survival rate if $\beta_3 > \beta_{12} > 0$.

With respect to the model construction, univariate multiple LR models were fitted using both categorical (Sex) and continuous (BW and the mRNA level of genes) predictors. In this case, a model with more than 10 predictors may suffer from multi-collinearity, making it seem that it is nearly redundant when all the others are in the model. Therefore, a correlation test was first performed among these data from gene expression to determine the potential collinear variables. Then a full logit model containing all identified predictors was constructed, and subsequently, reduced models were built by removing non-significant predictors utilizing the backward elimination method. The likelihood ratio test and akaike information criterion (AIC) were adapted to compare the fitness of full and reduced models (Burnham and Anderson, 2004). Lastly, the goodness-of-fit of the final model was determined by model adequacy assessment using the Hosmer-Lemeshow and sensitivity/specificity test (Hosmer et al., 2013).

2.9 Statistical analysis

The one-way ANOVA followed by Tukey's multiple comparisons test was used to determine the mean differences of the inheritance/integration, hatching rate, and fry survival. The survival curves of different genotypes following the *E. ictaluri* challenge were compared by the Kaplan-Meier plots followed by the Log-rank (Mantel-Cox) test. All statistical analysis was achieved via GraphPad Prism 9.4.1 (GraphPad Software, LLC). Final survival between single- and dual-AMG transgenic, gene expression between moribund and sacrificed fish were analyzed using the unpaired Student's two-sample *t*-test. Statistical significance was set at $P < 0.05$, and all data are presented as the mean ± standard error (SEM).

3. Results

3.1 Inheritance and integration rate

Traditional crossbreeding combined with CRISPR/Cas9-mediated targeted gene insertion produced *Cec*-, *Cath*- and *Cec*/*Cath*- transgenic individuals including one-AMG and two-AMG integrated fish in this study. We implanted positive *Chr1_Cath*- (Chr1_Cath$^+$: #1, #11 and #19) and *IR_Cec*-transgenic (IR_Cec$^+$: #23, #29 and #30) female founders with 75 µg/kg LHRHa for

spawning (Figure 15A), and one Chr1_Cath^{+} (#1) and two IR_Cec^{+} females (#23 and #30) gave eggs, respectively. Then three pairs of crossbreeding combined with genome editing were accomplished: #30 IR_Cec^{+} ♀ × #7 Chr1_Cath^{+} ♂, #30 IR_Cec^{+} ♀ × #24 IR_Cec^{+} ♂, and #1 Chr1_Cath^{+} ♀ × #21 Chr1_Cath^{+} ♂. All the positive individuals were determined by gel electrophoresis and DNA sequencing (Figure 15BC).

A total of five families containing three main groups were produced from the three-pair crossbreeding: non-AMG (*_Cec^{-}/*_Cath^{-}), one-AMG (*_Cec^{-}/*_Cath^{+}, *_Cec^{+}/*_Cath^{-}) and two-AMG integrated (*_Cec^{+}/*_Cath^{+}) (Table 9). Our results suggested that transgenes were multi-generational, and dual-AMG transgenic fish with a high integration rate that can be induced through genome editing using transgenic embryos combined with crossbreeding. Specifically, the F_1 generation of family 1 had four genotypes, including non-AMG (IR_Cec^{-}/Chr1_Cath^{-}), one-AMG (IR_Cec^{+}/Chr1_Cath^{-}, IR_Cec^{-}/Chr1_Cath^{+}) and two-AMG integrated (IR_Cec^{+}/Chr1_Cath^{+}), and the mean inheritance of each genotype was 27.0%, 27.0%, 25.0%, and 21.0%, respectively, without significant difference in statistics (P = 0.582) (Figure S12A in Appendix 2). Non-AMG (IR_Cec^{-}) and one-AMG (IR_Cec^{+}) transgenic progeny with similar inheritance (54.8% and 45.2%, P = 0.130) in the family 2 (Figure S12B in Appendix 2). Similarly, non-AMG (IR_Cec^{-}/LH_Cath^{-}), one-AMG (IR_Cec^{+}/LH_Cath^{-}, IR_Cec^{-}/*_Cath^{+}), and two-AMG (IR_Cec^{+}/*_Cath^{+}) transgenic fish were presented with the inheritance/integration rate of 20.9%, 22.4%, 19.4% and 37.3% in the family 3 (Figure S12C in Appendix 2). The percentage of dual-AMG transgenic individuals was significantly higher than that of single-AMG transgenesis (37.3% *vs.* 22.4%, P < 0.0001; 37.3% *vs.* 19.4%, P = 0.001) in the family 3. Besides, non-AMG (Chr1_Cath^{-}, 56.7%) and one-AMG (Chr1_Cath^{+}, 43.3%) transgenic progeny can be detected with no difference (P = 0.127) in the family 4 (Figure S12D in Appendix 2). In addition, non-AMG (Mstn_Cec^{-}/Chr1_Cath^{-}, 29.4%), one-AMG (*_Cec^{+}/Chr1_Cath^{-}, 17.6%; Mstn_Cec^{-}/Chr1_Cath^{+}, 19.1%) and two-AMG (*_Cec^{+}/Chr1_Cath^{+}, 33.8%) transgenic individuals with different inheritance/integration rate were identified in the family 5 (Figure S12E in Appendix 2), and a significantly increased integration rate in the two-AMG group compared to the one-AMG (33.8% *vs.* 17.6%, P = 0.038; 33.8% *vs.* 19.1%, P = 0.049).

Table 9. The inheritance and integration rate of cecropin (*Cec*) and cathelicidin (*Cath*) transgenes introduced by the CRISPR/Cas9-mediated microinjection in five families of channel catfish, *Ictalurus punctatus*. The embryos from family 3 and family 5 received microinjection of pUC57_Cath and pUC57_Cec plasmids, respectively. Concerning IR_Cec^{+}/*_Cath^{+} and *_Cec^{+}/Chr1_Cath^{+}, off-target positive individuals (*Cath* was inserted at non-*lh* locus; Cec was inserted at non-*mstn* locus) were included. N, the number of fingerlings of each genotype.

Family	Mating	Microinjection	Progeny Genotype	Genotyping		Integration (%)		Inheritance (%)	
				N	%	*Cec*	*Cath*	*Cec*	*Cath*
1	IR_Cec^{+}♀× Chr1_Cath^{+}♂	No	IR_Cec^{+}/Chr1_Cath^{-}	41	27.0	_	_	48.0	46.0
			IR_Cec^{-}/Chr1_Cath^{+}	38	25.0				
			IR_Cec^{+}/Chr1_Cath^{+}	32	21.0				
			IR_Cec^{-}/Chr1_Cath^{-}	41	27.0				
2	IR_Cec^{+}♀× IR_Cec^{+}♂	No	IR_Cec^{+}	63	45.2	_	_	45.2	_
			IR_Cec^{-}	76	54.8				
3	IR_Cec^{+}♀× IR_Cec^{+}♂	pUC57_Cath	IR_Cec^{+}/LH_Cath^{-}	15	22.4	_	56.7	59.7	_
			IR_Cec^{-}/LH_Cath^{+}	13	19.4				
			IR_Cec^{+}/*_Cath^{+}	25	37.3				
			IR_Cec^{-}/LH_Cath^{-}	14	20.9				
4	Chr1_Cath^{+}♀× Chr1_Cath^{+}♂	No	Chr1_Cath^{+}	39	43.3	_	_	_	43.3
			Chr1_Cath^{-}	51	56.7				
5	Chr1_Cath^{+}♀× Chr1_Cath^{+}♂	pUC57_Cec	Mstn_Cec^{+}/Chr1_Cath^{-}	12	17.6	51.4	_	_	52.9
			Mstn_Cec^{-}/Chr1_Cath^{+}	13	19.1				
			*_Cec^{+}/Chr1_Cath^{+}	23	33.8				
			Mstn_Cec^{-}/Chr1_Cath^{-}	20	29.4				

In the current study, the dual-AMG transgenic progeny was generated from a traditional crossbreeding IR_Cec^{+} ♀ × Chr1_Cath^{+} ♂ with an inheritance rate of 21.0% (IR_Cec^{+}/Chr1_Cath^{+}). Still, microinjection using the *IR_Cec*- or *Chr1_Cath*-transgenic embryos can yield a higher inheritance/integration rate as 37.3% (IR_Cec^{+}/*_Cath^{+}) and 33.8% (*_Cec^{+}/Chr1_Cath^{+}) in family 3 and family 5, respectively.

Notably, off-target events were observed in family 3 and family 5 when the transgenic embryos were subjected to microinjection. In detail, 76% [19/25] of on-target (IR_Cec^{+}/LH^{-}_Cath^{+}) and 24% [6/25] of off-target (IR_Cec^{+}/IR_Cath^{+}) integrated individuals from the two-AMG (IR_Cec^{+}/*_Cath^{+}) genotype were observed in the family 3, respectively. A total of 56.52% [13/23] on-target positive (Mstn_Cec^{+}/Chr1_Cath^{+}) fish and 43.48% [10/23] off-target positive (IR_Cec^{+}/Chr1_Cath^{+}) were detected from the two-AMG (*_Cec^{+}/Chr1_Cath^{+}) inserted individuals in family 5.

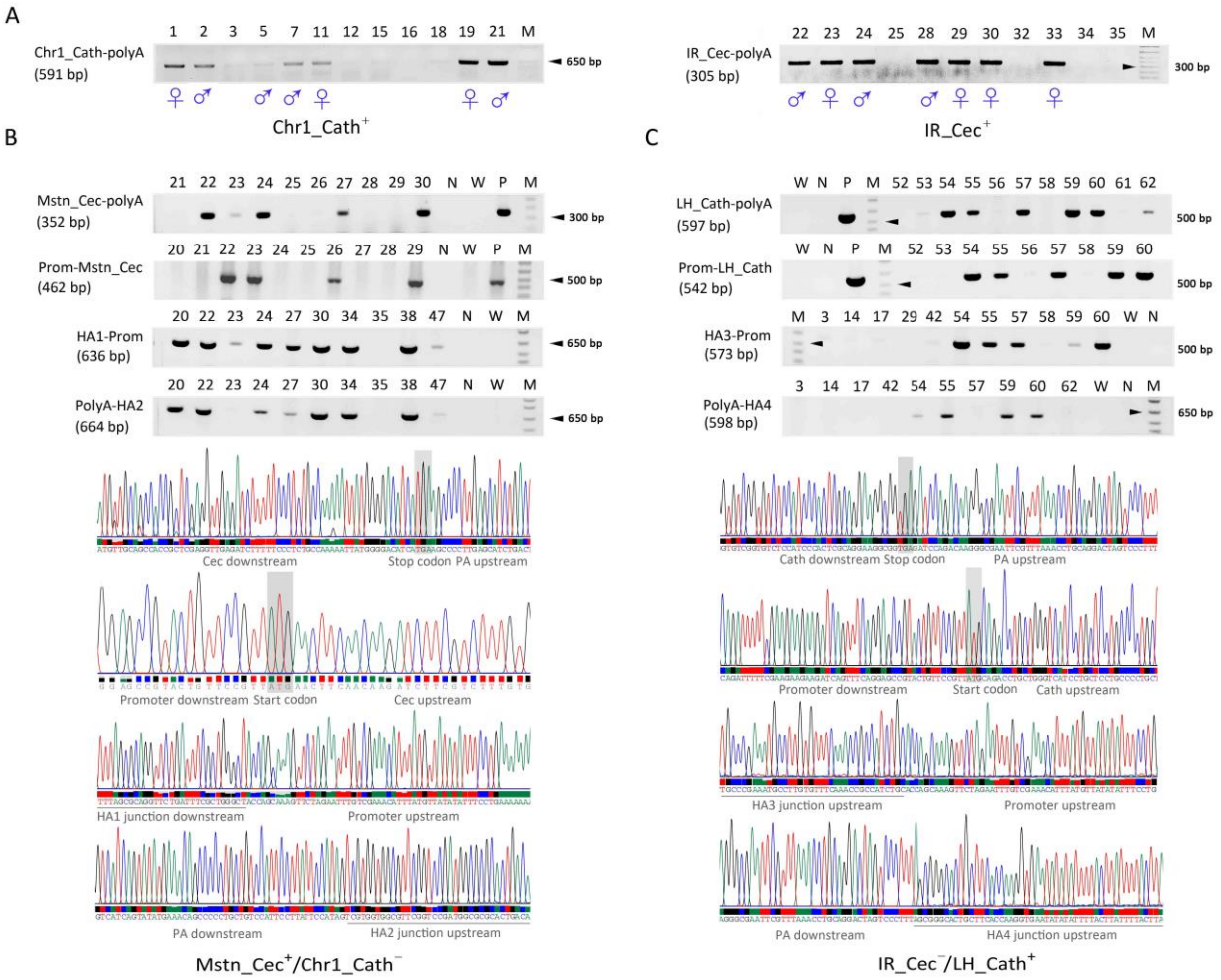


Figure 15. Determination of transgene integration and sequencing results from various genotypes in channel catfish (*Ictalurus punctatus*). (A) 1% TAE gel electrophoresis of PCR amplicons showed the detection of positive cathelicidin- (Chr1_Cath^{+}; Chr1_Cath-polyA, 591 bp) and cecropin-transgenic (IR_Cec^{+}; IR_Cec-polyA, 305 bp) founders. **(B)** 1% TAE agarose gel of PCR amplicons indicated the determination of insertion of the *Cec* transgene at the *mstn* locus from the Mstn_Cec^{+}/Chr1_Cath^{-} genotype. The CDS region (Mstn_Cec-polyA, 352 bp), promoter region (Prom-Mstn_Cec, 462 bp) and junctional regions (HA1-Prom, 636 bp and PolyA-HA2, 664 bp) were verified by sequencing. **(C)** Results from PCR amplicons showed the insertion of the *Cath* transgene at the *lh* locus from the IR_Cec^{-}/LH_Cath^{+} genotype. The CDS region (LH_Cath-polyA, 597 bp), promoter region (Prom-LH_Cath, 542 bp) and junctional regions (HA3-Prom, 573 bp and PolyA-HA4, 598 bp) were verified by sequencing. The numbers on the top of the gel images indicated the sample IDs of the fish. Lane N, negative control using water as a template; Lane W, wild-type control; Lane P, positive control using a plasmid donor as a template; Lane M, DNA marker (1 kb), 300, 500 and 650-bp bands were highlighted with black triangles; CDS, coding sequences; PA, polyA tail; Mstn_Cec^{+}/Chr1_Cath^{-}, *Cec* transgenic fish (the *Cec* transgene was integrated at the *mstn* locus. IR_Cec^{-}/LH_Cath^{+}, *Cath* transgenic fish (the *Cath* transgene was integrated at the *lh* locus); IR_Cec^{+}, F_3 cecropin-transgenic channel catfish with the integration of the *Cec* gene in a random manner. Chr1_Cath^{+}, P_1 cathelicidin-transgenic channel catfish produced by integrating

the *Cath* gene at a non-coding region in chromosome 1. Full gel electrophoresis photos are attached in Figure S11 in Appendix 2.

3.2 Hatching rate and fry survival rate

Families 1, 2, and 4 were produced by artificial fertilization. Families 3 and 5 were produced by artificial fertilization and microinjection. The hatching rate was 44.8, 42.4, 22.5, 44.8, and 22.7%, and the fry survival rate was 38.2, 40.4, 20.4, 33.1, and 14.2% for families, 1, 2, 3, 4 and 5, respectively. Although the CRISPR/Cas9-mediated genome editing increased the efficiency of a dual-AMG transgenesis using transgenic embryos, the microinjection significantly decreased the hatchability compared to the non-injection group (family 3 *vs.* family 2: 22.50% *vs.* 42.40%, $P < 0.0001$; family 5 *vs.* family 4: 22.68% *vs.* 44.80%, $P < 0.0001$; Figure 16A). Furthermore, the microinjection had a significant negative effect on the fry survival rate (family 3 *vs.* family 2: 28.44% *vs.* 39.45%, $P = 0.004$; family 5 *vs.* family 4: 29.24% *vs.* 37.54%, $P = 0.005$; Figure 16B).

Based on the inheritance of *Cec* and *Cath* in family 1, the percentage of dual transgenic individuals inheriting both transgenes should be 0.22, and the observed value was 0.21 (Table 9). Based on the integration and inheritance rates exhibited in family 3, the percentage of double transgenics predicted should be 34%, and it was 37%. Similarly, in family 5, the expected percentage of double transgenics was 27%, and that actually observed was 34%.

3.3 Bacterial resistance

Compared to the WT individuals, AMG-integrated fish exhibited enhanced resistance against *E. ictaluri* during the challenge experiment. The WT fish showed the lowest CSR at 20% after infection. Two types of *Cec*-transgenic fish (40% for IR_Cec^{+}/*_Cath^{-}, 50% for Mstn_Cec^{+}/Chr1_Cath^{-}; $P = 0.117$) had a similar CSR after the bacterial infection, which was significantly higher than the non-AMG inserted fish ($P < 0.0001$). *Cath*-transgenic individuals had higher observed CSR than *Cec*-transgenic channel catfish (50% for *_Cec^{-}/Chr1_Cath^{+}, 60% for IR_Cec^{-}/LH_Cath^{+}); however, there was no significant difference in the CSR among different types of *Cec*- and *Cath*-transgenic fish ($P = 0.800$). In addition, three types of two-AMG integrated fish showed higher CSRs (60% for IR_Cec^{+}/Chr1_Cath^{+}, 70% for

IR_Cec^{+}/LH_Cath^{+}, and 80% for Mstn_Cec^{+}/Chr1_Cath^{+}) than the *Cec*- or *Cath*-inserted fish, and these three types of two-AMG integrated groups had no difference in survival among themselves (P = 0.478) (Figure 16C).

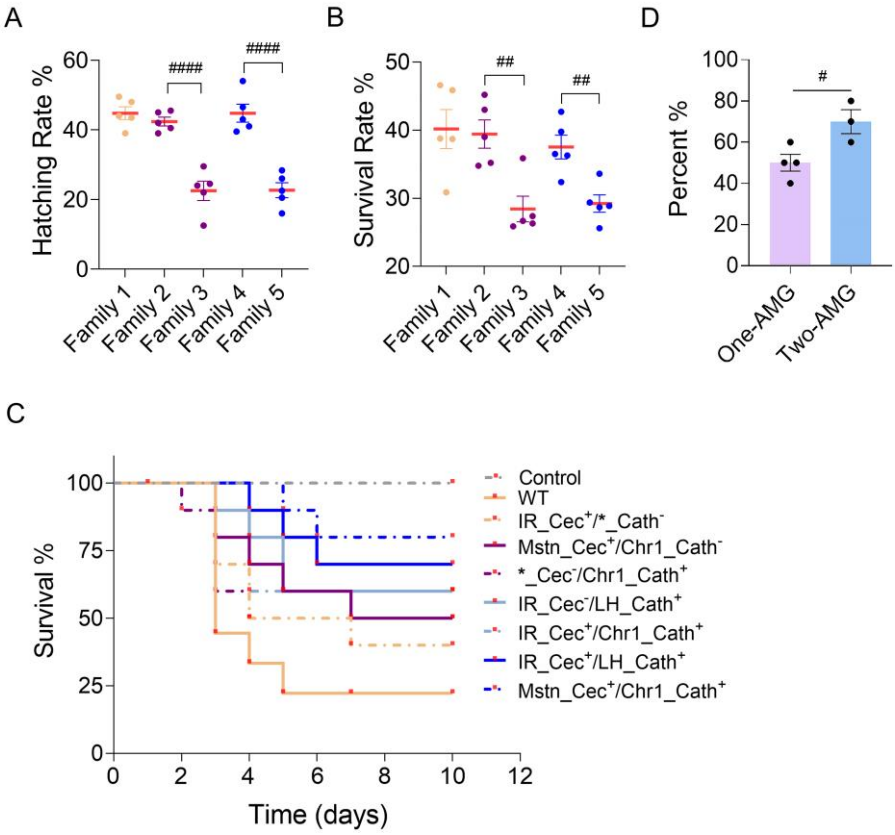


Fig 16. Hatch and fry survival of transgenic channel catfish (*Ictalurus punctatus*) and the resulting cumulative survival rate from different genotypes following challenge with *Edwardsiella ictaluri*. (A) The effect of microinjection on the hatchability. **(B)** The effect of the microinjection on the fry survival rate. **(C)** Kaplan-Meier plots of AMG-integrated channel catfish. A total of 8 genotypes were involved in the bacterial challenge. In addition to these bacterial infection groups, one control group with medium immersion was implanted for the challenge experiment. The comparison of different survival curves was determined by the Log-rank (Mantel-Cox) test. **(B)** Comparison of cumulative survival rate in one-AMG and two-AMG integrated individuals post *E. ictaluri* infection. One-AMG group had 4 genotypes: IR_Cec^{+}/*_Cath^{-}, Mstn_Cec^{+}/Chr1_Cath^{-}, *_Cec^{-}/Chr1_Cath^{+} and IR_Cec^{-}/LH_Cath^{+}; Two-AMG group contained 3 genotypes: IR_Cec^{+}/Chr1_Cath^{+}, IR_Cec^{+}/LH_Cath^{+} and Mstn_Cec^{+}/Chr1_Cath^{+}. WT, wild-type fish; IR_Cec^{+}/*_Cath^{-}, *Cec* transgenic fish (*Cec* was integrated in a random manner); Mstn_Cec^{+}/Chr1_Cath^{-}, *Cec* transgenic fish (*Cec* was integrated at the *mstn* locus); *_Cec^{-}/Chr1_Cath^{+}, *Cath* transgenic fish (*Cath* was integrated at the non-coding region of the chromosome 1); IR_Cec^{-}/LH_Cath^{+}, *Cath* transgenic fish (*Cath* was integrated at the *lh* locus); IR_Cec^{+}/Chr1_Cath^{+}, transgenic fish possessed both *Cec* and *Cath* transgenes (*Cec* was integrated randomly and *Cath* was integrated at the non-coding region of chromosome 1); IR_Cec^{+}/LH_Cath^{+}, transgenic fish possessed both *Cec* and *Cath*

transgenes (*Cec* was integrated randomly and *Cath* was integrated at the *lh* locus); Mstn_Cec^{+}/Chr1_Cath^{+}, transgenic fish possessed both *Cec* and *Cath* transgenes (*Cec* was integrated at the *mstn* locus and *Cath* was integrated at the non-coding region of chromosome 1). AMGs, antimicrobial peptide genes; one-AMG, one AMG (*Cec* or *Cath*) integrated in the genome of channel catfish; two-AMG, two AMGs (*Cec* and *Cath*) integrated in the genome of channel catfish. # = $P < 0.05$; ## = $P < 0.01$; ### = $P < 0.001$; #### = $P < 0.0001$; ns = not significant, by unpaired student's *t*-test or one-way ANOVA.

Overall, one-AMG (50% *vs.* 20%, $P = 0.023$) and two-AMG (70% *vs.* 20%, $P = 0.005$) transgenic channel catfish had significantly enhanced resistance against *E. ictaluri* compared to the WT individuals. The bacterial disease resistance of double AMG integration was greater when compared to that of single-AMG integration (mean CSR: 70% *vs.* 50%, r = 0.79; $P = 0.032$; Pearson Correlation analysis) (Figure 16D).

3.4 Synergistic expression of transgenes and innate AMGs

The mRNA level of 10 AMGs from 38 moribund and 42 sacrificed fish was determined individually (Figure S13-14 in Appendix 2), and the mRNA levels of *Cec*, *Cath*, *CCL3*, *H2A*, *LEAP2*, *UBI*, *BPI*, *TCP* and *NK-lysin* increased (all levels > 0) in the challenged group compared to the non-infected group. In addition, significantly elevated mRNA levels of the *Cath* ($P < 0.0001$) and *TCP* genes ($P < 0.0001$) were detected in the sacrificed individuals compared to the moribund fish. Compared with the moribund fish, the expression of the *Cec* transgene also showed up-regulated but not significant ($P = 0.1732$). In contrast, the expression of the *LEAP2* ($P = 0.0022$) and *NK-lysin* genes was significantly downregulated ($P < 0.0001$) in the moribund fish compared to sacrificed individuals (Figure 17A).

In our study, both categorical (Sex) and continuous (BW and the mRNA level of genes) variables were used as predictors to construct a univariate multiple LR model (Appendix 3). Because we employed more than 10 predictors in the LR model, which may cause multi-collinearity; we first performed a correlation test among the gene expression matrix to determine and remove the potential collinear variables. The results of the correlation analysis showed correlation coefficients of -0.33 (*Cath* and *NK-lysin*) to 0.42 (*LEAP2* and *Catpd*) between different genes, suggesting a weak correlation (Figure 17B). Therefore, we fitted a full model combining Sex, BW, *Cath*, *Cec*, *H2A*, *UBI*, *TCP Catpd*, *LEAP2*, *BPI*, *NK-lysin*, and *CCL3* as predictors. The full

model had an AIC value of 69.304 ($P < 0.0001$), and the Sex ($P = 0.678$), BW ($P = 0.130$), *UBI* ($P = 0.958$), *H2A* ($P = 0.103$), *Catpd* ($P = 0.181$), *BPI* ($P = 0.294$) and *CCL3* ($P = 0.214$) were not significant (Table S4-5 in Appendix 1). Then we refined a reduced model after removing these non-significant predictors. Notably, we kept the *Cec* for the reduced model since the $P = 0.0874$, and it was the transgene of focus. The reduced model had a smaller AIC value ($AIC_2 = 62.776$, $P < 0.0001$) than that of the original model ($AIC_1 = 69.304$) (Table S6-7 in Appendix 1). Therefore, this fitted model was more robust than the full model. Nonetheless, *LEAP2* was not significant ($P = 0.066$), and we discarded it to fit a second reduced model correspondingly. Finally, a model was built that only contained the *Cec* ($P = 0.034$), *Cath* ($P = 0.047$), *TCP* ($P = 0.007$), and *NK-lysin* ($P = 0.003$) with significant effects on the survival after *E. ictalurid* challenge: $\text{Logit}(\hat{\pi}) = -1.95 + 0.17Cath + 0.23Cec + 1.48TCP - 0.39NK\text{-}lysin$ (Figure 17C; Table S8-9 in Appendix 1).

In addition, we assessed the interactions among these four predictors, and there were no significant interactions among them (all $P > 0.05$, Table S10 in Appendix 1). Additionally, the AIC = 71.035 > 64.585 indicated that the fitness decreased when we took the interactions into the model (Table S11 in Appendix 1). Therefore, the final model was fitted without interactions. The Goodness-of-Fit test revealed that our final LR model fitted the data well ($P = 0.206 > 0.05$) (Table S12 in Appendix 1). Additionally, the ROC curve showed 92.04% confidence that high mRNA levels of *Cath*, *Cec*, *TCP*, and low mRNA levels of *NK-lysin* significantly improved the CSR post-*E. ictaluri* infection (Figure S15 in Appendix 2). Our final model displayed that the *Cath* and *Cec* transgenes had a similar contribution to the survival ($\beta_1 = 0.17$ for *Cath*, $\beta_2 = 0.23$ for *Cec*), and the expression of transgenes improved the mRNA level of the *TCP* ($\beta_3 = 1.48$) but decreased that of the *NK-lysin* ($\beta_4 = -0.39$). In particular, the BW and gender of fish did not affect the CSR during the bacterial challenge.

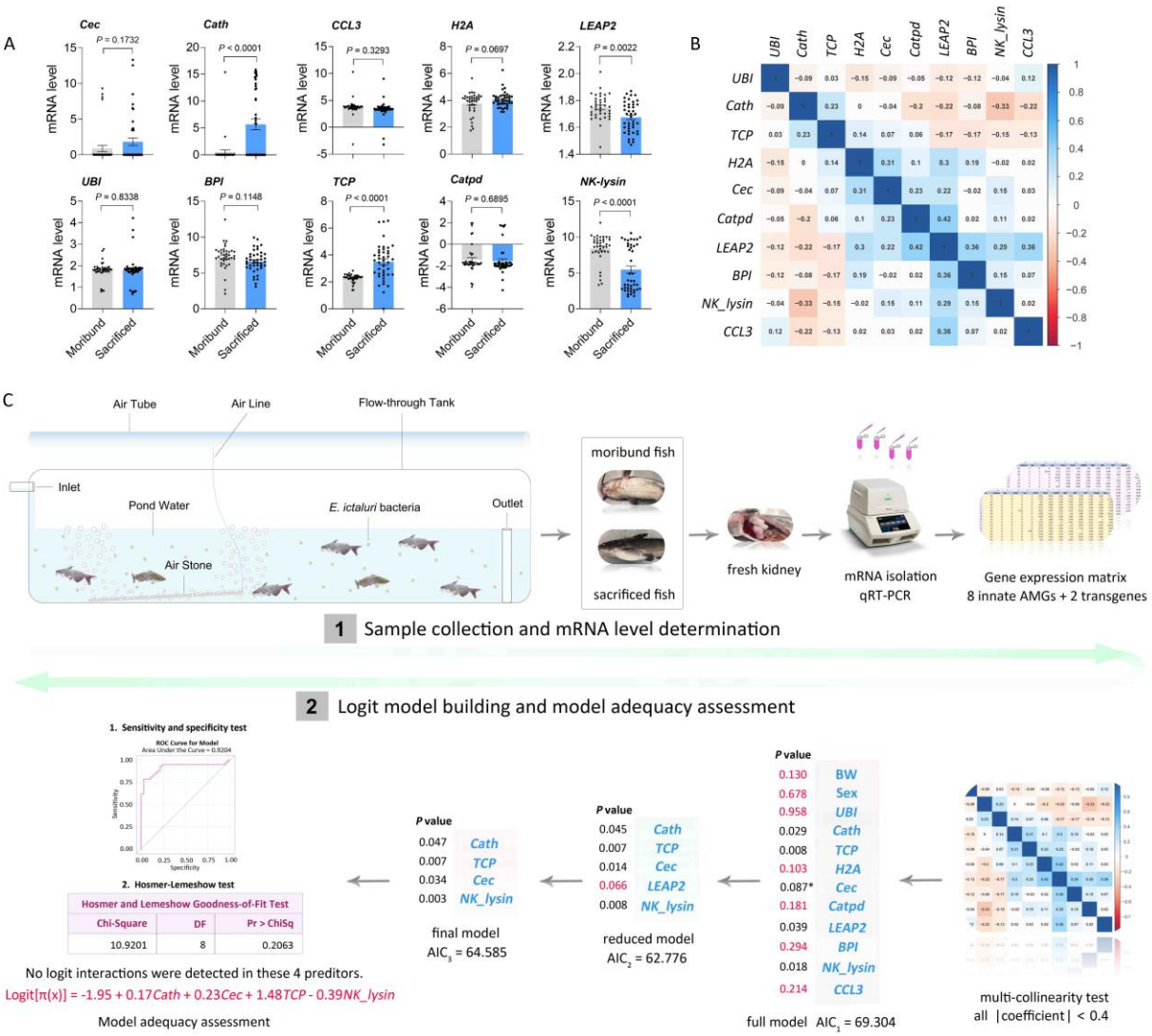


Figure 17. Determination of co-expression of transgenes and innate AMGs in channel catfish (*Ictalurus punctatus*) based on a univariate-multiple logistic regression model. (A) Comparison of the relative mRNA levels of innate AMGs (*Catpd*, *H2A*, *UBI*, *LEAP2*, *NK-lysin*, *BPI*, *TCP* and *CCL3*) and transgenes (*Cec* and *Cath*) in the head kidney from moribund and sacrificed individuals after *Edwardsiella ictaluri* infection. **(B)** Gene expression heatmap to detect potential multi-collinearity among these 10 AMGs. **(C)** A standardized procedure for building a logistic regression model using the gene expression matrix (10 AMGs), body weight (BW) and gender (Sex). AIC, akaike information criterion, judges a model by how close its fitted values tend to be to the true values, and the smaller, the better. AMGs, antimicrobial peptide genes; *Catpd*, cathepsin D (GU588646); *H2A*, histone H2A (XM_017462467); *UBI*, ubiquitin (NM_001200293); *LEAP2*, liver-expressed antimicrobial peptide-2 (AY845141); *NK-lysin*, natural killer lysin (AY934592); *BPI*, bactericidal permeability-increasing protein (AY816351); *TCP*, thrombin-derived C-terminal peptides (XM_017458593); *CCL3*, C-C motif chemokine 3.

4. Discussion

For the first time, we generated dual-gene integrated genetic lines using transgenic channel catfish founders coupled with a CRISPR/Cas9-assisted system for introduction of a second gene. Our results confirmed the reliability and efficiency of this approach, paving the way for novel aquatic genetic enhancement. Additionally, two-AMG-integrated channel catfish were more resistant to pathogenic bacteria than WT or one-AMG-inserted individuals. Furthermore, we applied univariate-multiple LR models to quantify the contribution of various transgenes and endogenous AMGs to fish survival after bacterial challenge, revealing potential synergistic expression of transgenes and endogenous AMGs.

In recent decades, much research has addressed AMG as transgenes to improve disease resistance in fish (Wang and Cheng, 2023). The pooled results of these studies revealed that AMG-transgenic fish were more resistant to disease than WT fish, even though there was high heterogeneity (I^2 = 97.73%) across these published papers (Figure 18A). Our conclusions were in line with those previously described, suggesting that *Cec* and *Cath* transgenes can significantly enhance channel catfish resistance to *E. ictaluri* (Figure 18B) compared to WT fish. Although AMGs as transgenes can significantly inhibit bacterial growth, the meta-analysis found that they did not effectively control some parasites and viruses (e.g., *Ichthyophthirius multifiliis* and grass carp hemorrhagic virus) (Figure 18C). In addition, we deduced from a collection of publications that the *Cath*-transgenic fish showed stronger resistance against pathogens than that in the *Cec*-transgenic lines (Figure 18D). Contrarily, our findings uncovered no significant difference in improving the CSR between *Cec*- and *Cath*-integrated channel catfish, implying that integration of different loci of the same transgene may cause varied genetic pleiotropy on disease resistance enhancement. Consequently, we applied an AMG-moderator meta-analysis of recruited studies to verify that the *Cec* integrated at the *mstn* locus conferred higher disease resistance than that in a random locus, which may imply that *mstn* knockout might have other biological functions, such as immunological response, or potentially non-negligible interaction between the *Cec* transgene and *mstn* locus. Besides, the *Cath* integrated at the non-coding area of chromosome 1 had a higher resistance to bacteria than that at the *lh* locus (Figure 18E). In addition to the locus-change effects, the authors (Figure 18F), fish species (Figure 18G), and pathogen types (Figure

18H) were also in response to the effect variation of the *Cec* and *Cath*. For instance, the metadata illustrated that *Cec* transgene showed significantly increased CSR in rainbow trout/channel catfish against *Aeromonas*, *Flavobacterium*, and *Edwardsiella* from three previous investigations (Dunham et al., 2002; Sarmasik et al., 2002; Chiou et al., 2013). However, the *Cec* did not enhance resistance against the Ich parasite in statistics when we combined all relevant publications to conduct a global assessment. Overall, although the *Cec*- or *Cath*-integrated fish exhibited enhanced resistance against a diversity of pathogens following previous works, the effects of these two transgenes varied depending on publications, fish species, pathogens, and insertion loci.

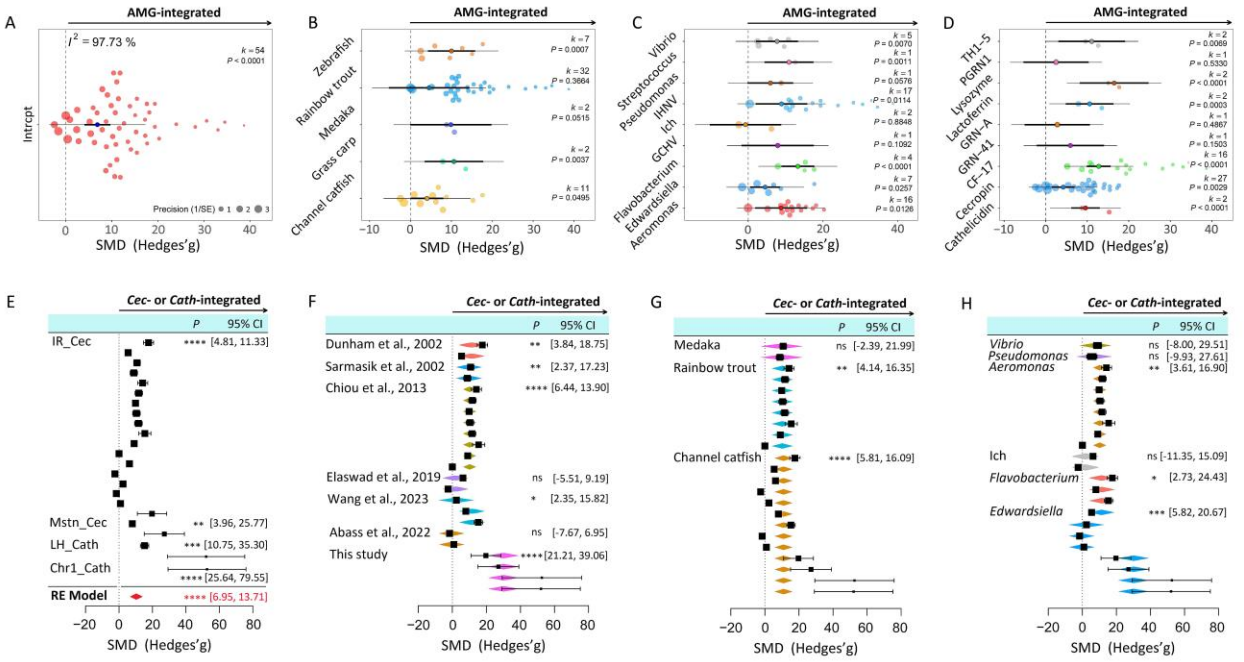


Figure 18. Higher cumulative survival rates (CSRs) post-infection in AMG-integrated fish compared to wild-type individuals based on the metadata. (A) The overall effect of AMGs as transgenes on fish CSRs combined with all AMG transgene-related studies. Although high heterogeneity (I^2 = 97.73%) existed across different studies, AMG-transgenic fish showed higher CSRs than the wild-type fish (effect size = 6.87, k = 54, $P < 0.0001$). **(B)** The effects of AMG transgenes on CSRs varied with the fish species. AMGs as transgenes significantly improved the CSRs in zebrafish ($P = 0.0114$), grass carp ($P = 0.0037$), and channel catfish ($P = 0.00495$). **(C)** Different effects were observed in a diversity of pathogenic challenge experiments when AMGs were integrated into the fish genome. AMGs were significantly effective against *Vibrio* ($P = 0.007$), *Streptococcus* ($P = 0.0011$), IHNV ($P = 0.0007$), *Flavobacterium* ($P < 0.0001$), *Edwardsiella* ($P = 0.0257$) and *Aeromonas* ($P = 0.0126$). **(D)** The type of AMG affected the improvement of CSRs, and TH1-5 ($P = 0.0069$), Lysozyme ($P < 0.0001$), Lactoferrin ($P = 0.0003$), CF-17 ($P < 0.0001$), Cecropin (*Cec*, $P = 0.0029$) and Cathelicidin (*Cath*, $P < 0.0001$) as transgenes significantly increased CSRs of fish after pathogen invasion. **(E)** Differences in improving

CSRs by *Cec* or *Cath* integrated at different loci. Fish integrating the *Cath* transgene at the non-coding region of chromosome1 (Chr1_Cath, $P = 0.0001$) demonstrated the robust resistance to disease, followed by integration of the *Cath* at the *lh* locus (LH_Cath, $P = 0.0002$), the *Cec* at the *mstn* locus (Mstn_Cec, $P = 0.0076$) and the *Cec* integrated randomly (IR_Cec, $P < 0.0001$). **(F)** The *Cec*- or *Cath*-transgenic fish displayed differences in the effect of CSRs in various studies. A significant increment of CSRs was observed in Dunham et al. 2002 ($P = 0.0030$), Sarmasik et al. 2002 ($P = 0.0098$), Chiou et al. 2013 ($P < 0.0001$), Wang et al. 2023 ($P = 0.0023$) and this study ($P < 0.0001$), respectively. **(G)** The effect of *Cec* or *Cath* as transgenes varied in different fish species, and significant enhancement in CSRs was displayed in rainbow trout ($P = 0.0010$) and channel catfish ($P < 0.0001$). **(H)** The *Cec*- or *Cath*-integrated fish showed different effects against various pathogens, and significant enhancement in CSRs was revealed in *Aeromonas* ($P = 0.0025$), *Flavobacterium* ($P = 0.0142$), and *Edwardsiella* ($P = 0.0005$). Ich, *Ichthyophthirius multifiliis*; GCHV, grass carp hemorrhage virus; IHNV, infectious hematopoietic necrosis virus; TH1-5, tilapia hepcidin 1-5; PGRN1, a type of progranulin gene from Mozambique tilapia; GRN-41/GRN-A, AMGs from Mozambique tilapia to produce secreted GRN peptides; CF-17, a synthetic cecropin B analog; SMD, standardized mean difference (Hedges'*g* was used to calculate SMD); k, the number of effect sizes for each category of different moderators; I^2, the heterogeneity index across studies; 95% CI, 95% confidence interval. AMG (*Cec* or *Cath*)-integrated, transgenic fish possess an exogenous AMG (*Cec* or *Cath*) integrated into the genome via genetic engineering. * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$; **** = $P < 0.0001$; ns = not significant.

As mentioned above, the *Cec*-, *Cath*-, and other AMG-transgenic fish were more resistant to pathogens compared to WT individuals. Nevertheless, no studies have documented the integration of two AMGs in aquatic animals. We successfully created channel catfish genetic lines harboring both *Cec* and *Cath* transgenes, and *_Cec^+/*_$Cath^+$ fish lines displayed heightened resistance against pathogenic bacteria compared to the *Cec*- or *Cath*-transgenic fish. These new genetic lines with dual-AMG integration should be evaluated as a strategy to control parasitic or viral diseases in the future.

Importantly, our work showed that transgene-based crossbreeding combined with CRISPR/Cas9-mediated microinjection significantly improved the efficiency of generating dual transgenics. In the current study, we can directly generate *_Cec^+/*_$Cath^+$ offspring by crossbreeding *_Cec^+ and *_$Cath^+$ founders, and the effectiveness of the double AMG integration depends on the heterozygosity of the parents. However, in a scenario, if one transgenic type of the founders is available, the present study offered the feasibility that the microinjection of transgenic fertilized eggs can produce dual-AMG transgenic fish/animals. Although one-step CRISPR/Cas9-assisted microinjection can result in dual- and multiple-insertions, generating transgenic animals harboring multiple transgenes have a low efficiency regarding the increased off-target events and

decreased integration rates as more knock-in genes are involved (Yang et al., 2013; Ota et al., 2014). Interestingly, other studies may have lent the possibility of direct editing genome to produce dual- or multiple-transgenic fish via delicate but complicated co-expression of multiple transgenes, such as introducing multiple plasmids either by targeting long repeated sequences (Xing et al., 2022c) or via polycistronic vector using 2A peptides (Liu et al., 2017), which requires further documentation. As a result, we believe that genome editing via transgenic embryos combined with crossbreeding to integrate multiple genes and reduce off-target effects is promising in boosting disease resistance and other consumer-valued traits in aquaculture.

Unlike previous studies that only enumerated gene expression matrices, we took advantage of the LR analysis to quantify the contribution of the body weight, gender, transgenes and innate AMGs to the CSR after bacterial infection. And our preliminary findings demonstrate that the LR model is capable of handling bivariate analysis based on gene expression matrices and categorial variables in such challenging experiments. According to our final LR model, fish body weight and gender had no impact on CSR during bacterial infection, which is consistent with a recent study that proves there is no correction needed between the body weight and the CSR (Abass et al., 2022). Concerning the AMG expression, although various expressions of eight endogenous AMGs were detected in moribund and sacrificed fish after the *E. ictaluri* challenge, only the *TCP* and *NK-lysin* significantly contributed to the fish survival based on the LR analysis. Earlier studies found that bacterial infection elevated the mRNA level of the *NK-lysin* gene in non-transgenic channel catfish (Wang et al., 2006; Pridgeon et al., 2012). However, in this study, the upregulation of the *Cec*, *Cath*, and *TCP*, as well as the downregulation of the *NK-lysin*, resulted in a high CSR. This may be due to the significant upregulation of the *Cec* and *Cath* transgenes reduced contribution of endogenous AMGs in the disease challenge experiments. However, the detailed effects of antimicrobial peptides as transgenes on the endogenous AMPs' immune modulation, and their potential interaction on regulating and conveying the more disease resitance in catfish require further investigation.

The current study only focused on the innate AMGs but did not include other immune-related genes. Eight endogenous AMGs were chosen based on genome-wide alignment using the channel catfish protein sequence and the antimicrobial peptide database (APD3,

https://aps.unmc.edu/AP/), and the significe of their expression was hypothesized by a pool of studies (Wang et al., 2022). Although LR models revealed that an expression matrix based on eight genes can effectively screen out AMGs with significant effects on survival, there is a high liklihood that more genes are implicated in the immune response. In addition to these innate AMGs, toll-like receptors (e.g., TLR2, TLR3, and TLR5) (Bilodeau and Waldbieser, 2005), NOD-like receptors (e.g., NOD1, NOD2, NOD3a, and NOD3b) (Sha et al., 2009; Rajendran et al., 2012), interferons (e.g., IFN-1, IFN-2, and IFN-4) (Long et al., 2006) and interleukins (IL-β1 and IL- β2) (Wang et al., 2006) also play vital roles in the host response to disease in channel catfish. In the future, a better LR model could be generated utilizing as many immune-related genes as possible. An alternative strategy to address this concern would be to apply RNA-Seq technology to the current bacterial challenge and screen out more immune-related genes to fit a more comprehensive LR model.

This study performed a univariate-multiple LR analysis as a binary response appeared in our bacterial infection experiment. In fact, a polytomous logit regression model with multiple responses for variables can still be applied to more complex cases. In teleost fish, the skin, liver, spleen, and kidney are known to be involved in host immune regulation (Uribe et al., 2011; Rauta et al., 2012). Previous studies have shown that immune-related genes are significantly expressed in the liver, spleen, and kidney in channel catfish (Bao et al., 2005; Xu et al., 2005; Peatman et al., 2006; Pridgeon et al., 2012) after pathogen infection, whereas the current study only investigated the expression of AMGs in the head kidney. In view of this, we could build a proportional-odds LR model (Fox and Hong, 2009) that includes the skin, liver, kidney, and spleen in the model, allowing us to determine which tissue is most favored by transgenic synergism with endogenous AMG expression. The ettects of temporal gene expression (regardless of dead or alive individuals) could be examined with a multi-temporal analysis using a baseline-category LR model (Brophy et al., 2011) allowing identification of which time is most favored by transgenic synergistic endogenous AMG expression. Finally, we can apply generalized linear models to assess the pattern of transgene synergistic expression with endogenous AMGs at different times in various tissues. Future studies should also consider the interactions between transgenic AMGs and endogenous AMGs may have implications in such

immune modulation and documenting their potential immunological function as a biological factor in conveying the disease resistance in catfish.

In this study, AMG transgenic fish demonstrated enhanced disease resistance, and dual-AMG integrated fish were more resistant than single-AMG-transgenic channel catfish. In aggregate, crossbreeding combined with genome editing via transgenic embryos is a promising strategy to control disease outbreaks in aquaculture as it can effectively create bi-AMG- or multi-AMG-transgenic fish lines with enhanced pathogen resistance. Logistic regression models are applicable for determining the contribution of transgenes and immune-related genes to survival rates as well as potential synergistic effects, which could be adapted for research and genetic enhancement application for other diseases and aquatic animal species.

5. References

Abass NY, Simora RMC, Wang J, Li S, Xing D, Coogan M, et al. Response of cecropin transgenesis to challenge with *Edwardsiella ictaluri* in channel catfish *Ictalurus punctatus*. Fish *Shellfish Immunol*. 2022;126:311–317. doi:10.1016/j.fsi.2022.05.050

Abdel-Wahab MM, Taha NM, Lebda MA, Elfeky MS, Abdel-Latif HMR. Effects of bovine lactoferrin and chitosan nanoparticles on serum biochemical indices, antioxidative enzymes, transcriptomic responses, and resistance of Nile tilapia against *Aeromonas hydrophila*. *Fish Shellfish Immunol*. 2021;111:160–169. doi:10.1016/j.fsi.2021.01.017

Armstrong JB, Malacinski GM. Developmental biology of the axolotl (New York: Oxford University Press), 1989.

Bae S, Park J, Kim J-S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*. 2014;30:1473–1475. doi:10.1093/bioinformatics/btu048

Bao B, Peatman E, Li P, He C, Liu Z. Catfish hepcidin gene is expressed in a wide range of tissues and exhibits tissue-specific upregulation after bacterial infection. *Dev Comp Immunol*. 2005;29：939–950. doi:10.1016/j.dci.2005.03.006

Bender R, Grouven U. Ordinal logistic regression in medical research. *J R Coll Physicians Lond*. 1997;31:546–551.

Bilodeau AL, Waldbieser GC. Activation of TLR3 and TLR5 in channel catfish exposed to virulent *Edwardsiella ictaluri*. *Dev Comp Immunol*. 2005;29:713–721. doi:10.1016/j.dci.2004.12.002

Blix TB, Dalmo RA, Wargelius A, Myhr AI. Genome editing on finfish: Current status and implications for sustainability. *Rev Aquac*. 2021;13:2344–2363. doi:10.1111/raq.12571

Brophy C, Connolly J, Fagerli IL, Duodu S, Svenning MM. A baseline category logit model for assessing competing strains of *Rhizobium* bacteria. *JABES*. 2011;16:409–421. doi:10.1007/s13253-011-0058-6

Burnham KP, Anderson DR. Multimodel Inference: Understanding AIC and BIC in model selection. *Sociol Method Res*. 2004;33:261–304. doi:10.1177/0049124104268644

Cal L, Suarez-Bregua P, Braasch I, Irion U, Kelsh R, Cerdá-Reverter JM, et al. Loss-of-function mutations in the melanocortin 1 receptor cause disruption of dorso-ventral countershading in teleost fish. *Pigment Cell Melanoma Res*. 2019;32:817–828. doi:10.1111/pcmr.12806

Chen Y, Cai S, Qiao X, Wu M, Guo Z, Wang R, et al. As-CATH1-6, novel cathelicidins with potent antimicrobial and immunomodulatory properties from *Alligator sinensis*, play pivotal roles in host antimicrobial immune responses. *Biochem J*. 2017;474:2861–2885. doi:10.1042/BCJ20170334

Chiou PP, Chen MJ, Lin C-M, Khoo J, Larson J, Holt R, et al. Production of homozygous transgenic rainbow trout with enhanced disease resistance. *Mar Biotechnol*. 2014;16:299–308. doi:10.1007/s10126-013-9550-z

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339:819–823. doi:10.1126/science.1231143

Coogan M, Alston V, Su B, Khalil K, Elaswad A, Khan M, et al. CRISPR/Cas-9 induced knockout of myostatin gene improves growth and disease resistance in channel catfish (*Ictalurus punctatus*). *Aquaculture*. 2022a;557:738290. doi:10.1016/j.aquaculture.2022.738290

Coogan M, Alston V, Su B, Khalil K, Elaswad A, Khan M, et al. Improved growth and high inheritance of melanocortin-4 receptor (*mc4r*) mutation in CRISPR/Cas-9 gene-edited channel catfish, *Ictalurus punctatus*. *Mar Biotechnol*. 2022b;24:843–855. doi:10.1007/s10126-022-10146-8

Dai J, Zheng J, Ou W, Xu W, Ai Q, Zhang W, et al. The effect of dietary cecropin AD on intestinal health, immune response and disease resistance of juvenile turbot (*Scophthalmus maximus* L.). *Fish Shellfish Immunol*. 2020;100:117–125. doi:10.1016/j.fsi.2020.02.052

Dunham RA, Warr GW, Nichols A, Duncan PL, Argue B, Middleton D, et al. Enhanced bacterial disease resistance of transgenic channel catfish *Ictalurus punctatus* possessing cecropin genes. *Mar Biotechnol*. 2022;4:338–344. doi:10.1007/s10126-002-0024-y

Elaswad A, Khalil K, Ye Z, Alsaqufi A, Abdelrahman H, Su B, et al. Effects of cecropin transgenesis and interspecific hybridization on the resistance to *Ichthyophthirius multifiliis* in channel catfish and female channel catfish × male blue catfish hybrids. *N Am J Aqualcult*. 2019;81:242–252. doi:10.1002/naaq.10096

Fox J, Hong J. Effect displays in R for multinomial and proportional-odds logit models: Extensions to the effects package. *J Stat Soft*. 2010;32:1–24. doi:10.18637/jss.v032.i01

Gratacap RL, Wargelius A, Edvardsen RB, Houston RD. Potential of genome editing to improve aquaculture breeding and production. *Trends Genet*. 2019;35:672–684. doi:10.1016/j.tig.2019.06.006

Hancock REW, Haney EF, Gill EE. The immunology of host defence peptides: beyond antimicrobial activity. *Nat Rev Immunol*. 2016;16:321–334. doi:10.1038/nri.2016.29

Hoskins RA, Carlson JW, Kennedy C, Acevedo D, Evans-Holm M, Frise E, et al. Sequence finishing and mapping of *Drosophila melanogaster* heterochromatin. *Science*. 2007;316:1625–1628. doi:10.1126/science.1139816

Hsieh J-C, Pan C-Y, Chen J-Y. Tilapia hepcidin (TH)2-3 as a transgene in transgenic fish enhances resistance to *Vibrio vulnificus* infection and causes variations in immune-related genes after infection by different bacterial species. *Fish Shellfish Immunol*. 2010;29:430–439. doi:10.1016/j.fsi.2010.05.001

Jiang D, Chen J, Fan Z, Tan D, Zhao J, Shi H, et al. CRISPR/Cas9-induced disruption of wt1a and wt1b reveals their different roles in kidney and gonad development in Nile tilapia. *Dev Biol*. 2017;428:63–73. doi:10.1016/j.ydbio.2017.05.017

Khalil K, Elayat M, Khalifa E, Daghash S, Elaswad A, Miller M, et al. Generation of myostatin gene-edited channel catfish (*Ictalurus punctatus*) via zygote injection of CRISPR/Cas9 system. *Sci Rep*. 2017;7:7301. doi:10.1038/s41598-017-07223-7

Kurita K, Burgess SM, Sakai N. Transgenic zebrafish produced by retroviral infection of in vitro-cultured sperm. *Proc Natl Acad Sci* USA. 2004;101:1263–1267. doi:10.1073/pnas.0304265101

Lee J. An insight on the use of multiple logistic regression analysis to estimate association between risk factor and disease occurrence. *Int J Epidemiol*. 1986;15:22–29. doi:10.1093/ije/15.1.22

Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. *Methods*. 2001;25:402–408. doi:10.1006/meth.2001.1262

Lo JH, Lin C-M, Chen MJ, Chen TT. Altered gene expression patterns of innate and adaptive immunity pathways in transgenic rainbow trout harboring Cecropin P1 transgene. *BMC Genomics*. 2014;15:887. doi:10.1186/1471-2164-15-887

Long S, Milev-Milovanovic I, Wilson M, Bengten E, Clem LW, Miller NW, et al. Identification and expression analysis of cDNAs encoding channel catfish type I interferons. *Fish Shellfish Immunol*. 2006;21:42–59. doi:10.1016/j.fsi.2005.10.008

Ma J, Fan Y, Zhou Y, Liu W, Jiang N, Zhang J, et al. Efficient resistance to grass carp reovirus infection in JAM-A knockout cells using CRISPR/Cas9. *Fish Shellfish Immunol*. 2018;76:206–215. doi:10.1016/j.fsi.2018.02.039

Mandal BK, Chen H, Si Z, Hou X, Yang H, Xu X, et al. Shrunk and scattered black spots turn out due to MC1R knockout in a white-black Oujiang color common carp (*Cyprinus carpio* var. color). *Aquaculture*. 2020;518:734822. doi:10.1016/j.aquaculture.2019.734822

Mao W, Wang Y, Wang W, Wu B, Feng J, Zhu Z. Enhanced resistance to *Aeromonas hydrophila* infection and enhanced phagocytic activities in human lactoferrin-transgenic grass carp (*Ctenopharyngodon idellus*). *Aquaculture*. 2004;242:93–103. doi:10.1016/j.aquaculture.2004.07.020

Mookherjee N, Anderson MA, Haagsman HP, Davidson DJ. Antimicrobial host defence peptides: functions and clinical potential. *Nat Rev Drug Discov*. 2020;19:311–332. doi:10.1038/s41573-019-0058-8

Mosimann C, Kaufman CK, Li P, Pugach EK, Tamplin OJ, Zon LI. Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish. *Development*. 2011;138:169–177. doi:10.1242/dev.059345

Ostir GV, Markides KS, Black SA, Goodwin JS. Emotional well-being predicts subsequent functional independence and survival. *J Am Geriatr Soc*. 2000;48:473–478. doi:10.1111/j.1532-5415.2000.tb04991.x

Ostir GV, Uchida T. Logistic regression: a nontechnical review. *Am J Phys Med Rehabil*. 2000;79:565–572. doi:10.1097/00002060-200011000-00017

Ota S, Hisano Y, Ikawa Y, Kawahara A. Multiple genome modifications by the CRISPR/Cas9 system in zebrafish. *Genes Cells*. 2014;19:555–564. doi:10.1111/gtc.12154

Peatman E, Bao B, Peng X, Baoprasertkul P, Brady Y, Liu Z. Catfish CC chemokines: genomic clustering, duplications, and expression after bacterial infection with *Edwardsiella ictaluri*. *Mol Genet Genomics*. 2006;275:297–309. doi:10.1007/s00438-005-0081-9

Peng K-C, Pan C-Y, Chou H-N, Chen J-Y. Using an improved Tol2 transposon system to produce transgenic zebrafish with epinecidin-1 which enhanced resistance to bacterial infection. *Fish Shellfish Immunol*. 2010;28:905–917. doi:10.1016/j.fsi.2010.02.003

Pridgeon JW, Mu X, Klesius PH. Expression profiles of seven channel catfish antimicrobial peptides in response to *Edwardsiella ictaluri* infection. *J Fish Dis*. 2012;35:227–237. doi:10.1111/j.1365-2761.2011.01343.x

Qin G. Gene Editing and Hormone Therapy to Control Reproduction in Channel Catfish, *Ictalurus punctatus*. Doctoral dissertation. Auburn University, AL, USA. 2019.

Rajendran KV, Zhang J, Liu S, Kucuktas H, Wang X, Liu H, et al. Pathogen recognition receptors in channel catfish: I. Identification, phylogeny and expression of NOD-like receptors. *Dev Comp Immunol*. 2012;37:77–86. doi:10.1016/j.dci.2011.12.005

Rauta PR, Nayak B, Das S. Immune system and immune responses in fish and their role in comparative immunity study: a model for higher organisms. *Immunol Lett*. 2012;148:23–33. doi:10.1016/j.imlet.2012.08.003

Rodrigues G, Maximiano MR, Franco OL. Antimicrobial peptides used as growth promoters in livestock production. *Appl Microbiol Biotechnol*. 2021;105:7115–7121. doi:10.1007/s00253-021-11540-3

Sarmasik A, Warr G, Chen TT. Production of transgenic medaka with increased resistance to bacterial pathogens. *Mar Biotechnol*. 2002;4:310–322. doi:10.1007/s10126-002-0023-z

Sha Z, Abernathy JW, Wang S, Li P, Kucuktas H, Liu H. NOD-like subfamily of the nucleotide-binding domain and leucine-rich repeat containing family receptors and their expression in channel catfish. *Dev Comp Immunol*. 2009;33:991–999. doi:10.1016/j.dci.2009.04.004

Shahi N, Mallik SK, Sarma D. Muscle growth in targeted knockout common carp (*Cyprinus carpio*) *mstn* gene with low off-target effects. *Aquaculture*. 2022;547:737423. doi:10.1016/j.aquaculture.2021.737423

Silveira RF, Roque-Borda CA, Vicente EF. Antimicrobial peptides as a feed additive alternative to animal production, food safety and public health implications: An overview. *Anim Nutr*. 2021;7:896–904. doi:10.1016/j.aninu.2021.01.004

Simora RMC, Xing D, Bangs MR, Wang W, Ma X, Su B, et al. CRISPR/Cas9-mediated knock-in of alligator cathelicidin gene in a non-coding region of channel catfish genome. *Sci Rep*. 2020;10:22271. doi:10.1038/s41598-020-79409-5

Sperandei S. Understanding logistic regression analysis. *Biochem Med*. 2014;24:12–18. doi:10.11613/BM.2014.003

Uribe C, Folch H, Enriquez R, Moran G. Innate and adaptive immunity in teleost fish: a review. *Vet Med*. 2011;56:486–503. doi:10.17221/3294-VETMED

Wang J, Cheng Y. Enhancing aquaculture disease resistance: Antimicrobial peptides and gene editing. *Rev Aquc*. 2023. doi:10.1111/RAQ.12845.

Wang J, Su B, Dunham RA. Genome-wide identification of catfish antimicrobial peptides: A new perspective to enhance fish disease resistance. *Rev Aquac*. 2022;14:2002–2022. doi:10.1111/raq.12684

Wang J, Wilson AE, Su B, Dunham RA. Functionality of dietary antimicrobial peptides in aquatic animal health: Multiple meta-analyses. *Anim Nutr*. 2023;12:200–214. doi:10.1016/j.aninu.2022.10.001

Wang Q, Bao B, Wang Y, Peatman E, Liu Z. Characterization of a NK-lysin antimicrobial peptide gene from channel catfish. *Fish Shellfish Immunol*. 2006a;20:419–426. doi:10.1016/j.fsi.2005.05.005

Wang Y, Wang Q, Baoprasertkul P, Peatman E, Liu Z. Genomic organization, gene duplication, and expression analysis of interleukin-1β in channel catfish (*Ictalurus punctatus*). *Mol Immunol*. 2006b;43:1653–1664. doi:10.1016/j.molimm.2005.09.024

Xing D, Su B, Li S, Bangs M, Creamer D, Coogan M. CRISPR/Cas9-mediated transgenesis of the Masu salmon (*Oncorhynchus masou*) elovl2 gene improves n-3 fatty acid content in channel catfish (*Ictalurus punctatus*). *Mar Biotechnol*. 2022a;24:513–523. doi:10.1007/s10126-022-10110-6

Xing D, Su B, Bangs M, Li S, Wang J, Bern L, et al. CRISPR/Cas9-mediated knock-in method can improve the expression and effect of transgene in P1 generation of channel catfish (*Ictalurus punctatus*). *Aquaculture*. 2022b;560:738531. doi:10.1016/j.aquaculture.2022.738531

Xing D, Li S, Shang M, Wang W, Zhang Q, Wang J, et al. A new strategy for increasing knock-in efficiency: Multiple elongase and desaturase transgenes knock-in by targeting long repeated sequences. *ACS Synth Biol*. 2022c;11:4210–4219. doi:10.1021/acssynbio.2c00252

Xu P, Bao B, He Q, Peatman E, He C, Liu Z. Characterization and expression analysis of bactericidal permeability-increasing protein (BPI) antimicrobial peptide gene from channel

catfish *Ictalurus punctatus*. *Dev Comp Immunol*. 2005;29:865–878. doi:10.1016/j.dci.2005.03.004

Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*. 2013;154:1370–1379. doi:10.1016/j.cell.2013.08.022

Yazawa R, Hirono I, Aoki T. Transgenic zebrafish expressing chicken lysozyme show resistance against bacterial diseases. *Transgenic Res*. 2006;15:385–391. doi:10.1007/s11248-006-0009-0

Zhang J-P, Li X-L, Li G-H, Chen W, Arakaki C, Botimer GD, et al. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol*. 2017;18:35. doi:10.1186/s13059-017-1164-8

CHAPTER FIVE

One-step knock-in of two antimicrobial peptide transgenes at multiple loci of channel catfish (*Ictalurus punctatus*) by CRISPR/Cas9-mediated multiplex genome engineering

Abstract

CRISPR/Cas9-mediated multiplex genome editing (MGE) conventionally uses multiple single-guide RNAs (sgRNAs) for gene-targeted mutagenesis via the non-homologous end joining (NHEJ) pathway. MGE has been proven to be highly efficient for functional gene disruption/knockout (KO) at multiple loci in mammalian cells or organisms. However, in the absence of a DNA donor, this approach is limited to small indels without transgene integration. Here, we establish the linear double-stranded DNA (dsDNA) and double-cut plasmid (dcPlasmid) combination-assisted MGE in channel catfish (*Ictalurus punctatus*), allowing combinational deletion mutagenesis and transgene knock-in (KI) at multiple sites through NHEJ/homology-directed repair (HDR) pathway in parallel. In this study, we used single-sgRNA-based genome editing (ssGE) and multi-sgRNA-based MGE (msMGE) to replace the luteinizing hormone (*lh*) and melanocortin-4 receptor (*mc4r*) genes with the cathelicidin (*As-Cath*) transgene and the myostatin (two target sites: *mstn1*, *mstn2*) gene with the cecropin (*Cec*) transgene, respectively. A total of 9,000 embryos were microinjected from three families, and 1,004 live fingerlings were generated and analyzed. There was no significant difference in hatchability (all $P > 0.05$) and fry survival (all $P > 0.05$) between ssGE and msMGE. Compared to ssGE, CRISPR/Cas9-mediated msMGE assisted by the mixture of dsDNA and dcPlasmid donors yielded a higher knock-in (KI) efficiency of *As-Cath* (19.93%, [59/296] *vs*. 12.96%, [45/347]; $P = 0.018$) and *Cec* (22.97%, [68/296] *vs*. 10.80%, [39/361]; $P = 0.003$) transgenes, respectively. The msMGE strategy can be used to generate transgenic fish carrying two transgenes at multiple loci. In addition, double and quadruple mutant individuals can be produced with high efficiency (36.3% ~ 71.1%) in one-step microinjection. In conclusion, we demonstrated that the CRISPR/Cas9-mediated msMGE allows the one-step generation of simultaneous insertion of the *As-Cath* and *Cec* transgenes at four sites, and the simultaneous disruption of the *lh*, *mc4r*, *mstn1* and *mstn2* alleles. This msMGE system, aided by the mixture donors, promises to pioneer a new dimension in the drive and selection of multiple designated traits in other non-model organisms.

Keywords: Multiplex genome editing, transgenesis, mutagenesis, cathelicidin, cecropin, aquaculture

1. Introduction

CRISPR/Cas9 genome editing technology has profoundly transformed the landscape of genetic engineering by providing an efficient, precise, targeted modification of DNA sequences (Jinek et al., 2012), enabling the creation of new genetic variants. In recent years, the application of this cutting-edge technology to aquaculture has emerged as a promising frontier for cultivating aquatic organisms, such as fish (Gratacap et al., 2019), shellfish (Yu et al., 2019), and shrimp (Sun et al., 2017). Aquaculture constitutes a critical sector of the global food industry, catering to a substantial fraction of the world's seafood demand. Nevertheless, it confronts multifaceted challenges such as disease outbreaks, environmental stressors, and genetic variation that can impede the growth, health, and quality of farmed species (FAO, 2020; Houston et al., 2020). CRISPR/Cas9 holds tremendous promise as a means of overcoming these challenges by facilitating the generation of disease-resistant, faster-growing, and higher-yielding aquaculture species (Gratacap et al., 2019; Wang et al., 2022a).

Cas enzymes can either be programmed to target many genes at once, or multiple single guide RNAs (sgRNAs) can be directed to a single genetic locus to enhance the efficiency of editing. In this context, one of the most exciting developments in genome editing is the advent of multiplex genome editing (MGE). Traditional genome editing tools make the modification of multiple genes a slow and challenging process with low efficiency (Wang et al., 2013; Cong et al., 2013). However, MGE is a groundbreaking technology that offers the simultaneous modification of multiple genes in parallel, enabling unparalleled control over genetic traits within an organism. This breakthrough has permitted researchers to create genetic modifications with high efficiency and investigate the genetic basis of complex traits more efficiently than ever before (Abdelrahman et al., 2021; McCarty et al., 2022). For instance, 80% biallelic mutations (*Tet1* and *Tet2* genes) were observed in the progeny when Cas9 mRNA and multiple sgRNAs were co-injected into mice (*Mus musculus*) zygotes (Wang et al., 2013). Campa et al. (2019) demonstrated that an increased efficiency (up to 60%) can be achieved in human 293T cells using CRISPR/Cas12a-mediated MGE compared to single-sgRNA-based genome editing (ssGE, ranges from 2 to 17%). In addition, multi-sgRNA-based MGE (msMGE) using a CRISPR/Cas9-

nuclease vector targeting seven genes can induce multi-gene mutations at rates ranging from 4.3% to 36.7% in human 293T cells (Sakuma et al., 2014).

In addition to the applications in mammalian cell lines, CRISPR/Cas9-mediated msMGE has monumental implications and myriad applications, which have undergone extensive development in aquaculture, rendering precise, programmable, and multiplex targeting of DNA sequences to generate new fish genetic lines with improved yield, disease resistance, and nutritional value (Elaswad et al., 2018; Gratacap et al., 2019; Coogan et al., 2022a; Xing et al., 2022ab). Not surprisingly, a higher single-locus mutation efficiency was found in zebrafish (*Danio rerio*) (90% *vs*. 80%, Kroll et al., 2021), blotched snakehead (*Channa maculate*) (53.3% *vs*. 40.0%, Ou et al., 2023), channel catfish (*Ictalurus punctatus*) (96.3% *vs*. 88.6%, Khalil et al., 2017; 64% *vs*. 46%, Coogan et al., 2022b) using CRISPR/Cas9-based msMGE compared to the ssGE strategy. Furthermore, msMGE can achieve more than one mutation in multiple genes with a relatively high efficiency. For example, co-injecting two sgRNAs and Cas9 protein targeting the *slc45a2* and *tyr* genes generated mutant Atlantic salmon (*Salmo salar*) with the efficiencies of 40% and 22%, respectively (Edvardsen et al., 2014). Additionally, 61.12% of biallelic mutations (*IGFBP-2b1* and *IGFBP-2b*2 genes) were observed in the offspring when Cas9 protein and two sgRNAs were co-injected into rainbow trout (*Oncorhynchus mykiss*) embryos (Cleveland et al., 2018). Recently, Krug et al. (2023) confirmed that msMGE coupled with three sgRNAs can induce a triple-gene mutation (*mitfa, ltk*, and *csf1ra*) in the African killifish (*Nothobranchius furzeri*) with a high efficiency of 23%. These studies indicated that msMGE can not only improve the efficiency of mutating a single gene, but also can achieve the simultaneous mutation of multiple genes with a high efficiency using multiple sgRNAs in aquacultural species.

Homologous recombination can generate transgenic animals with site-specific insertions if a donor DNA with homology to the ends flanking the double-stranded breaks is co-injected (Meyer et al., 2010; Cui et al., 2011). The ease of design, construction, and co-delivery of multiple CRISPR/Cas9-sgRNA complexes and donor templates suggests the possibility of msMGE-based transgenesis. Although msMGE has been applied to disrupt/knockout (KO) one or multiple genes with considerable efficiency, few studies have used msMGE coupled with donor templates to introduce foreign genes in organisms. Recently, Wang et al. (2022b)

developed an engineered double-stranded DNA (dsDNA)-assisted msMGE system (dReaMGE) and observed kilobase-scale sequence replacements at two loci in bacteria (*Schlegelella brevitalea*) with an efficiency of 80%, accompanied by simultaneous deletions (3 to 10 kb) with the efficiency from 11.5% to 19.1%. In addition, 33.8% and 7.3% of triple and quadruple mutants, respectively, were detected in *Paraburkholderia megapolitana* using the dReaMGE technology (Wang et al., 2022b). Subsequently, Xing et al. (2022a) demonstrated that msMGE coupled with three donor plasmids targeting long repeated sequences, was capable of achieving transgene transfer rates of 11.4% to 25% for single gene insertion, and 7% for triple gene insertion in channel catfish. Similarly, the integration rates of single or double genes at long repeated sequences were 28.1% and 7.8%, or 6.8%, respectively, in channel catfish by using cocktail-designed plasmids as donor templates (Xing et al., 2023). These reports suggest that CRISPR/Cas9-mediated msMGE is a promising approach to achieve multi-transgene insertions at different loci/genes aided by dsDNA or plasmid donors.

Luteinizing hormone (*lh*) serves as a regulator of the reproductive process and plays an important role in the final maturation of fish gametes (Gen et al., 2003; Chu et al., 2014). Previous studies have determined that *lh*-deficient channel catfish are sterile (Qin et al., 2016; Qin, 2019), and the sterilization can be temporarily reversed by hormone therapy in channel catfish (Qin, 2019). In addition, myostatin (*mstn*), a critical growth factor, inhibits myogenesis and hypertrophy for body weight gain (Thomas et al., 2000). Currently, two loci of the myostatin gene (*mstn-a* and *mstn-b*) have been identified in numerous teleost fish species, including channel catfish (Liu et al., 2016; Zhang et al., 2020). In diploid teleosts, *mstn-b* remains responsible for muscle development, whereas *mstn*-a is associated with immune function (Zhang et al., 2020; Coogan et al., 2022a). The focus of the present study was on *mstn-b* in channel catfish, and therefore, unless otherwise noted, all references to myostatin should be taken to refer to *mstn-b*. *Mstn* mutants showed a higher growth rate in channel catfish, red sea bream (*Pagrus major*), and common carp (*Cyprinus carpio*) compared to wild-type (WT) fish (Ohama et al., 2020; Coogan et al., 2022a; Shahi et al., 2022). Similarly, the melanocortin-4 receptor (*mc4r*) controls energy homeostasis (Cone, 2006), and *mc4r*-mutant channel catfish had a faster increase in body weight compared to WT individuals (Coogan et al., 2022b). Additionally, transgenic fish carrying an exogenous antimicrobial peptide gene (AMG) have shown heightened resistance to a variety of pathogens in

channel catfish, blue catfish (*I. furcatus*), grass carp (*Ctenopharyngodon idellus*), Atlantic salmon and other fish species (Wang et al., 2022a; Wang and Cheng, 2023). Although high disease resistance, fast growth or genetic sterility have been established in fish species, no studies have yet documented whether it would be possible to combine all these producer-favorable traits in single fish via one-step CRISPR/Cas9-mediated msMGE. Hypothetically, sterile, fast-growing lines with high disease could generate resistance by simultaneously integrating AMGs at both the reproduction- and growth-related loci.

The ability to perform msMGE helps to enhance our understanding of the genetic underpinnings of complex traits and create multiple producer-desired characteristics in crops and animals (Wang et al., 2013; Abdelrahman et al., 2021). Here, we replaced *lh* gene with the cathelicidin (*As-Cath*) transgene and *mstn* gene with the cecropin (*Cec*) transgene, respectively, using the ssGE method assisted by a double-cut plasmid (dcPlasmid). To confirm the feasibility of simultaneously KI/KO different transgenes/functional genes at multiple loci in one step, we used the CRISPR/Cas9-mediated msMGE to drive both NHEJ-based gene disruption and HDR-based transgene insertion to achieve highly efficient replacement of two AMGs (*As-Cath* and *Cec*) at four targets (*lh*, *mc4r*, *mstn1* and *mstn2*) in channel catfish. Hatchability, fry survival, integration efficiency, off-target events, mosaicism, and protein structure were compared between ssGE and msMGE.

2. Materials and methods

2.1 Ethical statement

Mature channel catfish brood stock (Kmix strain) were cultured in earthen ponds at the Fish Genetics Research Unit, E.W. Shell Fisheries Research Center, Auburn University, AL. All experiments were conducted using the Institutional Animal Care and Use Committee at Auburn University (AU-IACUC) IACUC and AU Biosafety Committee approved protocols.

2.2 Experimental design and preparation of Cas9/sgRNA

We selected the *lh* (GeneBank: NM001200080.1), *mc4r* (GeneBank: LBML01001141.1), and *mstn* (GeneBank: AF396747.1) (Liu et al., 2016) genes as transgene-targeting sites, expecting to generate sterile genetic lines with high growth and disease resistance. Specifically, four single

guide RNAs (sgRNAs) were selected: *lh*_sgRNA, *mc4r*_sgRNA, *mstn1*_sgRNA and *mstn2*_sgRNA targeting the *lh*, *mc4r*, *mstn1* and *mstn2* (exon1 and exon2 of the *mstn* gene) sites, respectively. We used the ssGE1/ssGE2 system to integrate the *As-Cath*/*Cec* transgene at the *mc4r* and *mstn1* loci, respectively. Given the msMGE system, we replaced the *lh* and *mc4r* genes with the *As-Cath* transgene, and *mstn* (*mstn1* and *mstn2*) with the *Cec* transgene, respectively, by co-injecting four designed donor templates. Therefore, the ultimate goal was to introduce two copies of *As-Cath* and *Cec* transgenes at multiple loci. Specifically, the plan involved inserting one copy of *As-Cath* at the *lh* locus and one copy of *As-Cath* at the *mc4r* locus. In addition, one copy of *Cec* was intended to be inserted at the exon1 of *mstn*, and another copy of *Cec* at the exon2 of *mstn*.

The sgRNAs were designed by CHOPCHOP (https://chopchop.cbu.uib.no/) (Labun et al., 2019) based on the scores. Putative off-target sites were excluded with the use of Cas-OFFinder (http://www.rgenome.net/cas-offinder/) (Bae et al., 2014). Selected sgRNAs were synthesized *in vitro* using the Maxiscript T7 Kit (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's instructions, then purified using the RNA Clean and Concentrator Kit (Zymo Research, Irvine, CA). The Nanodrop 2000 spectrophotometer (Thermo Scientific, USA) and a 2% agarose gel with 1 × TBE buffer were used to assess the concentration and quality of sgRNAs, respectively. Synthetic sgRNAs were diluted to a concentration of ~300 ng/μL and distributed into PCR tubes (2 μL/tube), then stored at -80 ℃ until needed. The Cas9 protein powder was obtained from PNA BIO Inc. (Newbury Park, CA), diluted to 100 ng/μL with DNase/RNase-free water, keeping at − 80 ℃ until use. The gene-specific oligonucleotides for sgRNA recognition and the universal primer used in this study are listed in Table S13 in Appendix 2.

2.3 Construction of donor templates

The inserted coding sequence (CDS) was derived from the cathelicidin gene of *Alligator sinensis* (*As-Cath*, GeneBank: XM_006037211.3) (Chen et al., 2017) and the cecropin gene of *Hyalophora cecropia* (*Cec*, GeneBank: M34924.1) (Boman et al., 1985). To achieve multiple KIs at four sites, we used a mixed donor of linear double-stranded DNA (dsDNA) and double-cut plasmid (dcPlasmid) based on our previous studies (Simora et al., 2020; Wang et al., 2023).

Therefore, four donors (HA1_UBI_*As-Cath*_pA_HA2, HA3_UBI_*As-Cath*_pA_HA4, HA5_UBI_*Cec*_pA_HA6 and HA7_UBI_*Cec*_pA_HA8) were provided to assist transgene integration at the *lh*, *mc4r*, *mstn1* and *mstn2* sites, respectively (Figure 19). In detail, the dsDNA donor was created by constructing the *As-Cath* CDS flanked by two 300 bp homology arms (HAs) derived from the *lh* gene of channel catfish on either side of the insert sequence. With respect to the dcPlasmid, an *As-Cath*- or *Cec*-dsDNA cassette flanked by two sgRNA recognition sequences (sgRNA-PAM, 23 bp each) was cloned into the pUC57_mini vector at the *Eco*RV enzyme digestion site (Appendix 7). Expression of the *Cec* and *As-Cath* transgenes was driven by the zebrafish ubiquitin (UBI) promoter (Mosimann et al., 2011). The dsDNA/plasmid donors were synthesized by Genewiz LLC (South Plainfield, NJ).

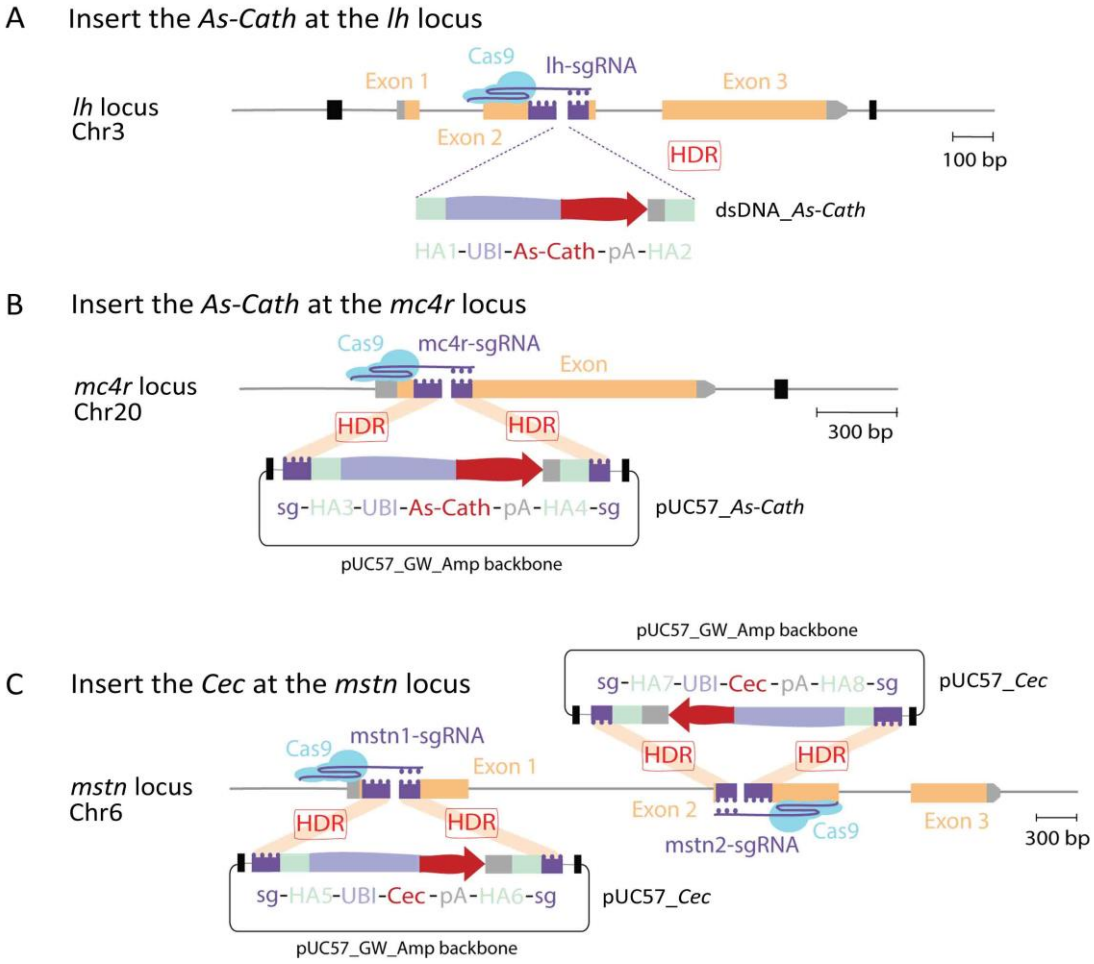


Figure 19. Schematics of homology-directed repair (HDR)-mediated knock-in (KI) multiple donors targeting four loci in the genome of channel catfish (*Ictalurus punctatus*). (**A**) HDR-mediated KI of the alligator cathelicidin (*As-Cath*) gene targeting the *lh* locus coupled with a linear dsDNA donor

(HA1_UBI_*As-Cath*_pA_HA2). (**B**) HDR-mediated KI of the *As-Cath* gene targeting the *mc4r* locus coupled with a double-cut plasmid (dcPlasmid) donor (sg_HA3_UBI_*As-Cath*_pA_HA4_sg). (**C**) HDR-mediated KI of the moth cecropin (*Cec*) gene targeting the *mstn1* and *mstn2* sites coupled with two dcPlasmid donors (sg_HA5_UBI_*Cec*_pA_HA6_sg and sg_HA7_UBI_*Cec*_pA_HA8_sg), respectively. *lh*, luteinizing hormone; *mc4r*, melanocortin-4 receptor; *mstn1*, the exon1 of myostatin; *mstn2*, the exon2 of myostatin; HA, homology arm; UBI, ubiquitin promoter from zebrafish (*Danio rerio*); sg, sgRNA; pA, polyA tail termination.

2.4 Microinjection, transgenic fish production and rearing

Channel catfish spawning, embryo preparation, and microinjection were performed according to previous procedures from our laboratory (Khalil et al., 2017; Wang et al., 2023a) with modifications. In brief, three strategies (ssGE1, ssGE2 and msMGE) were used to generate transgenic/gene-edited channel catfish in this study. CRISPR/Cas9-mediated ssGE1/ssGE2 was used to produce single transgenic (*As-Cath* or *Cec*)/gene-edited (*mc4r* or *mstn1*) by targeting the *mc4r* or *mstn1* site assisted by HA3_UBI_*As-Cath*_pA_HA4 or HA5_UBI_*Cec*_pA_HA6 donor. MsMGE was used to generate multiple transgenic/gene-edited individuals by targeting *lh*, *mc4r*, *mstn1* and *mstn2* in parallel assisted by the mixture of four donors (Figure 20A). In addition to these three microinjected groups (1,000 microinjected embryos per group), non-injected (nCT) (600 embryos) and injected (iCT, 60% phenol red solution only) (600 embryos) groups were utilized. The microinjection was conducted on 6/15/2021, 6/17/2021, and 6/19/2021 to generate fingerlings from three families. All injected embryos and their full-sibling controls were cultured in Holtfreter's solution (Bart and Dunham, 1996) until hatching. During this period, dead embryos and fry were collected for hatchability and fry survival calculations. The fry were subsequently moved into tanks within a recirculating system, maintaining a density of 2 fry/L. They were fed Aquamax fry powder (50% crude protein, 17% crude fat, 3% crude fiber and 12% ash; Purina Animal Nutrition LLC, Shoreview, MN) four times daily to satiation two months. The fingerlings were then fed Aquaxcel WW Fish Starter 4512 (45% crude protein, 12% crude fat, 3% crude fiber, and 1% phosphorus; Cargill Animal Nutrition, Minneapolis, MN) twice daily to satiation until sampling.

2.5 Integration analysis and mutation detection

After 4 months of rearing, all fingerlings (20 − 50 g) were pit-tagged (Biomark Inc., Boise, Idaho, USA) and individually weighed. Genomic DNA was extracted from the mixture of fin clips and

barbels from each injected fish. Specific primers (Table S13 in Appendix 1) were designed using the online software Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) to screen for transgenes by detecting the UBI promoter and CDS of transgenes in all microinjected fish. The 5′ and 3′ junction regions were then detected, and off-target events were determined by PCR. For the integrated *As-Cath*/*Cec*, promoter and junction sequences, PCR of positive samples was performed in a 50 μL volume of system. Then PCR products were purified using the QIAquick^{R} PCR Product Purification Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. Before sequencing, all purified DNA samples were quantified and identified by Nanodrop and 1.0% agarose gels. The sequencing results were blasted with transgenes using MAFFT (version 7, https://mafft.cbrc.jp/alignment/server/) to identify inserted DNA sequences (Figure 20B).

With respect to the potential mutant individuals, PCR products were inserted into the pCR^{TM}4-TOPO vector using the TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA) and transformed into One Shot TOP10F chemically competent *Escherichia coli* (Invitrogen, Carlsbad, CA) as previously described (Elaswad et al., 2018). Three colonies were then randomly picked for colony PCR, and liquid *E. coli* cultures were prepared for rolling circle amplification (RCA) sequencing by Sequetech (Mountain View, CA) (using a primer M13 forward primer). The sequencing results of the mixture of fin clips and barbels were analyzed using TIDE online (https://tide.nki.nl/) (Brinkman et al., 2014) to quantify the editing efficiency and to identify the types of insertions and deletions (indels) in the potential mutant individuals. TIDE analysis calculates the efficiency of gene editing by quantifying the percentage of indels in a target DNA sequence based on the sequencing chromatograms. This is typically achieved by comparing the sequence of edited samples to a control or reference sequence. The higher the editing efficiency, the greater the proportion of indels present in the edited samples. Finally, the 3D protein structures of the mutant genes were predicted using the online tool AlphaFold (https://alphafold.com/) (Jumper et al., 2021).

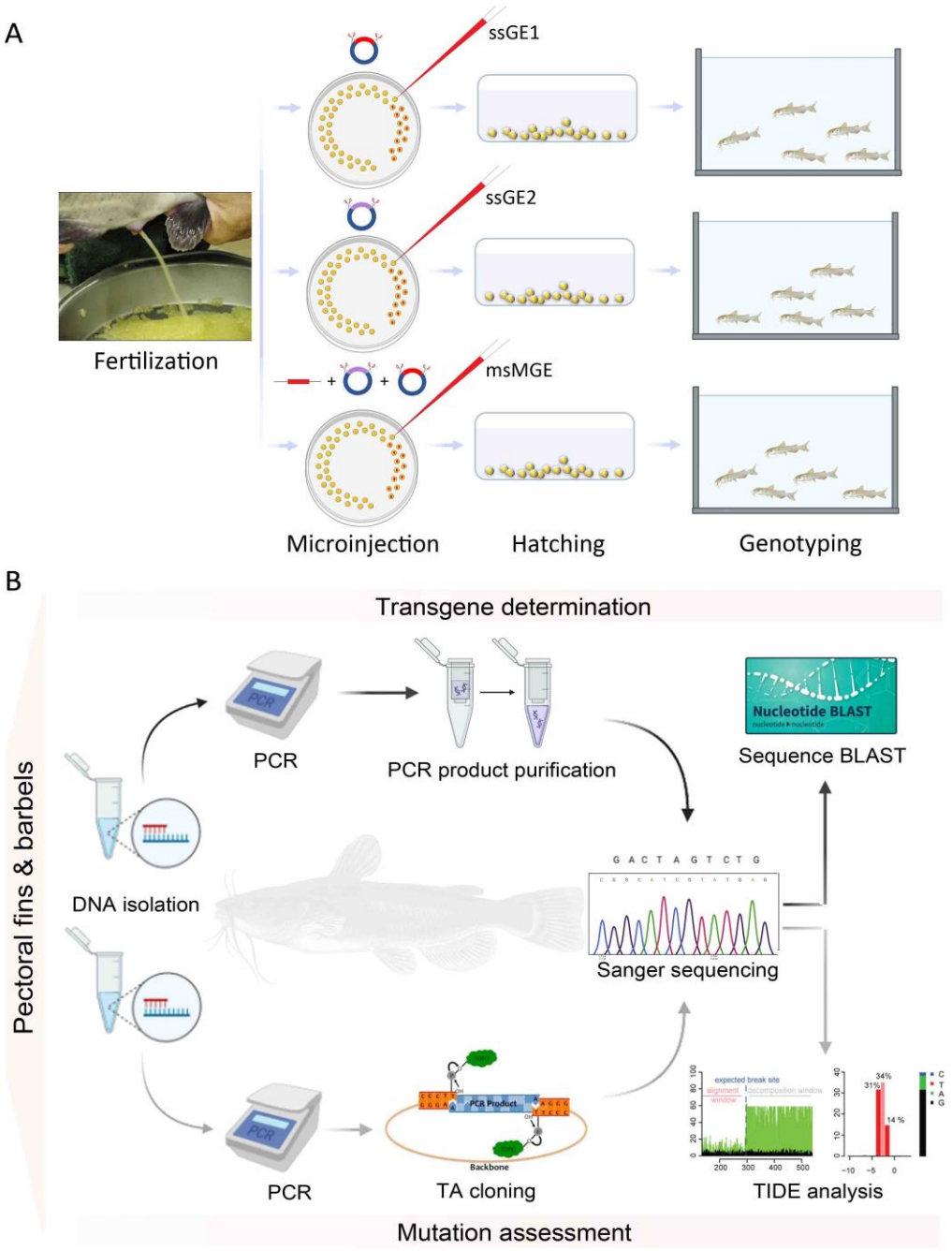


Figure 20. Experimental design and genotyping in channel catfish (*Ictalurus punctatus*) using CRISPR/Cas9-mediated ssGE and msMGE, respectively. (**A**) Microinjection and fish rearing for ssGE1, ssGE2 and msMGE systems. (**B**) Procedure for transgene determination and mutation evaluation of transgenic/gene-edited channel catfish. For the mutation test, TIDE analysis was implemented for each non-transgenic fish. ssGE1, single sgRNA-based genome editing coupled with the pUC57_*As-Cath* donor targeting the *mc4r* locus; ssGE2, single sgRNA-based genome editing coupled with the pUC57_*Cec* donor targeting the *mstn1* site; msMGE, multi-sgRNA-based multiplex genome editing coupled with the mixture of dsDNA_*As-Cath*, pUC57_*As-Cath* and pUC57_*Cec* donors targeting the *lh*, *mc4r*, *mstn1* and

mstn2 sites in parallel; *lh*, luteinizing hormone; *mc4r*, melanocortin-4 receptor; *mstn1*, the exon1 of myostatin; *mstn2*, the exon2 of myostatin; TIDE, tracking of indels by decomposition (http://tide.nki.nl).

2.6 Determination of mosaicism

Three 12-month-old positive transgenic channel catfish (*As-Cath*-, *Cec*- and *As-Cath*/*Cec*-transgenic fish) with a mean body weight of 72.4 g, were collected and euthanized with buffered tricaine methanesulfonate (MS222; Syndel, Ferndale, WA) at 200 ppm. The fin, barbel, skin, muscle, intestine, head kidney, stomach, liver, blood, gill, brain, eye, spleen, and gonad of each individual were collected in 1.5 mL tubes for DNA isolation. To determine the potential mosaicism of the transgene, PCR was performed using specific primers as described above.

2.7 Statistical analysis

Hatchability and fry survival rate from different groups in each family were analyzed using one-way ANOVA followed by Tukey's multiple comparison test. KI efficiency and off-target events between CRISPR/Cas9-mediated ssGE and msMGE (ssGE1 *vs*. msMGE, ssGE2 *vs*. msMGE) were compared using the unpaired Student's *t*-test. All statistical analyses were performed using GraphPad Prism 9.4.1 (GraphPad Software, LLC). The statistical significance level was set at $P < 0.05$, and all data were presented as the mean ± standard deviation (SD).

3. Results

3.1 Hatchability and fry survival

Different families demonstrated significant differences in hatchability and fry survival in both nCT and injected groups (iCT, ssGE1, ssGE2 and msMGE) (Table S14-16 in Appendix 1). Compared with the nCT group, the injected groups showed a significant decrease in hatchability for all three families ($P < 0.01$), but there were no significant differences between iCT, ssGE1, ssGE2, and msMGE (all $P > 0.05$) (Figure 21A). Although significantly lower hatchability was observed in injected groups compared to the nCT group in family 1 and family 3, there was no significant difference in fry survival of these two families ($P = 0.0551$ in family 1; $P = 0.3256$ in family 3). Family 2 showed a significant decrease in fry survival in the msMGE group compared to the nCT group ($P = 0.0136$), but there was no significant difference among ssGE1, ssGE2 and

msMGE (P = 0.4819) (Figure 21B). These results indicated that CRISPR/Cas9-mediated ssGE1, ssGE2 and msMGE significantly reduce the hatchability in channel catfish, but the microinjection did not change the fry survival rate in the early stage compared to the nCT group.

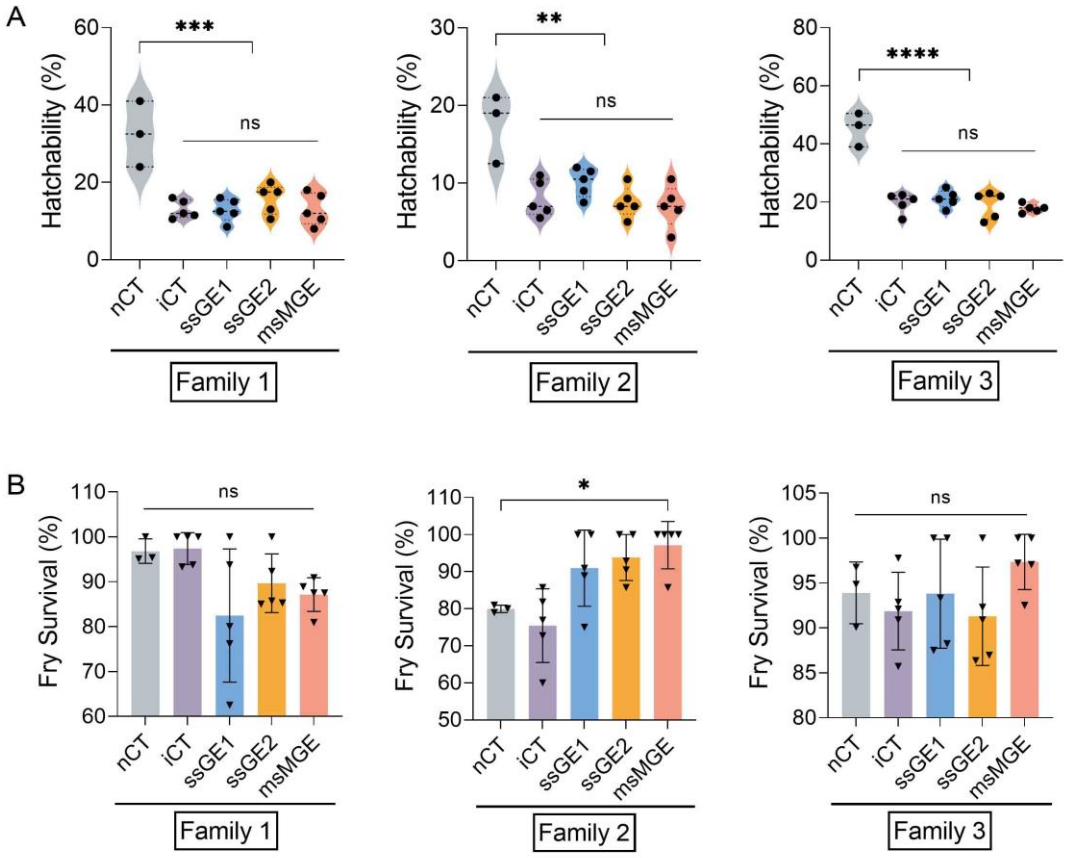


Figure 21. Comparison of hatchability (A) and fry survival (B) of channel catfish (*Ictalurus punctatus*) microinjected at the one-cell stage from three families using three CRISPR/Cas9-mediated knock-in (KI) strategies (ssGE1, ssGE2 and msMGE). Control groups included the injected control (iCT, 60% phenol red solution only) and non-injected control (nCT). Three families were generated on 6/15/2021 (family 1), 6/17/2021 (family 2) and 6/19/2021 (family 3), respectively. ssGE1, single sgRNA-based genome editing coupled with the pUC57_*As-Cath* donor targeting the *mc4r* locus; ssGE2, single sgRNA-based genome editing coupled with the pUC57_*Cec* donor targeting the *mstn1* site; msMGE, multi-sgRNA-based multiplex genome editing coupled with the mixture of dsDNA_*As-Cath*, pUC57_*As-Cath* and pUC57_*Cec* donors targeting the *lh*, *mc4r*, *mstn1* and *mstn2* sites in parallel; *lh*, luteinizing hormone; *mc4r*, melanocortin-4 receptor; *mstn1*, the exon1 of myostatin; *mstn2*, the exon2 of myostatin. **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$; ns, not significant.

3.2 Knock-in efficiencies of transgenes

In the present study, we obtained 347, 361 and 296 fry from the ssGE1, ssGE2 and msMGE groups after microinjecting 9,000 embryos respectively. All the positive individuals were confirmed by gel electrophoresis and sequencing (Figure 22; Figure S16-18 in Appendix 2). Both ssGE and msMGE strategies induced the transgenic/gene-edited channel catfish. The overall KI efficiency of *As-Cath* and *Cec* transgenes in the msMGE system was significantly higher than that of the ssGE1 (19.93% [59/296] *vs*. 12.96% [45/347], $P = 0.0179$ for *As-Cath*) or ssGE2 (22.97% [68/296] *vs*. 10.80% [39/361], $P = 0.0027$ for *Cec*) (Figure 23A). For both *As-Cath* and *Cec* transgenes, a higher off-target rate was observed in msMGE compared to ssGE1 (59.32% [35/59] *vs*. 37.78% [17/45], $P = 0.1163$ for *As-Cath*) or ssGE2 (66.18% [45/68] *vs*. 43.59% [17/39], $P = 0.0050$ for *Cec*) (Figure 23B). In addition, we obtained double (9.67%, 3/31), triple (6.45%, 2/31), and quadruple (9.67%, 3/31) transgenic fish using CRISPR/Cas9-mediated msMGE (Table 10).

Table 10. Genotyping of on-target insertion types for cathelicidin (*As-Cath*) or cecropin (*Cec*) transgenic channel catfish (*Ictalurus punctatus*) using CRISPR/Cas9-mediated msMGE. msMGE, multi-sgRNA-based multiplex genome editing coupled with the mixture of dsDNA_*As-Cath*, pUC57_*As-Cath* and pUC57_*Cec* donors targeting the *lh*, *mc4r*, *mstn1* and *mstn2* sites in parallel; *lh_As-Cath*, the *As-Cath* transgene is integrated at the *lh* locus; *mc4r_As-Cath*, the *As-Cath* transgene is integrated at the *mc4r* locus; *mstn1_Cec*, the *Cec* transgene is integrated at the *mstn1* site; *mstn2_Cec*, the *Cec* transgene is integrated at the *mstn2* site. The on-target positive fish were identified using specific primers for each DNA sample (the mixture of fin and barbel) of a single fish.

Fish ID	*lh_As-Cath*	*mc4r_As-Cath*	*mstn1_Cec*	*mstn2_Cec*	Copy of *As-Cath*	Copy of *Cec*
1	√	√			2	0
2				√	0	1
3		√	√	√	1	2
4		√			1	0
5				√	0	1
6	√	√	√	√	2	2
7				√	0	1
8		√			1	0
9				√	0	1
10	√	√	√	√	2	2
11		√	√	√	1	2
12		√			1	0
13				√	0	1
14		√			1	0
15		√			1	0
16			√		0	1
17		√			1	0
18				√	0	1
19		√			1	0
20			√	√	0	2
21		√		√	1	1

22		√			1	0
23	√	√	√	√	2	2
24				√	0	1
25		√			1	0
26		√			1	0
27				√	0	1
28		√			1	0
29		√			1	0
30				√	0	1
31		√			1	0
Total	**4**	**20**	**7**	**16**		

Although HDR-mediated KIs were dominant when donor templates were supplied, the NEEJ-mediated KOs were also detectable in our present study. In both ssGE1 and ssGE2 systems, the percentage of HDR-KI was significantly higher than that of NHEJ-KO (62.22% [28/45] *vs*. 4.97% [15/302], $P = 0.0037$ for ssGE1; 56.41% [22/39] *vs*. 2.48% [8/322], $P = 0.0001$ for ssGE2) (Figure 23C). Regarding the msMGE system, HDR-KI accounted for a significantly higher rate compared to NHEJ-KO for both *As-Cath* and *Cec* transgenes (40.68% [24/59] *vs*. 4.64% [11/237], $P = 0.0001$ for *As-Cath*; 33.82% [23/68] *vs*. 7.02% [16/228], $P = 0.0013$ for *Cec*) (Figure 23D). In this case, we not only obtained transgenic fish but also gene mutant individuals.

Table 11. Genotyping of mutations in gene-edited channel catfish (*Ictalurus punctatus*) using CRISPR/Cas9-mediated ssGE. The efficiency was calculated using the online tool TIDE (https://tide.nki.nl/) for each DNA sample (the mixture of fin and barbel) of a single fish. TIDE analysis calculates the efficiency of gene editing by quantifying the percentage of indels in a target DNA sequence based on the sequencing chromatograms. The "Protein" column indicates whether the mutated sequences result in an alteration in the protein structure. If "Yes", the protein structure has changed, otherwise not. ssGE1, single sgRNA-based genome editing coupled with the pUC57_*As-Cath* donor targeting the *mc4r* locus; ssGE2, single sgRNA-based genome editing coupled with the pUC57_*Cec* donor targeting the *mstn1* site.

Fish ID	ssGE1		Protein	Fish ID	ssGE2		Protein
	***mc4r*-Indel**	**Efficiency (%)**			***mstn1*-Indel**	**Efficiency (%)**	
32	+1, -2	39.2	Yes	47	+1, -5	35.1	Yes
33	+1	71.4	Yes	48	-3	47.1	No
34	-2/+2	45.7	No	49	-5	53.2	Yes
35	-2	35.3	Yes	50	-3	48.9	No
36	-2	34.6	Yes	51	+1	67.2	Yes
37	-2	33.0	Yes	52	+1, -5	84.0	Yes
38	+1, -2	43.4	Yes	53	-6	32.6	No
39	-4	71.1	Yes	54	-6	41.8	No
40	-7	82.4	Yes				
41	+1	69.3	Yes				
42	+4, -2	43.2	Yes				
43	+1, -2	45.3	Yes				

44	-2	28.6	Yes		
45	-7	84.1	Yes		
46	+1	68.9	Yes		
Total	**15**			**Total**	**8**

3.3 Determination of mutagenesis in sequences

Sanger sequencing and TIDE results with distinct indel spectra at different frequencies in the ssGE and msMGE systems confirmed that the target loci were mutated within or outside the Cas9 cleavage site. These sequence mutations included substitutions, deletions, and insertions that caused or did not cause alterations in the protein domain. Single-locus mutations were induced at *mc4r* in 15 fish and at *mstn1* in eight fish using ssGE1 and ssGE2 strategies, respectively. Similar mean mutation frequency was determined for the *mc4r* and *mstn1* sites in ssGE (*mc4r vs*. *mstn1*: 53.03% [28.6% ~ 84.1%] *vs*. 62.04% [35.1% ~91.6%], $P = 0.2511$) (Table 11). For example, TIDE determined that individual #32 showed that 29.2% of the *mc4r* sequences in the editing cell mix carried an indel, with 25.3% ($R^2 = 0.78$, $P < 0.001$) being a +1 bp insertion (a G nucleotide) on the sense strand. A 2-bp deletion with the frequency of 3.9% was not significantly detected by TIDE in #32 (Figure 24A). In addition, a 5 bp deletion and a 1 bp insertion were found at the *mstn1* site in fish #47 with editing efficiencies of 5.1% and 30.0%, respectively ($R^2 = 0.99$, $P < 0.001$) (Figure 24B).

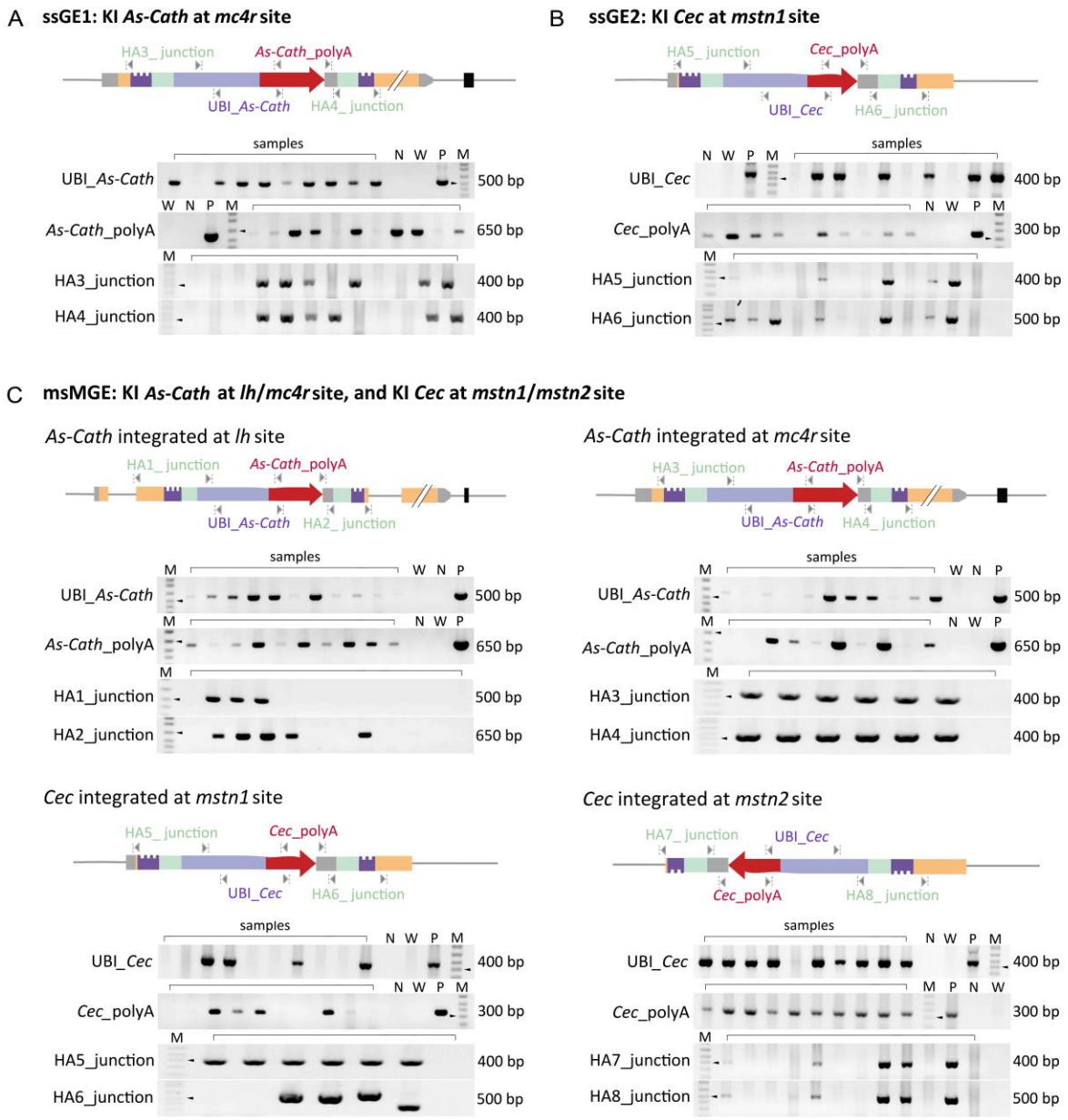


Figure 22. Genotyping strategy and PCR amplification for transgene determination of ssGE1 (A), ssGE2 (B) and msMGE (C) in channel catfish (*Ictalurus punctatus*). The transgenes (*As-Cath* and *Cec*) were driven by the zebrafish ubiquitin (UBI) promoter. To test the integrity of the inserted donors, the promoter-transgene region (UBI_*As-Cath*/UBI_*Cec*), transgene-polyA region (*As-Cath*_polyA/*Cec*_polyA) and the 5' and 3' junction regions (HA_junction) were amplified using specific primers. The 300 bp, 400 bp, 500 bp and 650 bp bands are indicated by black triangles on the M lane. ssGE1, single sgRNA-based genome editing coupled with the pUC57_*As-Cath* donor targeting the *mc4r* locus; ssGE2, single sgRNA-based genome editing coupled with the pUC57_*Cec* donor targeting the *mstn1* site; msMGE, multi-sgRNA-based multiplex genome editing coupled with the mixture of dsDNA_*As-Cath*, pUC57_*As-Cath* and pUC57_*Cec* donors targeting the *lh*, *mc4r*, *mstn1* and *mstn2* sites in parallel; *Cec*, the cecropin gene from moth; *As-Cath*, the cathelicidin gene from alligator; *lh*, luteinizing hormone; *mc4r*, melanocortin-4 receptor; *mstn1*, the exon1 of myostatin; *mstn2*, the exon2 of

myostatin; HA, homology arm; Lane N, water as the blank control; Lane W, a wild-type individual as the negative control; Lane P, a plasmid as the positive control; lane M, 1kb+ DNA marker. The presence of a distinct band indicates positive for transgene. The gel electrophoresis images shown here are cropped, and full-length gels are presented in Figure S16-18 in Appendix 2.

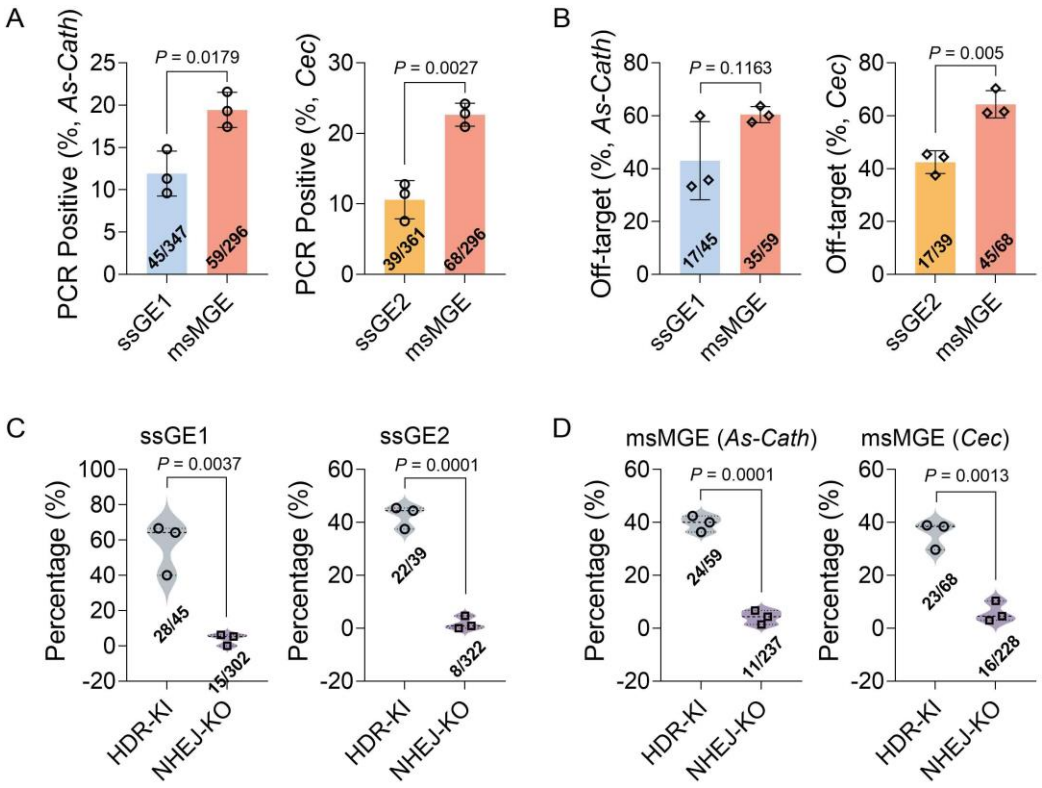


Figure 23. Knock-in (KI) efficiency and mutation rate of CRISPR/Cas9-mediated ssGE and msMGE in channel catfish (*Ictalurus punctatus*). (**A-B**) Comparison of integration rates and off-target events between ssGE and msMGE (ssGE1 *vs*. msMGE and ssGE2 *vs*. msMGE). (**C-D**) Comparison of HDR-KI (transgenesis) and NHEJ-KO (mutagenesis) between ssGE and msMGE (ssGE1 *vs*. msMGE and ssGE2 *vs*. msMGE). Sample size is displayed on each column. HDR-KI, homology-directed repair-mediated knock-in; NHEJ-KO, nonhomologous end joining-mediated knock out; ssGE1, single sgRNA-based genome editing coupled with the pUC57_*As-Cath* donor targeting the *mc4r* locus; ssGE2, single sgRNA based genome editing coupled with the pUC57_*Cec* donor targeting the *mstn1* site; msMGE, multi-sgRNA-based multiplex genome editing coupled with the mixture of dsDNA_*As-Cath*, pUC57_*As-Cath* and pUC57_*Cec* donors targeting the *lh*, *mc4r*, *mstn1* and *mstn2* sites in parallel; *lh*, luteinizing hormone; *mc4r*, melanocortin-4 receptor; *mstn1*, the exon1 of myostatin; *mstn2*, the exon2 of myostatin; *As-Cath*, the cathelicidin transgene from alligator; *Cec*, the cecropin transgene from moth.

The *lh* locus (29.81% [19.2% ~ 40.7%]) had a lower average mutation rate compared to the *mc4r* (51.97% [35.5% ~ 63.8%]) (P = 0.1512), *mstn1* (45.88% [19.4% ~ 71.1%]) (P = 0.1029) and *mstn2* (65.27% [50.0 % ~ 80.5%]) (P = 0.0090) sites in the msMGE strategy (Table 12). In detail,

fish #1’ showed an efficiency of 17.8% with a 1 bp insertion at the *lh* locus, and the +1 bp consisted almost exclusively of an A nucleotide on the sense strand (R^2 = 0.53, P < 0.001) (Figure 24C). Although the same sgRNA (*mc4r*-sgRNA) targeting the *mc4r* locus was used in ssGE1 and msMGE, TIDE analysis revealed different indels and efficiencies in ssGE1 and msMGE. For example, fish #3’ had a 35.5% editing efficiency in the *mc4r* locus, consisting of 19.4% insertion (+1 bp, a C nucleotide) and 16.1% deletion (-2 bp) (R^2 = 0.92, P < 0.001) (Figure 24D). Similarly, different indel types and efficiencies were observed at the *mstn1* site using ssGE2 and msMGE. For instance, fish #4’ showed that 50.5% of *mstn* sequences at the exon1 in the fin/barbel mix carried various indels, of which 46.6% were a T nucleotide insertion, 11.2% a 5 bp deletion, and 5.8% a 2 bp deletion (R^2 = 0.93, P < 0.001) (Figure 24E). Two types of deletions (-1 bp and -3 bp) with efficiencies of 66.7% and 13.8%, respectively, were detected in fish #6’ (R^2 = 0.96, P < 0.001) (Figure 24F). In addition to single-locus mutant fish, msMGE induced double (#15’ and #18’) and quadruple (#5’ and #8’) mutant individuals with distinct indel spectra (Table 12).

3.4 Prediction of mutant proteins

Both ssGE (ssGE1 and ssGE2) and msMGE induced single-locus mutants, and the mutations were predicted to lead to frame shifts yielding non-functional proteins or to amino acid (AA) substitutions that alter protein structures. Fourteen of the 15 *mc4r* mutants showed structure-modified proteins based on the AlphaFold prediction from the ssGE1 strategy. For example, a G nucleotide insertion and a 2 bp deletion (TC) were detected in fish #32, leading to large amounts of mutated AA sequences at the position 44 and beyond. This mutation resulted in a non-functional melanocortin-4 receptor protein. By contrast, 50% of the *mstn1* mutations (4/8) altered the protein structure using the ssGE2 strategy (Table 11, Appendix 8). Specifically, a 1 bp insertion and a 5 bp deletion were observed at the *mstn1* site in individual #47. These indels altered AA sequences at position 11 and beyond, which also silenced myostatin function (Figure 25A).

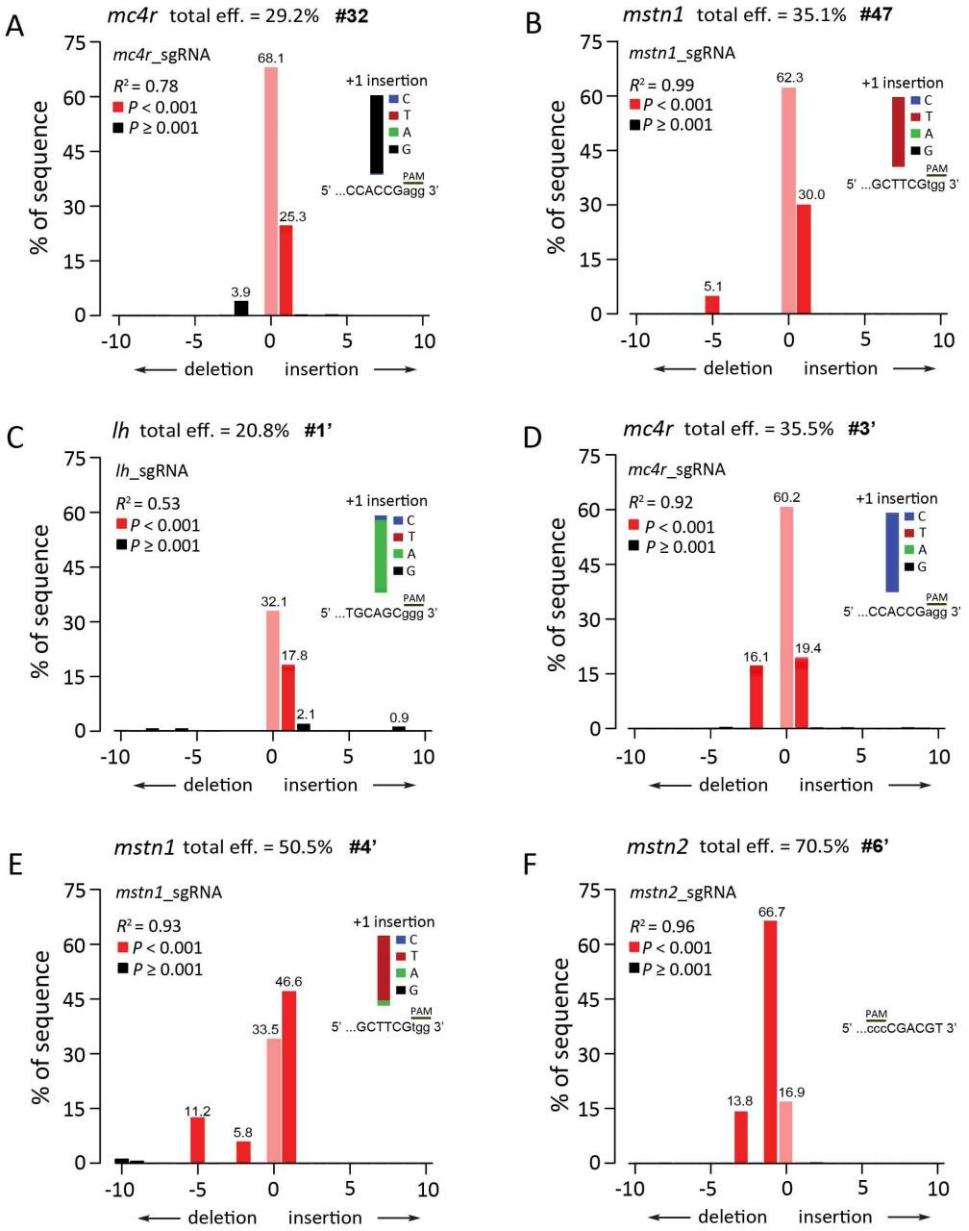


Figure 24. Indel spectrum and the frequency determined by TIDE analysis in channel catfish (*Ictalurus punctatus*) from CRISPR/Cas9-mediated ssGE and msMGE. (**A-B**) TIDE analysis of the *mc4r* and *mstn* genes targeted by Cas9/sgRNA complex using ssGE1 and ssGE2, respectively. (**C-F**) TIDE analysis of the *lh*, *mc4r*, *mstn1* and *mstn2* sites targeted by Cas9/sgRNA complex using msMGE. ssGE1, single sgRNA-based genome editing coupled with the pUC57_*As-Cath* donor targeting the *mc4r* locus; ssGE2, single sgRNA-based genome editing coupled with the pUC57_*Cec* donor targeting the *mstn1* site; msMGE, multi-sgRNA-based multiplex genome editing coupled with the mixture of dsDNA_*As-Cath*, pUC57_*As-Cath* and pUC57_*Cec* donors targeting the *lh*, *mc4r*, *mstn1* and *mstn2* sites in parallel; TIDE, tracking of indels by decomposition (http://tide.nki.nl); *lh*, luteinizing hormone; *mc4r*, melanocortin-4 receptor; *mstn1*, the exon1 of myostatin; *mstn2*, the exon2 of myostatin; PAM, protospacer adjacent motif; R^2 and *P* values indicate the statistical significance for each indel based on the goodness-of-fit test.

Table 12. Genotyping of mutations in gene-edited channel catfish (*Ictalurus punctatus*) using CRISPR/Cas9-mediated msMGE. The efficiency was calculated using the online tool TIDE (https://tide.nki.nl/) for each DNA sample (the mixture of fin and barbel) of a single fish. TIDE analysis calculates the efficiency of gene editing by quantifying the percentage of indels in a target DNA sequence based on the sequencing chromatograms. The "Protein" column indicates whether the mutated sequences result in an alteration in the protein structure. If "Yes", the protein structure has changed, otherwise not. msMGE, multi-sgRNA-based multiplex genome editing coupled with the mixture of dsDNA_*As-Cath*, pUC57_*As-Cath* and pUC57_*Cec* donors targeting the *lh*, *mc4r*, *mstn1* and *mstn2* sites in parallel.

Fish ID	*lh*		*mc4r*		*mstn1*		*mstn2*		Mutagenesis	Protein
	Indel	Efficiency	Indel	Efficiency	Indel	Efficiency	Indel	Efficiency		
1'	+1, +2, -4/+4	20.8%							*lh*	Yes
2'	+1, -3	19.2%							*lh*	Yes
3'			+1, -2	35.5%					*mc4r*	Yes
4'					+1, -2, -5	63.6%			*mstn1*	No
5'	+1, -3	40.7%	+2, +3	63.8%	-1/+1	71.1%	+1, -8	50.0%	*lh*, *mc4r*, *mstn1*, *mstn2*	-
6'							-1, -7	80.5%	*mstn2*	Yes
7'					+1, +3, -5	56.9%			*mstn1*	Yes
8'	+1, -5	33.3%	-5	56.6%	+1, +3, -5	61.0%	+1, -8	65.3%	*lh*, *mc4r*, *mstn1*, *mstn2*	-
9'					+3	34.4%			*mstn1*	No
10'					+1, -3	66.1%			*mstn1*	Yes
11'					+3	58.2%			*mstn1*	No
12'					+1, -3	40.1%			*mstn1*	Yes
13'					-1/+1	32.3%			*mstn1*	No
14'	-2/+2, -3	29.1%							*lh*	Yes
15'	+1, -5	30.3%			+1,-3	29.8%			*lh*, *mstn1*	-
16'					+1, -3/+3	19.4%			*mstn1*	Yes
17'					+1, -3	21.2%			*mstn1*	Yes
18'	+1, +2, -2	36.3%			+3	42.4%			*lh*, *mstn1*	-
19'	+1, -5	28.8%							*lh*	Yes
Total	**8**		**3**		**13**		**3**			

Fifteen gene-edited fish exhibited mutations at a single site (4 *lh*-mutant, 1 *mc4r*-mutant, 9 *mstn1*-mutant and 1 *mstn2*-mutant), and the induced single-locus mutations at the *lh*, *mc4r*, *mstn1* or *mstn2* site showed a different spectrum of indels in the msMGE. All mutations of *lh*, *mc4r* and *mstn2* caused alternations in the tertiary structure of the corresponding proteins. Intriguingly, 44.44% (4/9) of individuals carrying *mstn1* mutations had no change in the structure of the myostatin protein (Table 12, Appendix 8). Representative individual sequences of the mutants and the corresponding predicted protein structures are shown in Figure 25B. Gene-edited individuals #1', #3' and #6' produced non-functional/altered luteinizing hormone, melanocortin-4 receptor, and myostatin proteins, respectively. However, even though a 5 bp deletion of *mstn1* was induced in individual #4', the myostatin protein retained the same domain as the WT control. This was because the mutation at position 12 (near the start codon) resulted in only a few AA substitutions (FVV changed to L), which did not cause any subsequent AA alterations (Figure 25B).

The msMGE strategy also generated double and quadruple mutant individuals, and these mutations mutagenized protein loss of function. Fish #18' showed the *lh* (+1 bp, +2 bp, -2 bp) and *mstn1* (+3 bp) mutations with different editing efficiencies of 36.3% and 42.4%, respectively (Figure 26AB). Fish #5' had quadruple mutations at the *lh* (+1 bp, -3 bp), *mc4r* (+2 bp, +3 bp), *mstn1* (-1/+1 bp), and *mstn2* (+1 bp, -8 bp) with parallel efficiencies of 40.7%, 63.8%, 71.1% and 50.0% (Figure 26AC). The mutant *lh* in fish #5'/#18', and mutant *mc4r* in fish #5' induced non-functional luteinizing hormone and melanocortin-4 receptor proteins. Furthermore, combined variants of *mstn1* and *mstn2* in fish #5' resulted in a domain-altered myostatin protein. However, mutant *mstn1* did not change the myostatin protein in fish #18' (Figure 26D). In addition, individuals #15' (30.3% at the *lh* and 29.8% at the *mstn1*) and #8' (33.3% at the *lh*, 56.6% at the *mc4r*, 61.0% at the *mstn1*, and 65.3% at the *mstn2*) had double and quadruple mutations (Table 12). These modified sequences lead to a non-functional luteinizing hormone and a domain-altered myostatin in fish #15'. And fish #8' generated defective luteinizing hormone, melanocortin-4 receptor, and myostatin in parallel due to the modified sequences (Appendix 8).

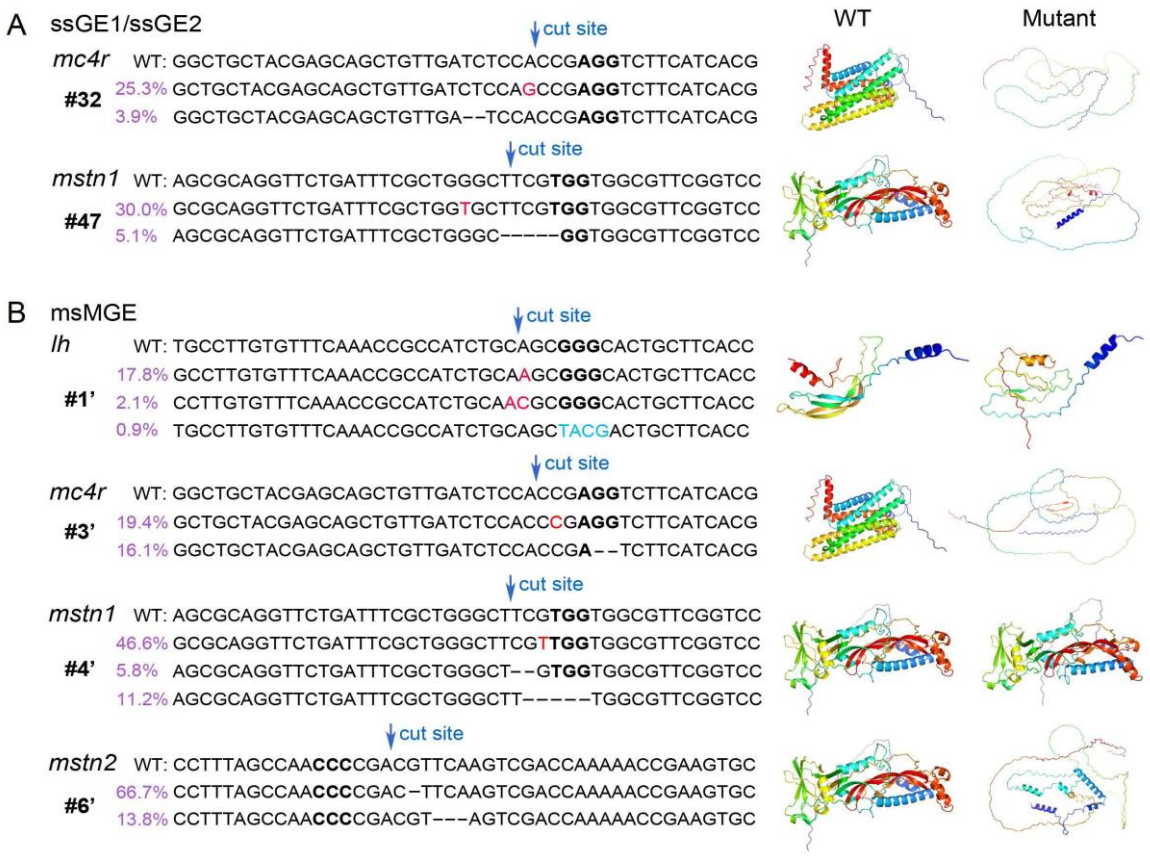


Figure 25. Mutated sequences and corresponding predicted 3D protein structures of the single gene in channel catfish (*Ictalurus punctatus*) using CRISPR/Cas9-mediated ssGE and msMGE. Sanger sequencing and AlphaFold were used for sequence determination and protein structure prediction, respectively. Insertions and substitutions are highlighted in red and blue, respectively. Deletions are indicated by dashes. Protospacer adjacent motif is highlighted in bold. The mutation efficiency is shown on the left of each sequence. WT, wild-type control; ssGE1, single sgRNA-based genome editing coupled with the pUC57_*As-Cath* donor targeting the *mc4r* locus; ssGE2, single sgRNA-based genome editing coupled with the pUC57_*Cec* donor targeting the *mstn1* site; msMGE, multi-sgRNA-based multiplex genome editing coupled with the mixture of dsDNA_*As-Cath*, pUC57_*As-Cath* and pUC57_*Cec* donors targeting the *lh*, *mc4r*, *mstn1* and *mstn2* sites in parallel; TIDE, tracking of indels by decomposition (http://tide.nki.nl); *lh*, luteinizing hormone; *mc4r*, melanocortin-4 receptor; *mstn1*, the exon1 of myostatin; *mstn2*, the exon2 of myostatin.

3.5 Detection of mosaicism

To assess the potential mosaicism of transgenic channel catfish, transgenes were detected by examining 14 tissues from three representative transgenic individuals (#2, *Cec* transgenic; #4, *As-Cath* transgenic; #21, *As-Cath*/*Cec* transgenic). The PCR results showed that fish #21 had detectable *As-Cath* and *Cec* transgenes in all 14 tissues, including liver, kidney, spleen, blood, gill, skin, intestine, fin, barbel, muscle, stomach, brain, eye, and gonad. Additionally, the *As-*

Cath transgene can be detected in all 14 tissues of individual #4. However, in fish #2, the *Cec* transgene was present in 13 tissues but for the stomach, indicating a mosaic (Figure 27, Figure S19 in Appendix 2).

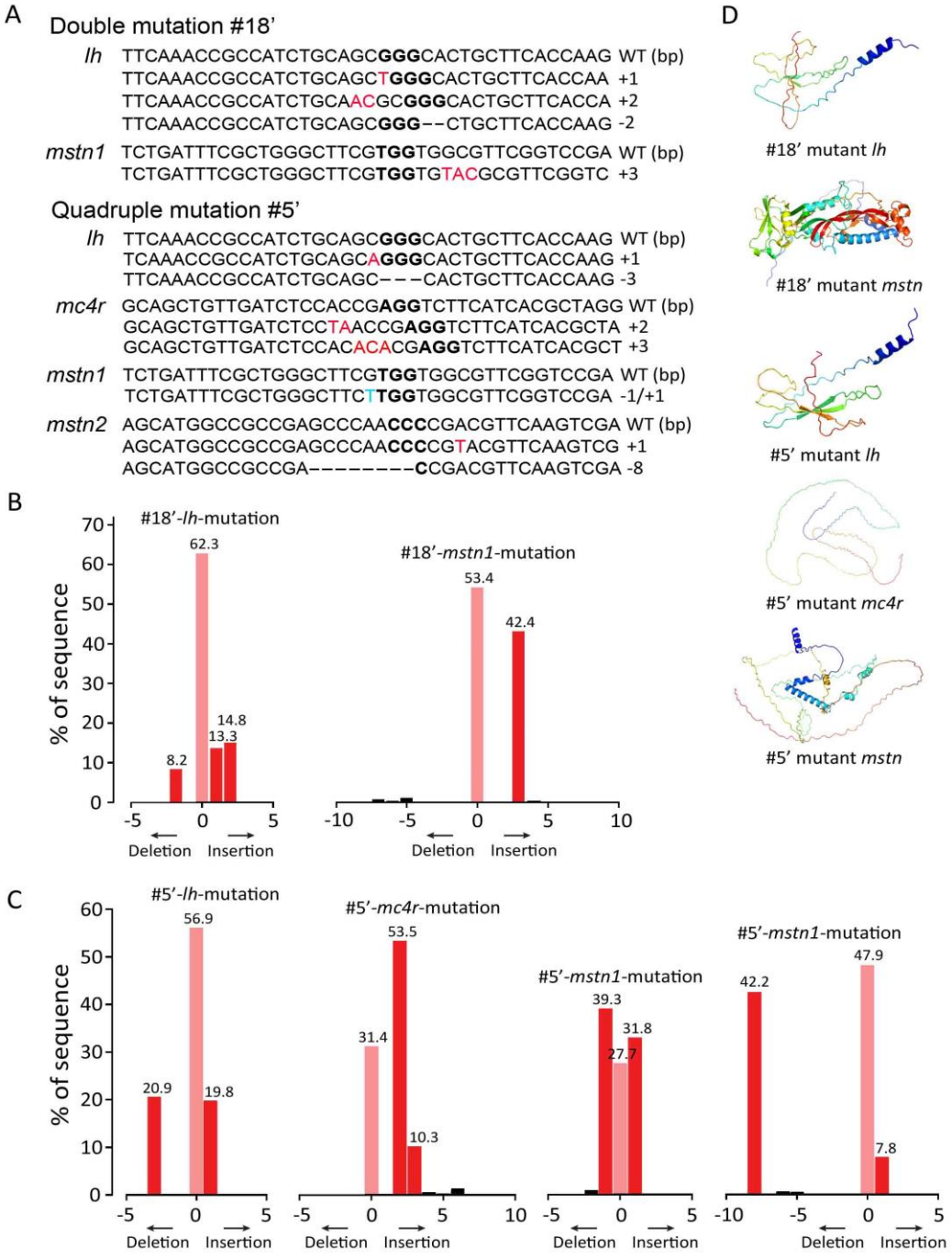


Figure 26. Mutated sequences and corresponding predicted 3D protein structures of the dual and quadruple genes in channel catfish (*Ictalurus punctatus*) using CRISPR/Cas9-mediated msMGE. (**A**) The sequences of dual (#18') and quadruple mutants (#5') were determined using Sanger sequencing. Insertions and substitutions are highlighted in red and blue, respectively. Deletions are indicated by

dashes. Protospacer adjacent motif is highlighted in bold. (**B**-**C**) Indel spectrum and the frequency determined by TIDE analysis in individual #18' and #5', respectively. (**D**) The 3D structures of mutant proteins were predicted using AlphaFold. WT, wild-type control; ssGE1, single sgRNA-based genome editing coupled with the pUC57_*As-Cath* donor targeting the *mc4r* locus; ssGE2, single sgRNA-based genome editing coupled with the pUC57_*Cec* donor targeting the *mstn1* site; msMGE, multi-sgRNA-based multiplex genome editing coupled with the mixture of dsDNA_*As-Cath*, pUC57_*As-Cath* and pUC57_*Cec* donors targeting the *lh*, *mc4r*, *mstn1* and *mstn2* sites in parallel; TIDE, tracking of indels by decomposition (http://tide.nki.nl); *lh*, luteinizing hormone; *mc4r*, melanocortin-4 receptor; *mstn1*, the exon1 of myostatin; *mstn2*, the exon2 of myostatin.

Similarly, mosaic individuals were presented in mutant fish. Compared to the ssGE strategy, msMGE induced more gene-deficient mosaic fish (26.67% [4/15] in ssGE1, 25.00% [2/8] in ssGE2, and 77.78% [21/27] in msMGE). The results of TIDE analysis in the ssGE1/ssGE2 strategy indicated that *mc4r* mutant individuals #32 (+1 bp, -2 bp), #38 (+1 bp, -2 bp), #42 (+4 bp, -2 bp), and #43 (+1 bp, -2 bp) had varied indels in the mixture of fin clips and barbels. Specifically, fish #32 and #38 had the same indels of type and sequence (+G, -TC). Although individual #43 had the same type of indels as #32 and #38, it was different in the sequence (+T, -CT). In addition, the same type and sequence of indels (+1 bp, -5 bp; +T, -TTCGT) in the *mstn1* locus were found in mosaic individuals #47 and #52 (Table 11). With respect to msMGE, in addition to the single-locus mutant mosaicism (fish #1' ~ #4', #6' ~ #8', #10', #12', #14', #16', #17' and #19'), double- and quadruple-loci mosaic mutants (#15' and #18'; #5' and #8') were also detected (Table 12, Appendix 8).

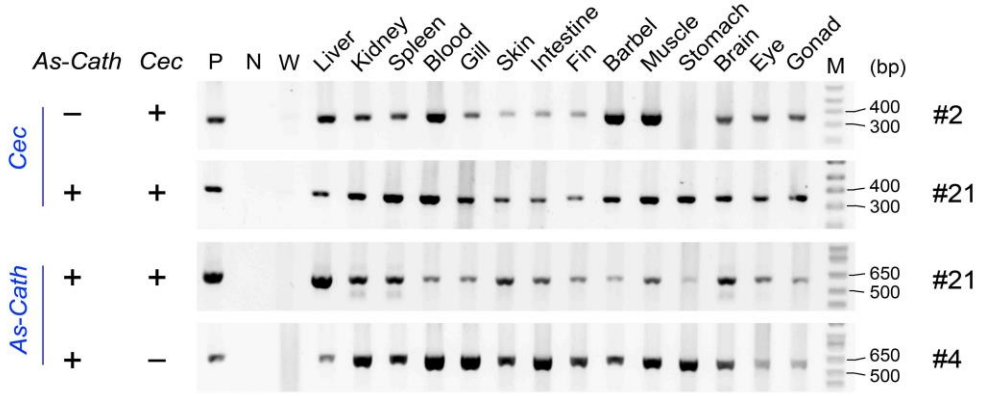


Figure 27. Determination of mosaicism in transgenic channel catfish (*Ictalurus punctatus*) by detecting the transgene in fourteen tissues from three individuals (#2, #4 and #21). The gel electrophoresis images shown here are cropped, and full-length gels are presented in Figure S19 in Appendix 2. *Cec*, the cecropin gene from moth; *As-Cath*, the cathelicidin gene from alligator; Lane N,

water as the blank control; Lane W, a wild-type individual as the negative control; Lane P, a plasmid as the positive control; lane M, 1kb+ DNA marker.

4. Discussion

To enable a precise knock-in strategy in fish species through CRISPR/Cas9-mediated multiplex genome editing, a multipronged system was implemented to introduce multiple antimicrobial peptide genes in channel catfish. A blend of donor constructs comprising linear dsDNA and double-cut plasmids was designed to concurrently generate transgenic/gene-edited fish lines, with the goal of replacing the reproduction-associated gene (*lh*) and growth-regulating genes (*mc4r* and *mstn*) of channel catfish were modified utilizing the CRISPR/Cas9 system. Subsequently, disease-resistance genetic lines carrying alligator cathelicidin and moth cecropin transgenes at three loci (two targets at the *mstn* locus) were established that have the potential for improved disease resistance with high growth rates and reversible sterility. This is the first report of two targeted gene insertions at three loci in an aquatic species, which holds promise for application to other organisms.

The msMGE system exhibited low toxicity towards channel catfish embryos, as reflected by comparable hatchability observed in injected control and msMGE groups. Additionally, the fry survival rates were identical between WT control and msMGE groups, further confirming the low toxicity. Although lower hatch rates were observed in the injection groups (ssGE and msMGE) compared to the WT group, there was no difference between the ssGE and msMGE groups. These findings suggested that CRISPR/Cas9 complex was minimally toxic to channel catfish embryos, even when multiple Cas9/sgRNA complexes and donor templates were co-injected into a single embryo. Hence, the heavy mortality observed in the embryos was primarily attributed to the physical damage caused by microinjection or pleiotropic effects resulting from gene disruption (Elaswad et al., 2018; Simora et al., 2020), rather than the inherent toxicity of the CRISPR/Cas9 complex itself. Schubert et al. (2014) found that zebrafish zygote mortality was positively correlated with the amount of reagent injected. Notably, fertilized zebrafish eggs are only 0.6 to 0.8 mm in size. The volume of solution for microinjection is typically 10% to 20% of the egg volume, which increases the mortality of the injected eggs to some extent (Xu, 1999; Rosen et al., 2009). However, in channel catfish, approximately 3-mm embryos were injected with 25 nL (8 μL for 320 eggs) of the solution, which was only 0.02% of the volume, greatly

reducing the contribution of injected volume to mortality in our work. Additionally, gametes of varying quality due to parental effects also affect hatchability and fry survival. In the present study, the hatchability and number of surviving fry of family 2 were significantly lower than those of other families, which was likely related to gamete quality.

The notable result of the current project was the precise integration of multiple transgenes at multiple sites, inducing double, triple, and quadruple transgenic fish using CRISPR/Cas9-mediated msMGE. Our data showed that 6.45% and 9.67% of the fish carried triple or quadruple transgenes, respectively, and this efficiency was comparable to results of Xing et al. (2022a) for which a 7% integration rate of channel catfish carrying three transgenes was achieved using a cocktail design involving the simultaneous introduction of the three plasmids. Compared to targeting only one insertion locus, our study demonstrated an increased transgene efficiency of 1.5- to 2.1-X by targeting multiple loci coupled with the same transgene donor using the msMGE strategy. However, the same transgene had varied integration rates at different loci. For example, the KI efficiency of the *As-Cath* transgene at the *mc4r* locus was 33.89%, which was significantly higher than that at the *lh* locus (6.78%). The difference in KI rates is mainly caused by the type of donor templates (Zhang et al., 2017). In this study, we used two types of donor templates in one-step microinjection, while different types of donor templates contributed to various KI efficiencies (Yoshimi et al., 2016; Zhang et al., 2017; Boel et al., 2018). A dsDNA containing the *As-Cath* coding sequences was used to replace the *lh* gene, while a double-cut plasmid was designed to replace the *mc4r* gene. Previous studies have reported that the CRISPR/Cas9-mediated system assisted by a double-cut plasmid can increase the KI efficiency up to 26.7% in medaka (*Oryzias latipes*) (Murakami et al., 2017). More recently, Simora et al. (2020) demonstrated that plasmid donors induced higher integration rates than linear dsDNA donors in channel catfish. These findings indicated that the transgene had a higher integration rate when a double-cut plasmid was provided rather than a linear dsDNA.

Mutagenesis was achieved as demonstrated by DNA sequencing and TIDE analysis in both ssGE and msMGE systems targeting single or multiple loci. Interestingly, although the same sgRNAs and constructs were used in ssGE1 and msMGE to target the *mc4r* locus, the TIDE results manifested different indel spectra with varied gene-edited efficiencies. Specifically, *mc4r*_sgRNA targeting the *mc4r* gene produced mainly small insertions in the ssGE1 strategy,

whereas insertions were major in the msMGE. Overall, small indels appeared to be the most common mutations induced by Cas9 combined with one sgRNA, which is in agreement with other studies (Yang et al., 2013; Canver et al., 2014; Kim et al., 2019; Wang et al., 2023b). Furthermore, even in the msMGE system, our data suggested that not all genomic loci were equally accessible to mutagenization. Similar observations have been made in previous studies (Jao et al., 2013; Hwang et al., 2013). Compared to the *lh* locus, the *mc4r* and *mstn* genes had higher mutagenesis rates, suggesting that the *mc4r* and *mstn* genes are more amenable to gene editing. Previous studies have also shown that different editing systems have selectivity for KI or mutation sites, and even the same system produces various types and efficiencies of indels (Sakuma et al., 2014; Xing et al., 2022a). These findings indicate that the efficacy and mutation spectra vary greatly depending on the sgRNA target site and CRISPR system used (Hsu et al., 2013; Fu et al., 2014; Brinkman et al., 2014). Future experiments are needed to elucidate how the DSB repair machinery selects bases for insertion.

Despite the promising applications of multiplex genome editing, there are several obstacles that need to be addressed. One of the biggest challenges is off-target effects. While multiplex genome editing creates complex genetic modifications, it also increases the risk of unintended genetic modifications at sites other than the intended targets (McCarty et al., 2020; Abdelrahman et al., 2021). Compared to the CRISPR/Cas9-mediated ssGE system, msMGE induced a higher rate of off-target insertions in our current study. On the one hand, the efficiency of on-target KI via HDR is hindered by the competing NHEJ pathway (Maruyama et al., 2015; Chu et al., 2015), while the use of multiple sgRNAs increases the risk of off-target effects (Fortin et al., 2019; McCarty et al., 2020). On the other hand, Xing et al. (2022a) demonstrated that some on-target inserted sequences at the 5’ and 3’ joints did not fully match theoretical expectations due to the complex of multiple constructs, indicating that using PCR only to detect the junction regions may underestimate the on-target efficiency. Recently, some engineered Cas9 variants (St1Cas9 and SaCas9) with altered PAM specificities (Kleinstiver et al., 2015), as well as the enhanced specificity Cas9 (eSpCas9) (Slaymaker et al., 2016) and the high-fidelity SpCas9-HF1 (Kleinstiver et al., 2016) have been developed to reduce off-target by making the Cas9/sgRNA complex less tolerant to mismatches. Another possible strategy for off-target reduction is the use of a 2A peptide that links two or more coding sequences to independently drive the expression of multiple transgenes. This reduces off-target events by decreasing the number of target loci (Tan

et al., 2010; Liu et al., 2017). Continued research endeavors to discover or engineer high-fidelity nuclease variants, refine construct/sgRNA design, and establish exceptionally sensitive detection techniques, will ultimately render these pioneering CRISPR/Cas9 systems viable for implementation in non-model animals.

Another challenge is the potential for somatic mosaicism. Induced mosaicism from CRISPR/Cas9-mediated gene mutation/transgenesis is common due to the continuous targeting and cleavage of genes during various embryonic developmental stages (Jao et al., 2013; Mizuno et al., 2014; Oliver et al., 2015; Mehravar et al., 2019). Different levels of mosaicism have been observed in zebrafish (Blitz et al., 2013), channel catfish (Simora et al., 2020; Xing et al., 2022b), blue catfish (Wang et al., 2023b) and other fish species (Blix et al., 2021). In this study, mosaicism was observed by both donor-dependent NHEJ repair and HDR-directed insertion, which is consistent with our previous study (Wang et al., 2023b) and other studies (Jao et al., 2013; Auer et al., 2014; Zuo et al., 2017). The frequency of mosaicism varies between species and CRISPR/Cas9 systems, and there is currently no clear and effective strategy to completely eliminate mosaicism. To overcome mosaicism in non-model animals, the initial step is to generate a founder animal harboring the desired modifications. Subsequently, the generation of new mutant strains is accomplished through outcrossing these mosaic founders (Mehravar et al., 2019). Notwithstanding these constraints, the findings articulated in this work furnish a robust foundation for future inquiries and engender new avenues of exploration.

Not all sequence mutations resulted in non-functional or structurally altered proteins based on AlphaFold prediction. Some mutations have minimal or no impact on protein structure or function (Schaefer and Rost, 2012). Based on our findings, the *mstn1* mutants had a lower probability of acquiring a loss-of-function myostatin protein compared to the *mstn2* mutants. For example, the loss of 3 bases (in-frame deletions) at the *mstn1* locus in fish #48 and #50 resulted in the deletion of an alanine (A) at position 15, but no subsequent AA changes. Thus, this 3-bp deletion did not perturb the structure of myostatin (Appendix 8). Mutations at the *mstn1* locus are unlikely to result in structural changes in the protein, possibly because the protein encoded by exon1 is distant from the myostatin growth factor domain (Zhang et al., 2020). In this case, unless it is a nonsense mutation or a large number of AA sequence substitutions, it is less likely to affect the protein's structure (Jumper et al., 2021). In contrast, this work observed a discernible

loss of function when both exons of the *mstn* gene (*mstn1* and *mstn2*) were altered. This efficient mutation was demonstrated in a previous report by Ou et al. (2023), where they detected a higher mutation rate by simultaneously targeting exon1 and exon2 with two sgRNAs in blotched snakehead. However, silent mutations, for example, a previous study showed that the substitution of a single base did not change the AA sequence in the protein due to the redundancy of the genetic code in blue catfish (Wang et al., 2023b). In the present study, fish #34 had a 2-bp substitution at the cleavage site of the *mc4r* gene, while this modification did not alter the AA sequence due to codon degeneracy. Additionally, Jao et al. (2013) reported that *mitfa*-null zebrafish did not show any phenotypic change in pigmentation. The lack of phenotypic change is likely due to the mutated sequences still encoding a functional protein. In view of this, after the base sequence mutation is realized, follow-up work including the analysis of AA sequence and protein structure, and the verification of protein function is essential. Naturally, the acquisition of phenotype depends on the integrity of protein functional structure.

Previous studies have established *mstn*-, *lh*-, and *mc4r*-deficient or *As-Cath* and *Cec* transgenic channel catfish (Dunham et al., 2002; Qin et al., 2016; Khalil et al., 2017; Simora et al., 2020; Coogan et al., 2022a, 2022b; Wang et al., 2023a). In addition, previous work has demonstrated growth advantages in *mc4r*/*mstn*-deficient fish, including channel catfish (Khalil et al., 2017; Coogan et al., 2022a, 2022b), common carp (Zhong et al., 2016; Shahi et al., 2022), olive flounder (*Paralichthys olivaceus*) (Kim et al., 2019), and red sea bream (*Pagrus major*) (Kishimoto et al., 2018). Meanwhile, *lh*-mutant channel catfish with reversible sterility, and cecropin/cathelicidin transgenic channel catfish with enhanced resistance to bacteria were established (Dunham et al., 2002; Qin et al., 2016; Qin, 2019; Wang et al., 2023a). In this study, we first achieved the integration of dual antimicrobial peptide genes at three loci (reproduction- and growth regulation-related) by one-step microinjection. That is, the *lh*/*mc4r* was replaced with the cathelicidin transgene and the *mstn* with the cecropin transgene. We hypothesize that this new transgenic line has the combination of these breeder-preferred traits based on the supported publications. However, multiple traits including disease resistance, growth rate and reproductive capacity should be evaluated in the future.

In conclusion, we succeeded in knocking in two transgenes at three loci and four targets in the genome using the CRISPR/Cas9-mediated msMGE system, coupled with a mixture of dsDNA

and double-cut plasmids, which is perfectly applicable in channel catfish. It is very promising to replace functional alleles with transgenes at multiple loci in parallel. The findings of this study contribute to the existing body of knowledge on the application of genome editing in aquaculture and provide valuable insights into producing multiple consumer-valued qualities. Given its simplicity and high efficiency, we propose that the described method could become a standard technique for the generation of transgenic and mutant fish.

5. References

Abdelrahman M, Wei Z, Rohila JS, Zhao K. Multiplex genome-editing technologies for revolutionizing plant biology and crop improvement. *Front Plant Sci*. 2021;12:721203. doi:10.3389/fpls.2021.721203

Auer TO, Duroure K, Cian AD, Concordet JP, Bene FD. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res*. 2014;24:142–53. doi:10.1101/gr.161638.113

Bae S, Park J, Kim J-S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*. 2014;30:1473–75. doi:10.1093/bioinformatics/btu048

Bart AN, Dunham RA. Effects of sperm concentration and egg number on fertilization efficiency with channel catfish (*Ictalurus punctatus*) eggs and blue catfish (*I. furcatus*) spermatozoa. *Theriogenology*. 1996;45:673–82. doi:10.1016/0093-691X(95)00413-3

Blitz IL, Biesinger J, Xie X, Cho KW. Biallelic genome modification in F0 *Xenopus tropicalis* embryos using the CRISPR/Cas system. *Genesis*. 2013;51:827–34. doi:10.1002/dvg.22719

Blix TB, Dalmo RA, Wargelius A, Myhr AI. Genome editing on finfish: Current status and implications for sustainability. *Rev Aquac*. 2021;13:2344–63. doi:10.1111/raq.12571

Boel A, Saffel HD, Steyaert W, Callewaert B, Paepe AD, Coucke PJ, et al. CRISPR/Cas9-mediated homology-directed repair by ssODNs in zebrafish induces complex mutational patterns resulting from genomic integration of repair-template fragments. *Dis Model Mech*. 2018;18:dmm035352. doi:10.1242/dmm.035352

Boman HG, Faye I, Hofsten P, Kockum K, Lee J-Y, Xanthopoulos KG, et al. On the primary structures of lysozyme, cecropins and attacins from *Hyalophora cecropia*. *Dev Comp Immunol*. 1985;9:551–8. doi:10.1016/0145-305X(85)90018-7

Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res*. 2014;42:e168. doi:10.1093/nar/gku936

Campa CC, Weisbach NR, Santinha AJ, Incarnato D, Platt RJ. Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nat Methods*. 2019;16:887–93. doi:10.1038/s41592-019-0508-6

Canver MC, Bauer DE, Dass A, Yien YY, Chung J, Masuda T, et al. Characterization of genomic deletion efficiency mediated by CRISPR/Cas9 in mammalian cells. *J Biol Chem*. 2014;289:21312–24. doi:10.1074/jbc.M114.564625

Chen Y, Cai S, Qiao X, Wu M, Guo Z, Wang R, et al. As-CATH1-6, novel cathelicidins with potent antimicrobial and immunomodulatory properties from Alligator sinensis, play pivotal roles in host antimicrobial immune responses. *Biochem J*. 2017;474:2861–85. doi:10.1042/BCJ20170334

Chu L, Li J, Liu Y, Hu W, Cheng CHK. Targeted gene disruption in zebrafish reveals noncanonical functions of LH signaling in reproduction. *Mol Endocrinol*. 2014;28:1785–95. doi:10.1210/me.2014-1061

Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotechnol*. 2015;33:543−48. doi:10.1038/nbt.3198

Cleveland BM, Yamaguchi G, Radler LM, Shimizu M. Editing the duplicated insulin-like growth factor binding protein-2b gene in rainbow trout (*Oncorhynchus mykiss*). *Sci Rep*. 2018;8:16054. doi:10.1038/s41598-018-34326-6

Cone RD. Studies on the physiological functions of the melanocortin system. *Endocr Rev*. 2006;27:736–49. doi:10.1210/er.2006-0034

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339:819–23. doi:10.1126/science.1231143

Coogan M, Alston V, Su B, Khalil K, Elaswad A, Khan M, et al. CRISPR/Cas-9 induced knockout of myostatin gene improves growth and disease resistance in channel catfish (*Ictalurus punctatus*). *Aquaculture*. 2022a;557:738290. doi:10.1016/j.aquaculture.2022.738290

Coogan M, Alston V, Su B, Khalil K, Elaswad A, Khan M, et al. Improved growth and high inheritance of melanocortin-4 receptor (*mc4r*) mutation in CRISPR/Cas-9 gene-edited channel catfish, *Ictalurus punctatus*. *Mar Biotechnol*. 2022b;24:843–55. doi:10.1007/s10126-022-10146-8

Cui X, Ji D, Fisher DA, Wu Y, Briner DM, Weinstein EJ. Targeted integration in rat and mouse embryos with zinc-finger nucleases. *Nat Biotechnol*. 2011;29:64–7. doi:10.1038/nbt.1731

Dunham RA, Warr GW, Nichols A, Duncan PL, Argue B, Middleton D, et al. Enhanced bacterial disease resistance of transgenic channel catfish *Ictalurus punctatus* possessing cecropin genes. *Mar Biotechnol*. 2002;4:338–44. doi:10.1007/s10126-002-0024-y

Edvardsen RB, Leininger S, Kleppe L, Skaftnesmo KO, Wargelius A. Targeted mutagenesis in Atlantic salmon (*Salmo salar* L.) using the CRISPR/Cas9 system induces complete knockout individuals in the F0 generation. *PLoS One*. 2014;9:e108622. doi:10.1371/journal.pone.0108622

Elaswad A, Khalil K, Ye Z, Liu Z, Liu S, Peatman E, et al. Effects of CRISPR/Cas9 dosage on TICAM1 and RBL gene mutation rate, embryonic development, hatchability and fry survival in channel catfish. *Sci Rep*. 2018;8:16499. doi:10.1038/s41598-018-34738-4

FAO. The State of World Fisheries and Aquaculture 2020. Food and Agriculture Organization; 2020. doi:10.4060/ca9229en.

Fortin JP, Tan J, Gascoigne KE, Haverty PM, Forrest WF, Costa MR, et al. Multiple-gene targeting and mismatch tolerance can confound analysis of genome-wide pooled CRISPR screens. *Genome Biol*. 2019;20:21. doi:10.1186/s13059-019-1621-7

Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol*. 2014;32:279–84. doi:10.1038/nbt.2808

Gen K, Yamaguchi S, Okuzawa K, Kumakura N, Tanaka H, Kagawa H. Physiological roles of FSH and LH in red seabream, *Pagrus major*. *Fish Physiol and Biochem*. 2003;28:77–80. doi:10.1023/B:FISH.0000030480.97947.ba

Gratacap RL, Wargelius A, Edvardsen RB, Houston RD. Potential of genome editing to improve aquaculture breeding and production. *Trends Genet*. 2019;35:672–84. doi:10.1016/j.tig.2019.06.006

Houston RD, Bean TP, Macqueen DJ, Gundappa MK, Jin YH, Jenkins TL, et al. Harnessing genomics to fast-track genetic improvement in aquaculture. *Nat Rev Genet*. 2020;21:389–409. doi:10.1038/s41576-020-0227-y

Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol*. 2013;31:227–29. doi:10.1038/nbt.2501

Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, AgarwalaV, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. 2013;31:827–32. doi:10.1038/nbt.2647

Jao LE, Wente SR, Chen W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc Natl Acad* Sci USA. 2013;110:13904–09. doi:10.1073/pnas.1308335110

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337:816–21. doi:10.1126/science.1225829

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596:583–9. doi:10.1038/s41586-021-03819-2

Khalil K, Elayat M, Khalifa E, Daghash S, Elaswad A, Miller M, et al. Generation of myostatin gene-edited channel catfish (*Ictalurus punctatus*) via zygote injection of CRISPR/Cas9 system. *Sci Rep*. 2017;7:7301. doi:10.1038/s41598-017-07223-7

Kim J, Cho J, Kim J, Kim HC, Noh JK, Kim YO, et al. CRISPR/Cas9-mediated myostatin disruption enhances muscle mass in the olive flounder *Paralichthys olivaceus*. *Aquaculture*. 2019;512:734336. doi:10.1016/j.aquaculture.2019.734336

Kishimoto K, Washio Y, Yoshiura Y, Toyoda A, Ueno T, Fukuyama H, et al. Production of a breed of red sea bream *Pagrus major* with an increase of skeletal muscle mass and reduced body

length by genome editing with CRISPR/Cas9. *Aquaculture*. 2018;495:415–27. doi:10.1016/j.aquaculture.2018.05.055

Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*. 2015;523:481–85. doi:10.1038/nature14592

Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*. 2016;529:490–95. doi:10.1038/nature16526

Kroll F, Powell GT, Ghosh M, Gestri G, Antinucci P, Hearn TJ, et al. A simple and effective F0 knockout method for rapid screening of behaviour and other complex phenotypes. *eLife*. 2021;10:e59683. doi:10.7554/eLife.59683

Krug J, Perner B, Albertz C, Mörl H, Hopfenmüller VL, Englert C. Generation of a transparent killifish line through multiplex CRISPR/Cas9-mediated gene inactivation. *eLife*. 2023;12:e81549. doi:10.7554/eLife.81549

Labun K, Montague TG, Krause M, Torres Cleuren YN, Tjeldnes H, Valen E. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res*. 2019;47:W171–4. doi:10.1093/nar/gkz365

Liu Z, Chen O, Wall JBJ, Zheng M, Zhou Y, Wang L, et al. Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci Rep*. 2017;7:2193. doi:10.1038/s41598-017-02460-2

Liu Z, Liu S, Yao J, Bao L, Zhang J, Li Y, et al. The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts. *Nat Commun*. 2016;7:11757. doi:10.1038/ncomms11757

Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol*. 2015;33:538−42. doi:10.1038/nbt.3190

McCarty NS, Graham AE, Studená L, Ledesma-Amaro R. Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nat Commun*. 2020;11:1281. doi:10.1038/s41467-020-15053-x

Mehravar M, Shirazi A, Nazari M, Banan M. Mosaicism in CRISPR/Cas9-mediated genome editing. *Dev Biol*. 2019;445:56–62. doi:10.1016/j.ydbio.2018.10.008

Meyer M, de Angelis MH, Wurst W, Kühn R. Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. *Proc Natl Acad Sci* USA. 2010;107:15022–6. doi:10.1073/pnas.1009424107

Mizuno S, Dinh TTH, Kato K, Mizuno-Iijima S, Tanimoto Y, Daitoku Y, et al. Simple generation of albino C57BL/6J mice with G291T mutation in the tyrosinase gene by the CRISPR/Cas9 system. *Mamm Genome*. 2014;25:327–34. doi:10.1007/s00335-014-9524-0

Mosimann C, Kaufman CK, Li P, Pugach EK, Tamplin OJ, Zon LI. Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish. *Development*. 2011;138:169–77. doi:10.1242/dev.059345

Murakami Y, Ansai S, Yonemura A, Kinoshita M. An efficient system for homology-dependent targeted gene integration in medaka (*Oryzias latipes*). *Zoological Lett*. 2017;3:10. doi:10.1186/s40851-017-0071-x

Ohama M, Washio Y, Kishimoto K, Kinoshita M, Kato K. Growth performance of myostatin knockout red sea bream *Pagrus major* juveniles produced by genome editing with CRISPR/Cas9. *Aquaculture*. 2020;529:735672. doi:10.1016/j.aquaculture.2020.735672

Oliver D, Yuan S, Mcswiggin H, Yan W. Pervasive genotypic mosaicism in founder mice derived from genome editing through pronuclear injection. *PLoS One*. 2015;10:e0129457. doi:10.1371/journal.pone.0129457

Ou M, Wang F, Li K, Wu Y, Huang S, Luo Q, et al. Generation of myostatin gene-edited blotched snakehead (*Channa maculata*) using CRISPR/Cas9 system. *Aquaculture*. 2023;563:738988. doi:10.1016/j.aquaculture.2022.738988

Qin Z, Li Y, Su B, Cheng Q, Ye Z, Perera DA, et al. Editing of the luteinizing hormone gene to sterilize channel catfish, *Ictalurus punctatus*, using a modified zinc finger nuclease technology with electroporation. *Mar Biotechnol*. 2016;18:255–63. doi:10.1007/s10126-016-9687-7

Qin, G., 2019. Gene Editing and Hormone Therapy to Control Reproduction in Channel Catfish, *Ictalurus punctatus*. Doctoral dissertation. Auburn University, AL, USA. http://hdl.handle.net/10415/6690

Rosen JN, Sweeney MF, Mably JD. Microinjection of zebrafish embryos to analyze gene function. *J Vis Exp*. 2009;e1115. doi:10.3791/1115

Sakuma T, Nishikawa A, Kume S, Chayama K, Yamamoto T. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci Rep*. 2014;4:5400. doi:10.1038/srep05400

Schaefer C, Rost B. Predict impact of single amino acid change upon protein structure. *BMC Genomics*. 2012;13:S4. doi:10.1186/1471-2164-13-S4-S4

Schubert S, Keddig N, Hanel R, Kammann U. Microinjection into zebrafish embryos (*Danio rerio*)-a useful tool in aquatic toxicity testing? *Environ Sci Eur*. 2014;26:1. doi:10.1186/s12302-014-0032-3

Shahi N, Mallik SK, Sarma D. Muscle growth in targeted knockout common carp (*Cyprinus carpio*) *mstn* gene with low off-target effects. *Aquaculture*. 2022;547:737423. doi:10.1016/j.aquaculture.2021.737423

Simora RMC, Xing D, Bangs MR, Wang W, Ma X, Su B, et al. CRISPR/Cas9-mediated knock-in of alligator cathelicidin gene in a non-coding region of channel catfish genome. *Sci Rep*. 2020;10:22271. doi:10.1038/s41598-020-79409-5

Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang, F. Rationally engineered Cas9 nucleases with improved specificity. *Science*. 2016;351:84–8. doi:10.1126/science.aad5227

Sun Y, Zhang J, Xiang J. A CRISPR/Cas9-mediated mutation in chitinase changes immune response to bacteria in *Exopalaemon carinicauda*. *Fish Shellfish Immunol*. 2017;71:43–9. doi:10.1016/j.fsi.2017.09.065

Tan Y, Liang H, Chen A, Guo X. Coexpression of double or triple copies of the rabies virus glycoprotein gene using a 'self-cleaving' 2a peptide-based replication-defective human adenovirus serotype 5 vector. *Biologicals*. 2010;38:586–3. doi:10.1016/j.biologicals.2010.06.001

Thomas M, Langley B, Berry C, Sharma M, Kirk S, Bass J, et al. Myostatin, a negative regulator of muscle growth, functions by inhibiting myoblast proliferation. *J Biol Chem*. 2000;275:40235–43. doi:10.1074/jbc.M004356200

Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*. 2013;153:910–8. doi:10.1016/j.cell.2013.04.025

Wang J, Cheng Y. Enhancing aquaculture disease resistance: Antimicrobial peptides and gene editing. 2023. doi:10.1111/RAQ.12845

Wang J, Su B, Bruce TJ, Wise AL, Zeng P, Cao G, et al. CRISPR/Cas9 microinjection of transgenic embryos enhances the dual-gene integration efficiency of antimicrobial peptide genes for bacterial resistance in channel catfish, *Ictalurus punctatus*. *Aquaculture*. 2023a:739725. doi:10.1016/j.aquaculture.2023.739725

Wang J, Su B, Al-Armanazi J, Wise AL, Shang M, Bern L, et al. Integration of alligator cathelicidin gene via two CRISPR/Cas9-assisted systems enhances bacterial resistance in blue catfish, *Ictalurus furcatus*. *Aquaculture*. 2023b.

Wang J, Su B, Dunham A. Genome-wide identification of catfish antimicrobial peptides: A new perspective to enhance fish disease resistance. *Rev Aquac*. 2022;4:2002–2022. doi:10.1111/raq.12684

Wang X, Zheng W, Zhou H, Tu Q, Tang YJ, Stewart AF, et al. Improved dsDNA recombineering enables versatile multiplex genome engineering of kilobase-scale sequences in diverse bacteria. *Nucleic Acids Res*. 2021;50:e15. doi:10.1093/nar/gkab1076

Xing D, Li S, Shang M, Wang W, Zhang Q, Wang J, et al. A new strategy for increasing knock-in efficiency: Multiple elongase and desaturase transgenes knock-in by targeting long repeated sequences. *ACS Synth Biol*. 2022a;11:4210–19. doi:10.1021/acssynbio.2c00252

Xing D, Su B, Li S, Bangs M, Creamer D, Coogan M, et al. CRISPR/Cas9-mediated transgenesis of the Masu salmon (*Oncorhynchus masou*) *elovl2* gene improves n-3 fatty acid content in channel catfish (*Ictalurus punctatus*). *Mar Biotechnol*. 2022b;24:513–23. doi:10.1007/s10126-022-10110-6

Xing D, Shang M, Li S, Wang W, Hasin T, Hettiarachchi D, et al. CRISPR/Cas9-mediated precision integration of *fat-1* and *fat-2* from *Caenorhabditis elegans* at long repeated sequence in channel catfish (*Ictalurus punctatus*) and the impact on n-3 fatty acid level. *Aquaculture*. 2023;567:739229. doi:10.1016/j.aquaculture.2023.739229

Xu Q. Microinjection into zebrafish embryos. Molecular Methods in Developmental Biology: Xenopus and Zebrafish. Methods in Molecular Biology. 1999;125–32.

Yang L, Guell M, Byrne S, Yang JL, De Los Angeles A, Mali P, et al. Optimization of scarless human stem cell genome editing. *Nucleic Acids Res*. 2013;41:9049–61. doi:10.1093/nar/gkt555

Yoshimi K, Kunihiro Y, Kaneko T, Nagahora H, Voigt B, Mashimo T. ssODN-mediated knock-in with CRISPR-Cas for large genomic regions in zygotes. *Nat Commun*. 2016;7:10431. doi:10.1038/ncomms10431

Yu H, Li H, Li Q, Xu R, Yue C, Du S. Targeted gene disruption in Pacific oyster based on CRISPR/Cas9 ribonucleoprotein complexes. *Mar Biotechnol*. 2019;21:301–9. doi:10.1007/s10126-019-09885-y

Zhang JP, Li XL, Li GU, Chen W, Arakaki C, Botimer GD, et al. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol*. 2017;18:35. doi:10.1186/s13059-017-1164-8

Zhang S, Li Y, Shao J, Liu H, Wang J, Wang M, et al. Functional identification and characterization of *IpMSTNa*, a novel orthologous myostatin (*MSTN*) gene in channel catfish *Ictalurus punctatus*. *Int J Biol Macromol*. 2020;152:1–10. doi:10.1016/j.ijbiomac.2020.02.060

Zhong Z, Niu P, Wang M, Huang G, Xu S, Sun Y, et al. Targeted disruption of *sp7* and *myostatin* with CRISPR-Cas9 results in severe bone defects and more muscular cells in common carp. *Sci Rep*. 2016;6:22953. doi:10.1038/srep22953

Zuo E, Cai YJ, Li K, Wei Y, Wang BA, Sun Y, et al. One-step generation of complete gene knockout mice and monkeys by CRISPR/Cas9-mediated gene editing with multiple sgRNAs. *Cell Res*. 2017;27:933. doi:10.1038/cr.2017.81

Future Perspectives

Genome editing, dominated by CRISPR-based platforms, has risen rapidly in the life sciences over the past decade. In aquaculture, the CRISPR/Cas9-mediated system holds promise for the enhancement of favorable traits, especially growth, disease resistance, sterility, and fatty-acid profile. In addition to being affordable and effective, CRISPR/Cas9 has the property of being widely applicable because it allows for simultaneous modifications to various genomic sites by delivering multiple sgRNAs with the Cas9 protein/mRNA (Yang et al., 2013; Ota et al., 2014). As previously established, AMGs can successfully improve fish disease resistance through transgenic integration. However, documental studies primarily concentrated on single-AMG integration. Theoretically, it is conceivable to introduce two or more AMGs into the genome to acquire higher heritable resistance to diseases by using CRISPR/Cas9-mediated transgenesis.

In addition to integrating exogenous AMGs into the genome, experiments have demonstrated that knockout of immune-related genes can enhance resistance against pathogens by negatively regulating gene expression or disrupting pathways. For instance, a few representative genes, like rhamnose-binding lectin (*RBL*), signal transducer and activator of transcription 2 (*STAT2*), junctional adhesion molecule-A (*JAM-A*), the repressor of RNA polymerase III transcription of *Paralichthys olivaceus* (*PoMaf1*), and GRB2-associated binding protein 3 (*GAB3*) have undergone mutations that mimicked and altered the immunity of the fish and improved the host's resistance to disease (Wang et al., 2022; Yang et al., 2022). Indeed, *RBL* is a critical component of fish's innate immunity as an antibacterial and non-self-recognition molecule (Booy et al., 2005; Watanabe et al., 2009), especially in the protection of teleost eggs as well as in the mucosa (Beck et al., 2012). Interestingly, a potential negative regulation of *RBL* is involved in the immunity of some fishes against pathogenic invasion. Beck et al. (2012) confirmed that columnaris susceptibility was negatively linked with *RBL* expression levels. Furthermore, vulnerable fish's gills showed higher up-regulated levels of *RBL* than those of resistant fish (Peatman et al., 2013). Recently, an RBL-mutated channel catfish line was established (Elaswad et al., 2018), and a higher survival rate was represented in their F_2 individuals compared to those WT fish after being infected with *Flavobacterium columnare*. In this regard, integrating AMGs into these susceptibility loci can bidirectionally boost disease resistance based on gene pleiotropy.

Alternatively, some studies have proved that MSTN-deficient fish not only grow faster, but also reduced disease susceptibility to *Edwardsiella ictaluri* in channel catfish (Coogan et al., 2022). Therefore, *MSTN* is also an alternative locus of an AMG integration for enhanced disease resistance.

Beyond the successful achievement of one desired trait, it is possible to use CRISPR/Cas9 to simultaneously improve multiple characteristics based on empirical data and theoretical foundations (Alimuddin et al., 2008; Qin et al., 2016; Qin et al., 2022; Xing et al., 2022), which means that we could alter other traits through the construction of different vectors while we focus on enhancement of disease resistance. Documentary investigations have already shown that single transgene or multiple transgenes can be integrated into the fish genome by single sgRNA-mediated genome editing (ssGE) or multi-sgRNA-mediated multiple genome editing (msMGE) assisted using plasmid donors (Xing et al., 2022, 2023). Wang et al. (unpublished data) demonstrated that the cathelicidin transgene can be knocked in the LH locus of channel catfish by double-stranded DNA (dsDNA)-mediated ssGE. In addition, a double-cut plasmid containing a foreign AMG transgene as a donor is beneficial to improve integration rates using CRISPR/Cas9-mediated ssGE (Xing et al., 2022; Wang et al., 2023). In general, several points need to be considered in the implementation of AMG transgenesis: 1) The AMG candidates should first ensure that the encoded AMPs are non-toxic or less toxic to fish cells, which should be confirmed by *in vitro* experiments (Dunham et al., 2002; Sarmasik et al., 2002; Hsieh et al., 2010). 2) An appropriate cassette containing the transgene, promoter, and termination sequences should be constructed to ensure that the transgene is expressed after integration (Wang et al., 2022). 3) Selected sgRNA should be subjected to online tools to minimize the off-target events, and the cleavage efficiency of synthetic sgRNA should be examined *in vitro* (Hallerman et al., 2022). 4) Microinjection or electroporation of the CRISPR-Cas9/sgRNA complex should be performed at the one-cell stage to reduce mosaic effects (Mehravar et al., 2019).

From a genetic perspective, our hypothesis is that replacing the original functional genes with AMGs in specific coding regions of the chromosome would confer multi-generational antimicrobial activities of the host and improve multiple producer-favor traits. This strategy will hopefully allow us to create new fish lines possessing multiple traits, such as sterile and disease-

resistant, growth-boosted, and disease-resistant, or docosahexaenoic acid (DHA)-enriched and disease-resistant, or hybrid lines that have all these traits (Figure 28). In this vein, an example from our work demonstrates that it is highly feasible to insert the cathelicidin gene at the luteinizing hormone (*lh*) locus and cecropin gene at *MSTN* locus using a one-step CRISPR/Cas9-mediated system, resulting in gene-edited fish with increased disease resistance and growth, but decreased fecundity (unpublished data). However, site-directed knock-in of multi-locus genes tends to increase mosaicism and off-target effects (Hsu et al., 2013; Yang et al., 2013). In this scenario, the implementation and dissemination of gene-edited fish to industry needs to be well planned.

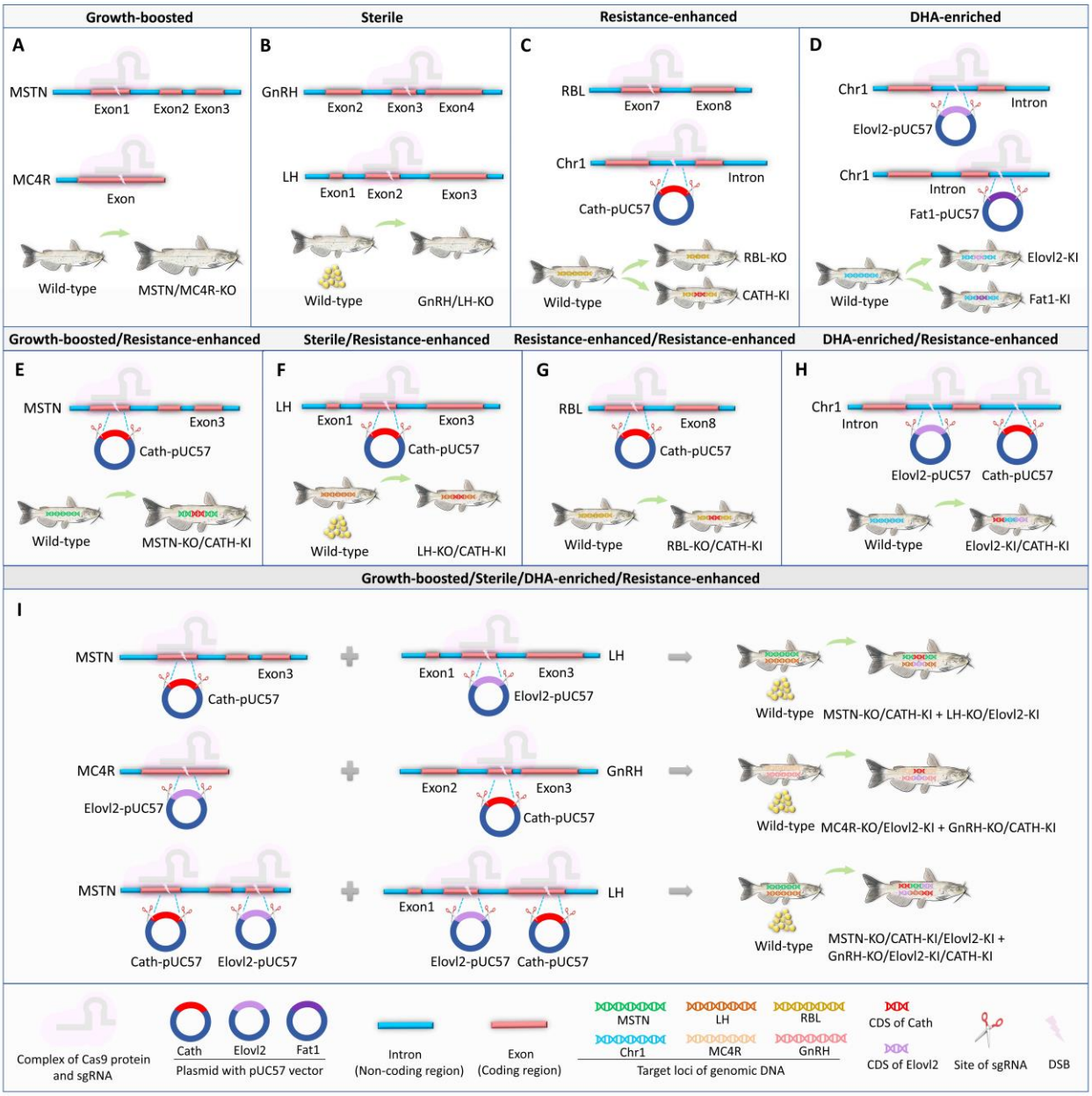


Figure 28. CRISPR/Cas9-mediated transgenesis induces traits of interest to disease resistance combined with sterile, growth-boosted, and DHA-enriched traits in channel catfish (*Ictalurus punctatus*). (**A**) A Growth-boosted fish genetic lines were created by knocking out the *MSTN* or *MC4R* gene. (**B**) Sterile fish lines were produced via knocking out the *GnRH* or *LH* gene. (**C**) Disease resistance-enhanced fish lines were created by knocking out RBL or knocking in a cathelicidin (*Cath*) gene at the non-coding region of chromosome 1. (**D**) DHA-enriched fish line was generated through knocking in the *Elovl2* or *Fat1* gene at the non-coding region of chromosome 1. (**E**) Disease resistance-enhanced fish with fast-growth was produced by knocking in the *Cath* gene at the *MSTN* locus. (**F**) Disease resistance-enhanced fish with sterility were produced by knocking in the *Cath* gene at the *LH* locus. (**G**) A higher disease resistance-enhanced fish line was created by knocking in the *Cath* gene at the *RBL* locus. (**H**)

Disease resistance-enhanced fish with high DHA content was produced by knocking the *Cath* and *Elovl2* genes at the non-coding region of chromosome 1. (**I**) Multiple CRISPR/Cas9 systems produce fish lines that contain enhanced-disease resistance, fast-growth, sterility and enriched-DHA traits. DHA, docosahexaenoic acid; *RBL*, like rhamnose-binding lectin; *Elovl2*, ELOVL fatty acid elongase 2; *Fat1*, humanized omega-3 desaturase; *MSTN*, myostatin; *MC4R*, melanocortin 4 receptor; *GnRH*, gonadotropin-releasing hormone; *LH*, luteinizing hormone; Chr1, chromosome 1; KO, knock out; KI, knock in; Cath-/Elovl2-/Fat1-pUC57, a plasmid containing the *Cath*/*Elovl2*/*Fat1* genes constructed with pUC57 as the vector; CDS, coding sequences; DSB, double-stranded break.

References

Alimuddin, Kiron V, Satoh S, Takeuchi T, Yoshizaki G. Cloning and over-expression of a masu salmon (*Oncorhynchus masou*) fatty acid elongase-like gene in zebrafish. *Aquaculture*. 2008;282:13-18. doi:10.1016/j.aquaculture.2008.06.033

Beck BH, Farmer BD, Straus DL, Li C, Peatman E. Putative roles for a rhamnose binding lectin in *Flavobacterium columnare* pathogenesis in channel catfish *Ictalurus punctatus*. *Fish Shellfish Immunol*. 2012;33:1008-1015. doi:10.1016/j.fsi.2012.08.018

Booy A, Haddow JD, Olafson RW. Isolation of the salmonid rhamnose-binding lectin STL2 from spores of the microsporidian fish parasite *Loma salmonae*. *J Fish Dis*. 2005;28:455-462. doi:10.1111/j.1365-2761.2005.00648.x

Coogan M, Alston V, Su B, et al. CRISPR/Cas-9 induced knockout of myostatin gene improves growth and disease resistance in channel catfish (*Ictalurus punctatus*). *Aquaculture*. 2022;557:738290. doi:10.1016/j.aquaculture.2022.738290

Dunham RA, Warr GW, Nichols A, et al. Enhanced bacterial disease resistance of transgenic channel catfish *Ictalurus punctatus* possessing cecropin genes. *Mar Biotechnol*. 2002;4:338-344. doi:10.1007/s10126-002-0024-y

Elaswad A, Khalil K, Ye Z, et al. Effects of CRISPR/Cas9 dosage on TICAM1 and RBL gene mutation rate, embryonic development, hatchability and fry survival in channel catfish. *Sci Rep*. 2018;8:16499. doi:10.1038/s41598-018-34738-4

Hallerman EM, Bredlau JP, Camargo LSA, et al. Towards progressive regulatory approaches for agricultural applications of animal biotechnology. *Transgenic Res*. 2022;31:167-199. doi:10.1007/s11248-021-00294-3

Hsieh JC, Pan CY, Chen JY. Tilapia hepcidin (TH)2-3 as a transgene in transgenic fish enhances resistance to *Vibrio vulnificus* infection and causes variations in immune-related genes after infection by different bacterial species. *Fish Shellfish Immunol*. 2010;29:430-439. doi:10.1016/j.fsi.2010.05.001

Hsu PD, Scott DA, Weinstein JA, et al. DNA targeting specificity of RNA-guided Cas9

nucleases. *Nat Biotechnol*. 2013;31:827-832. doi:10.1038/nbt.2647

Mehravar M, Shirazi A, Nazari M, Banan M. Mosaicism in CRISPR/Cas9-mediated genome editing. *Dev Biol*. 2019;445:156-162. doi:10.1016/j.ydbio.2018.10.008

Ota S, Hisano Y, Ikawa Y, Kawahara A. Multiple genome modifications by the CRISPR/Cas9 system in zebrafish. *Genes Cells*. 2014;19:555-564. doi:10.1111/gtc.12154

Peatman E, Li C, Peterson BC, Straus DL, Farmer BD, Beck BH. Basal polarization of the mucosal compartment in *Flavobacterium columnare* susceptible and resistant channel catfish (*Ictalurus punctatus*). *Mol Immunol*. 2013;56:317-327. doi:10.1016/j.molimm.2013.04.014

Qin G, Qin Z, Lu C, et al. gene editing of the catfish gonadotropin-releasing hormone gene and hormone therapy to control the reproduction in channel catfish, *Ictalurus punctatus*. *Biology*. 2022;11:649. doi:10.3390/biology11050649

Qin Z, Li Y, Su B, et al. Editing of the luteinizing hormone gene to sterilize channel catfish, *Ictalurus punctatus*, using a modified zinc finger nuclease technology with electroporation. *Mar Biotechnol*. 2016;18:255-263. doi:10.1007/s10126-016-9687-7

Sarmasik A, Warr G, Chen TT. Production of transgenic medaka with increased resistance to bacterial pathogens. *Marine Biotechnology*. 2002;4:310-322. doi:10.1007/s10126-002-0023-z

Wang J, Su B, Bruce TJ, et al. CRISPR/Cas9 microinjection of transgenic embryos enhances the dual-gene integration efficiency of antimicrobial peptide genes for bacterial resistance in channel catfish, *Ictalurus punctatus*. *Aquaculture*. 2023;575:739725. doi:10.1016/j.aquaculture.2023.739725

Wang J, Su B, Dunham RA. Genome-wide identification of catfish antimicrobial peptides: A new perspective to enhance fish disease resistance. *Rev Aquac*. 2022;14:2002-2022. doi:10.1111/raq.12684

Watanabe Y, Tateno H, Nakamura-Tsuruta S, et al. The function of rhamnose-binding lectin in innate immunity by restricted binding to Gb3. *Dev Comp Immunol*. 2009;33:187-197. doi:10.1016/j.dci.2008.08.008

Xing D, Shang M, Li S, et al. CRISPR/Cas9-mediated precision integration of fat-1 and fat-2 from Caenorhabditis elegans at long repeated sequence in channel catfish (*Ictalurus punctatus*) and the impact on n-3 fatty acid level. *Aquaculture*. 2023;567:739229. doi:10.1016/j.aquaculture.2023.739229

Xing D, Su B, Li S, et al. CRISPR/Cas9-mediated transgenesis of the masu salmon (*Oncorhynchus masou*) elovl2 gene improves n-3 fatty acid content in channel catfish (*Ictalurus punctatus*). *Mar Biotechnol*. 2022;24:513-523. doi:10.1007/s10126-022-10110-6

Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. One-step generation of mice

carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*. 2013;154:1370-1379. doi:10.1016/j.cell.2013.08.022

Yang Z, Yu Y, Wang L, Wong SM, Yue GH. Silencing asian seabass gab3 inhibits nervous necrosis virus replication. *Mar Biotechnol*. 2022;24(6):1084-1093. doi:10.1007/s10126-022-10169-1

Appendix 1: Supplement tables

Table S1. Oligonucleotide primers for PCR and qPCR were used in Chapter one.

Construct/Gene	Purpose	Name	Nucleotide sequence (5' → 3')
ssODN1-As-Cath-ssODN2	PCR: Cathelicidin region (591 bp)	Cath1-F	CTCTCGACCATCGGCAGATT
		Cath1-R	GTCTGGATCTCACCGCCTTC
	PCR: Promoter region (519 bp)	Prom1-F	CCGCTTTTCCTGGTCCAGAT
		Prom1-R	TTGGTTGTAGACGTCGACGG
	PCR: SSODN1 region (234 bp)	ssODN1-F	CGGGCTCTTGTACACAGGTT
		ssODN1-R	ATCGACGCTCAAGTCAGAGG
	PCR: SSODN2 region (296 bp)	ssODN2-F	CAGTGGCGATAAGTCGTGTC
		ssODN2-R	GCTACATTCTGCCACACTGC
HA1-As-Cath-HA2	PCR: Cathelicidin region (597 bp)	Cath2-F	TTCAGGAGCCGTACTGTTCC
		Cath2-R	GCATTCTAGTTGTGGTTTGTCCA
	PCR: Promoter region (542 bp)	Prom2-F	ACCCTTTGCCACAGTTCTCC
		Prom2-R	GGCCCTTGGTTGTAGACG
	PCR: Left HA region (573 bp)	HA1-F	TAAGGCCACGTTTCGATTCT
		HA1-R	TCATTTTGCCGTCTGTTGTT
	PCR: Right HA region (598 bp)	HA2-F	TGAGTTTGGACAAACCACAAC
		HA2-R	TTGATTGAAAATGTTTCCCTGTT
	qPCR: Cathelicidin (125 bp)	Cath_RT-F	GCAGGGGTCTATTCAAGAAGC
		Cath_RT-R	GTCTGGATCTCACCGCCTTC
lh gene	PCR (594 bp)	LH-F	TGAGCGATCACAGCAAAATC
		LH-R	GCAGCTTAGTGCGACAGGAT
	qPCR (147 bp)	LH_RT-F	TGAGCGATCACAGCAAAATC
		LH_RT-R	TTAACAGGTTCGCAGTGTGG
18S rRNA	qPCR (128 bp)	18s-F	GAGAAACGG CTACCACATCC
		18s-R	GATACGCTCATT CCGATTACAG

Table S2. The summary of total knock-in (KI) efficiency and on-target KI efficiency of different CRISPR/Cas9-mediated systems.

System	Percent % (number)		
	On-target KI	Off-target KI	Only KO
2H2OP50	0.75%(1/134)	26.87%(36/134)	5.56%(3/54)
2H2OP100	0(/152)	17.76%(27/152)	6.67%(4/60)
dsDNA50	10.80%(23/213)	1.41%(3/213)	3.33%(2/60)
dsDNA100	6.97(17/244)	3.28%(8/244)	3.33%(2/60)

Table S3. The odds ratio (OR) calculation of total knock-in (KI) efficiency and on-target KI efficiency.

		2H2OP	dsDNA	50 ng/μL	100 ng/μL
Total KI	Positive	64	51	63	52
	Negative	222	406	284	344
On-target KI	Positive	1	40	24	17
	Negative	285	427	323	379

Comparison of total KI efficiency:
2H2OP *vs.* dsDNA: OR = (64×406)/(222×51) = 2.30 times
50 ng/μL *vs.* 100 ng/μL: OR = (63×344)/(284×52) = 1.47 times

Comparison of on-target KI efficiency:
dsDNA *vs.* 2H2OP: OR = (285×40)/(1×427) = 26.70 times
50 ng/μL *vs.* 100 ng/μL: OR = (24×379)/(323×17) = 1.66 times

Table S4. A full model based on all 12 predictors involved since there was no multi-collinearity in the gene expression matrix.

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	13.5313	10.1977	1.7606	0.1845
BW	1	0.1743	0.1150	2.2955	0.1298
Sex	1	0.3805	0.9167	0.1723	0.6781
Cath	1	0.2573	0.1178	4.7686	0.0290
Cec	1	0.2420	0.1416	2.9211	0.0874
CCL3	1	0.3034	0.2442	1.5438	0.2140
LEAP2	1	-15.2284	7.3658	4.2743	0.0387
H2A	1	1.4132	0.8672	2.6557	0.1032
UBI	1	-0.0350	0.6614	0.0028	0.9578
BPI	1	0.3067	0.2922	1.1021	0.2938
TCP	1	1.2674	0.4773	7.0520	0.0079
Catpd	1	0.5337	0.3991	1.7885	0.1811
NK_lysin	1	-0.3884	0.1639	5.6156	0.0178

Table S5. Fit statistics of the full model with AIC value.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	112.703	69.304
SC	115.085	100.270
-2 Log L	110.703	43.304

Table S6. The first reduced model after removing non-significant ($P > 0.05$) predictors (BW, Sex, *CCL3*, *H2A*, *UBI*, *BPI* and *Catpd*). Here, we kept the *Cec* for the reduced model since the $P = 0.0874$ and it was the transgene that we focused on.

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	8.5466	5.9352	2.0736	0.1499
Cath	1	0.1788	0.0893	4.0101	0.0452
Cec	1	0.2801	0.1140	6.0362	0.0140
LEAP2	1	-6.2088	3.3751	3.3841	0.0658
TCP	1	1.4105	0.5255	7.2051	0.0073
NK_lysin	1	-0.3597	0.1363	6.9674	0.0083

Table S7. Fit statistics of the full model with AIC value. As we observed, this fitted model had a smaller AIC value (62.776) than that of the full model (AIC = 69.304). Therefore, we believed that this reduced model was more robust than the full model. Nonetheless, we found LEAP2 was not significant ($P = 0.0658$), and we removed it to fit a reduced model correspondingly.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	112.703	62.776
SC	115.085	77.068
-2 Log L	110.703	50.776

Table S8. A second reduced model only contained *Cath*, *Cec*, *TCP* and *NK-lysin*.

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.9481	1.7098	1.2982	0.2545
Cath	1	0.1666	0.0840	3.9359	0.0473
Cec	1	0.2331	0.1102	4.4772	0.0344
TCP	1	1.4811	0.5469	7.3347	0.0068
NK_lysin	1	-0.3892	0.1314	8.7717	0.0031

Table S9. Fit statistics of the reduced model with AIC value.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	112.703	64.585
SC	115.085	76.495
-2 Log L	110.703	54.585

Table S10. Potential logit interactions among these predictors were assessed using a new model. The results indicated that there were no significant interactions among them (all $P > 0.05$).

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	5.9369	3.4003	3.0485	0.0808
Cath	1	0.1822	1.1568	0.0248	0.8749
Cec	1	-35.4348	31.8365	1.2388	0.2657
Cath*Cec	1	-52.3986	833.9	0.0039	0.9499
TCP	1	-1.1059	1.1651	0.9009	0.3425
Cath*TCP	1	0.0235	0.4174	0.0032	0.9552
Cec*TCP	1	13.2537	11.9656	1.2269	0.2680
Cath*Cec*TCP	1	17.5803	280.9	0.0039	0.9501
NK_lysin	1	-1.7925	0.6970	6.6145	0.0101
Cath*NK_lysin	1	-0.1257	0.2606	0.2327	0.6296
Cec*NK_lysin	1	3.7484	3.2716	1.3127	0.2519
Cath*Cec*NK_lysin	1	6.3180	99.3898	0.0040	0.9493
TCP*NK_lysin	1	0.4758	0.2443	3.7940	0.0514
Cath*TCP*NK_lysin	1	0.0294	0.0853	0.1191	0.7300
Cec*TCP*NK_lysin	1	-1.3699	1.2123	1.2769	0.2585
Cath*Cec*TCP*NK_lysi	1	-2.1410	33.8270	0.0040	0.9495

Table S11. Fit statistics of the model with AIC value when we took interactions into the model. The AIC = 71.035 > 64.585 indicated that the fitness was reduced when we took the interactions into the model. Therefore, we fitted the final model without interactions involved.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	112.703	71.035
SC	115.085	109.148
-2 Log L	110.703	39.035

Table S12. The final model was subjected to the sensitivity/specificity test and the Hosmer-Lemeshow test to determine the sensitivity and robustness. And the Goodness-of-Fit test revealed that our final logit model fitted the data well ($P = 0.2063 > 0.05$).

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
10.9201	8	0.2063

Table S13. Oligonucleotide sequences for single guide RNA (sgRNA) synthesis and transgene/mutagenesis detection. PCR primers were used to detect *As-Cath*/*Cec* transgene, and *lh*/*mc4r*/*mstn1*/*mstn2* mutation in putative transgenic/gene-edited channel catfish, *Ictalurus punctatus*. *Cec*, the cecropin gene from moth; *As-Cath*, the cathelicidin gene from alligator; *lh*, luteinizing hormone; *mc4r*, melanocortin-4 receptor; *mstn1*, the exon1 of myostatin; *mstn2*, the exon2 of myostatin.

Oligo name	Nucleotide sequence (5' → 3')	Product Size (bp)	Purpose
sgRNA synthesis			
sgRNA1	TTCAAACCGCCATCTGCAGC**GGG**	—	*lh*_sgRNA synthesis
sgRNA2	GCAGCTGTTGATCTCCACCG**AGG**	—	*mc4r*_sgRNA synthesis
sgRNA3	TCTGATTTCGCTGGGCTTCG**TGG**	—	*mstn1*_sgRNA synthesis
sgRNA4	**CCC**CGACGTTCAAGTCGACCAAA	—	*mstn2*_sgRNA synthesis
Universal Primer	TTTTGCACCGACTCGGTGCCACTT TTTCAAGTTGATAACGGACTAGCC TTATTTTAACTTGCTATTTCTAGCT CTAAAAC	—	Scaffold of the sgRNA synthesis
PCR Primers			
*As-Cath*_F	TTCAGGAGCCGTACTGTTCC	597	Determine the *As-Cath*_polyA region of *As-Cath* transgenic fish
*As-Cath*_R	GCATTCTAGTTGTGGTTTGTCCA		
*Cec*_F	GGAGCCGTACTGTTCCGTTA	352	Determine the *Cec*_polyA region of *As-Cath* transgenic fish
*Cec*_R	CCCATATGTCCTTCCGAGTG		
Prom1_F	GCAGCCAATCACTGCTTGTA	462	Determine the UBI_*Cec* region of *Cec* transgenic fish
Prom1_R	ATTCCGAGGACCTGGATTG		
Prom2_F	ACCCTTTGCCACAGTTCTCC	542	Determine the UBI_*As-Cath* region of *As-Cath* transgenic fish
Prom2_R	GGCCCTTGGTTGTAGACG		
*lh*_F	TGAGCGATCACAGCAAAATC	594	Determine the *lh* mutation
*lh*_R	GCAGCTTAGTGCGACAGGAT		
*mc4r*_F	CTGCTCTTCCTCATCCTTCG	598	Determine the *mc4r* mutation
*mc4r*_R	ATGCTTTTCACGACGTCTCC		
*mstn1*_F	CATGACATCTCGCGCTACCT	390	Determine the *mstn1* mutation
*mstn1*_R	GCAGCTGCTTGACCACATC		
*mstn2*_F	CCGTGTTCGGTTGTGTGTAG	480	Determine the *mstn2* mutation
*mstn2*_R	ATCTCAATCCCCCAGTTGGT		
HA1_F	TAAGGCCACGTTTCGATTCT	533	Determine the junction of HA1 at the *lh* locus
HA1_R	TCATTTTGCCGTCTGTTGTT		
HA2_F	TGAGTTTGGACAAACCACAA	598	Determine the junction of HA2 at the *lh* locus
HA2_R	TTGATTGAAAATGTTTCCCTGTT		
HA3_F	AAAGTGCCAATCTGCCAAAG	467	Determine the junction of HA3 at the *mc4r* locus
HA3_R	CCGGCTTTGTTTCCAATCT		
HA4_F	ATCACGCTAGGGTTGGTCAG	463	Determine the junction of HA4 at the *mc4r* locus
HA4_R	GCATGGTGAAGAACATGCTG		
HA5_F	TGGAGAAAGTTGTGGGTCTGT	334	Determine the junction of HA5

HA5_R	CCCAGCGAAATCAGAACCT		at the *mstn1* locus
HA6_F	ACATATGGGAGGGCAAATCA	574	Determine the junction of HA6
HA6_R	AAGCAGTAGTAAAGGGACTCACG		at the *mstn1* locus
HA7_F	GAATCGTTTCAGAATGGACGA	395	Determine the junction of HA7
HA7_R	GGGGTTGGCTAAAGGAGAGA		at the *mstn2* locus
HA8_F	ACCAAAAACCGAAGTGCTGT	498	Determine the junction of HA8
HA8_R	ATCCCATTTCAACCAGCAAA		at the *mstn2* locus

Table S14. The hatchability and fry survival of channel catfish (*Ictalurus punctatus*) in the family 1.

Group	Hatchability					Fry survival				
	nCT	iCT	ssGE1	ssGE2	msMGE	nCT	iCT	ssGE1	ssGE2	msMGE
Rep1	48	23	30	34	24	48	23	30	34	21
Rep2	65	30	25	26	33	62	28	20	24	30
Rep3	82	24	32	34	21	78	24	30	29	17
Rep4		32	24	21	36		30	20	18	32
Rep5		21	17	40	16		21	16	34	14
Survival	195	130	128	155	130	188	126	106	119	86

Table S15. The hatchability and fry survival of channel catfish (*Ictalurus punctatus*) in the family 2.

Group	Hatchability					Fry survival				
	nCT	iCT	ssGE1	ssGE2	msMGE	nCT	iCT	ssGE1	ssGE2	msMGE
Rep1	38	13	18	14	14	30	10	16	12	14
Rep2	25	20	15	16	13	20	12	15	16	13
Rep3	42	14	21	14	21	34	12	20	13	18
Rep4		22	24	21	16		18	18	19	16
Rep5		11	13	10	6		8	13	10	6
Survival	105	80	91	75	70	84	60	52	70	57

Table S16. The hatchability and fry survival of channel catfish (*Ictalurus punctatus*) in the family 3.

Group	Hatchability					Fry survival				
	nCT	iCT	ssGE1	ssGE2	msMGE	nCT	iCT	ssGE1	ssGE2	msMGE
Rep1	78	45	50	44	34	74	44	50	40	33
Rep2	93	28	45	46	40	90	26	42	40	37
Rep3	101	38	40	44	32	91	35	35	38	32
Rep4		42	34	30	36		36	30	30	36
Rep5		44	42	26	36		40	42	24	35
Survival	272					255	181	189	172	153

PCR positive individuals from ssGE1, ssGE2 and msMGE:

ssGE1: *mc4r_As-cath* (106+52+189=347 fish)

PCR positive *As-Cath*: 12/106=11.32%; 5/52=9.62%, 28/189=14.81% (45/347=12.96%)

Off-target events: 4/106=3.77% (4/12=33.33%); 3/52=5.77% (3/5=60%); 10/189=5.29% (10/28=35.71%) (17/45=37.78%)

45 fish have *As-Cath* transgene. Specifically, HDR-mediated KI: 8/12=66.67%; 2/5=40%; 18/28=64.29% (28/45=62.22%)

15 fish have *mc4r* mutation. NHEJ-mediated KO: 5/94=5.32%; 0/47=0; 10/161=6.21% (15/302=4.97%)

ssGE2: *mstn1_Cec* (119+70+172=361 fish)

PCR positive *Cec*: 9/119=7.56%; 8/70=11.43%, 22/172=12.79% (39/361=10.80%)

Off-target events: 4/119=3.36% (4/9=44.44%); 3/70=4.29% (3/8=37.5%); 10/172=5.81% (10/22=45.45%) (17/39=43.59%)

39 fish have *Cec* transgene. Specifically, HDR-mediated KI: 5/9=55.56%; 5/8=62.5%; 12/22=54.55% (22/39=56.41%)

8 fish have *mstn1* mutation. NHEJ-mediated KO: 1/110=0.91%; 0/62=0; 7/150=4.67% (8/322=2.48%)

msMGE: 86+57+153=296 fish

PCR positive *As-Cath*: 15/86=17.44%; 11/57=19.30%, 33/153=21.57% (59/296=19.93%)

Off-target events: (9/15=60%); (7/11=63.63%); (19/33=57.58%) (35/59=59.32%)

59 fish have *As-Cath* transgene. Specifically, HDR-mediated KI: 6/15=40% (*mc4r_As-Cath*, 6 and *lh_As-Cath*, 0); 4/11=36.36%(*mc4r_As-Cath*, 4 and *lh_As-Cath*, 0); 14/33=42.42% (*mc4r_As-Cath*, 10 and *lh_As-Cath*, 4) (24/59=40.68%)

NHEJ-mediated KO: 1/71=1.41% (*lh* KO, 1); 2/46=4.35% (*lh* KO 1, *mc4r* KO 1); 8/120=6.67% (*lh* KO 6, *mc4r* KO, 2) (11/237=4.64%)

24 fish have *As-Cath* insertion on target: 20 target at *mc4r*, 4 target at *lh*

11 fish have *lh* or *mc4r* mutation: 8 fish have *lh* KO, 3 fish have *mc4r* KO

PCR positive *Cec*: 18/86=20.93%; 13/57=22.81%, 37/153=24.18% (68/296=22.97%)

Off-target events: (11/18=61.11%); (8/13=61.54%); (26/37=70.27%) (45/68=66.18%)

68 fish have *Cec* transgene. Specifically, HDR-mediated KI: 7/18=38.89% (*mstn1_Cec* 2, *mstn2_Cec* 5); 5/13=38.46% (*mstn1_Cec* 1, *mstn2_Cec* 4); 11/37=29.73% (*mstn1_Cec* 4, *mstn2_Cec* 7) (23/68=33.82%)

NHEJ-mediated KO: 2/68=2.94% (*mstn1* KO 2, *mstn2* KO 0); 2/44=4.55% (*mstn1* KO 2, *mstn2* KO 0); 12/116=10.34% (*mstn1* KO 9, *mstn2* KO 3) (16/228=7.02%)

23 fish have *Cec* insertion on target: 7 target at *mstn1*, 16 target at *mstn2*

16 fish have *mstn1* or *mstn2* mutation: 13 fish have *mstn1* KO, 3 fish have *mstn2* KO

Both: PCR positive *As-Cath* and *Cec*: 11/86=12.79%; 6/57=10.53%; 15/153=9.80% (32/296=10.81%)

32 fish have both *As-Cath* and *Cec* transgenes, and 6 fish are on-target. On-target positive *As-Cath* and *Cec*: 2/11=18.18% [#7: 2 copy of *As-Cath* (at *lh* and *mc4r* loci) and 2 copy of *Cec* (at *mstn1* and *mstn2*); #12: 1 copy of *As-Cath* (at *mc4r*) and 2 copy of *Cec* (1 copy at *mstn1* and 1 copy at *mstn2*)];

0/6=0; 4/15=6.67% [#11 and #25: 2 copy of *As-Cath* (1 copy at *lh* locus and 1 copy at *mc4r*) and 2 copy of *Cec* (1 copy at *mstn1* and 1 copy at mstn2); #22 and #3: 1 copy of *As-Cath* (at *mc4r* locus), 2 copy of *Cec* (1 copy at *mstn1* and 1 copy at *mstn2*)] (6/32=18.75%)

Totally, 347+361+296=1004 fish

Appendix 2: Supplement figures

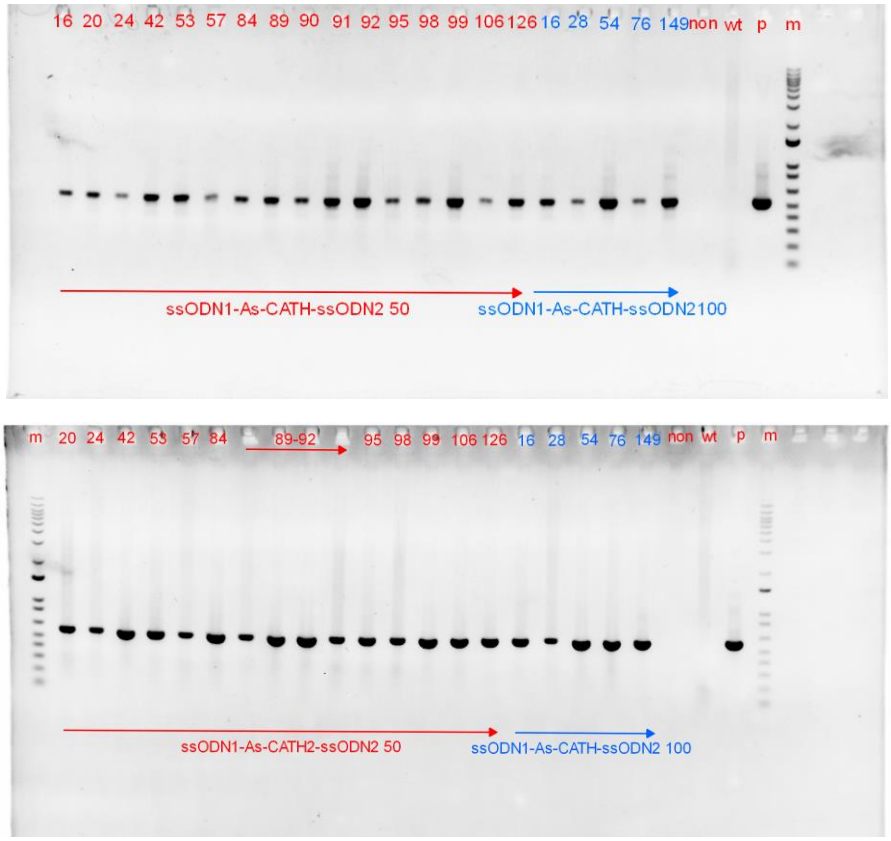


Figure S1. Full 1% TAE gel pictures of transgene detection from the ssODN1-As-Cath-ssODN2 construct (2H2OP system).

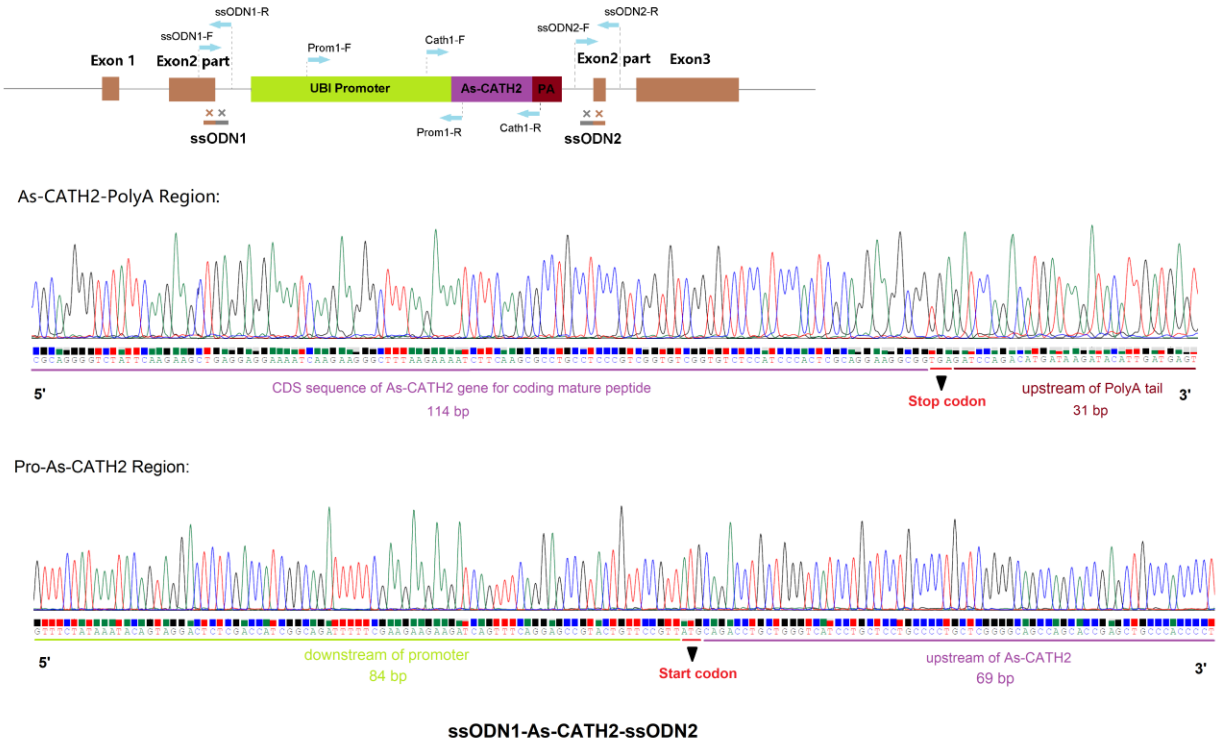


Figure S2. Full sequencing results of transgene detection from the ssODN1-As-Cath-ssODN2 construct (2H2OP system).

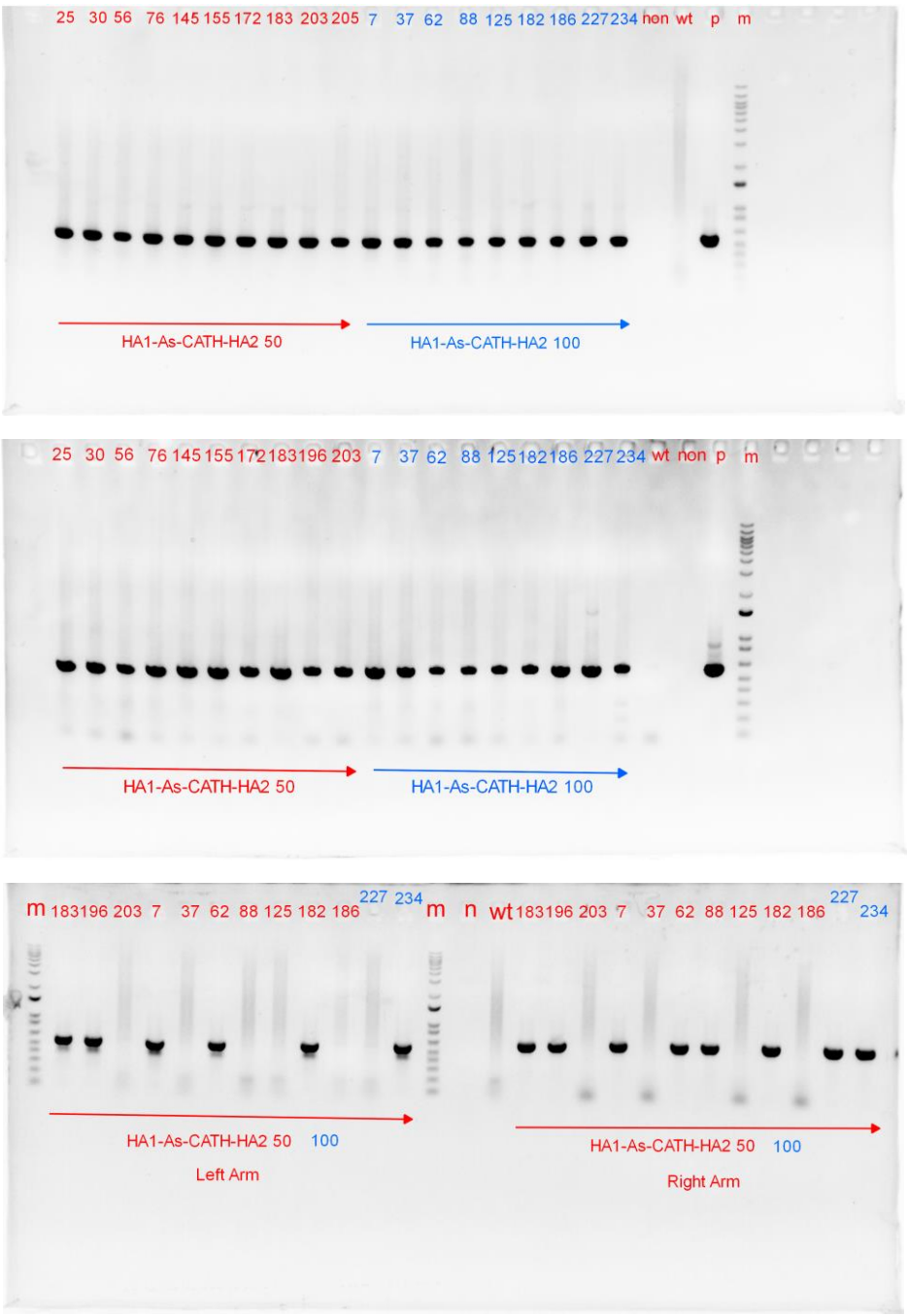


Figure S3. Full 1% TAE gel pictures of transgene detection from the HA1-As-Cath-HA2 construct (dsDNA system).

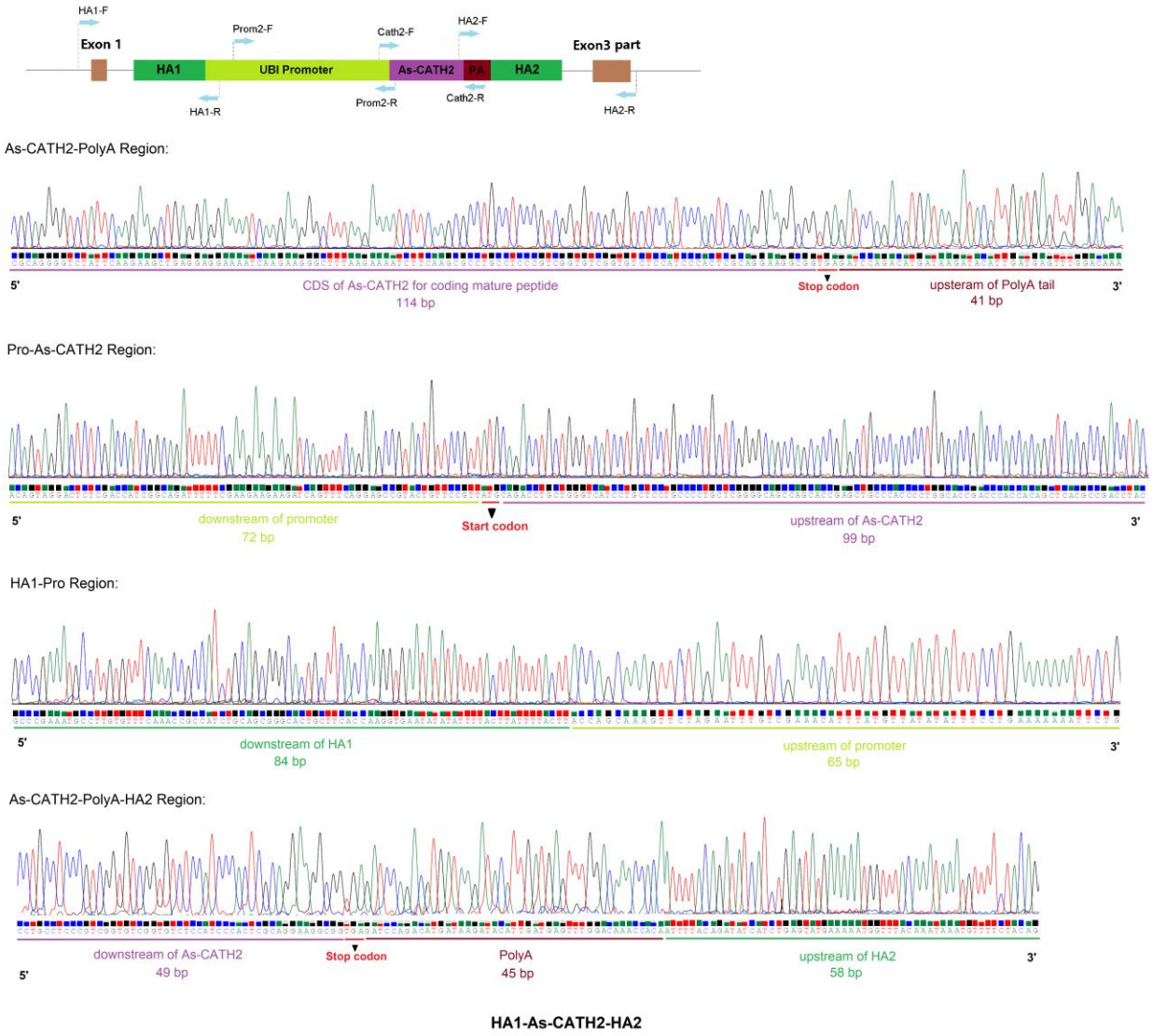


Figure S4. Full sequencing results of transgene detection from the HA1-As-Cath-HA2 construct (dsDNA system).

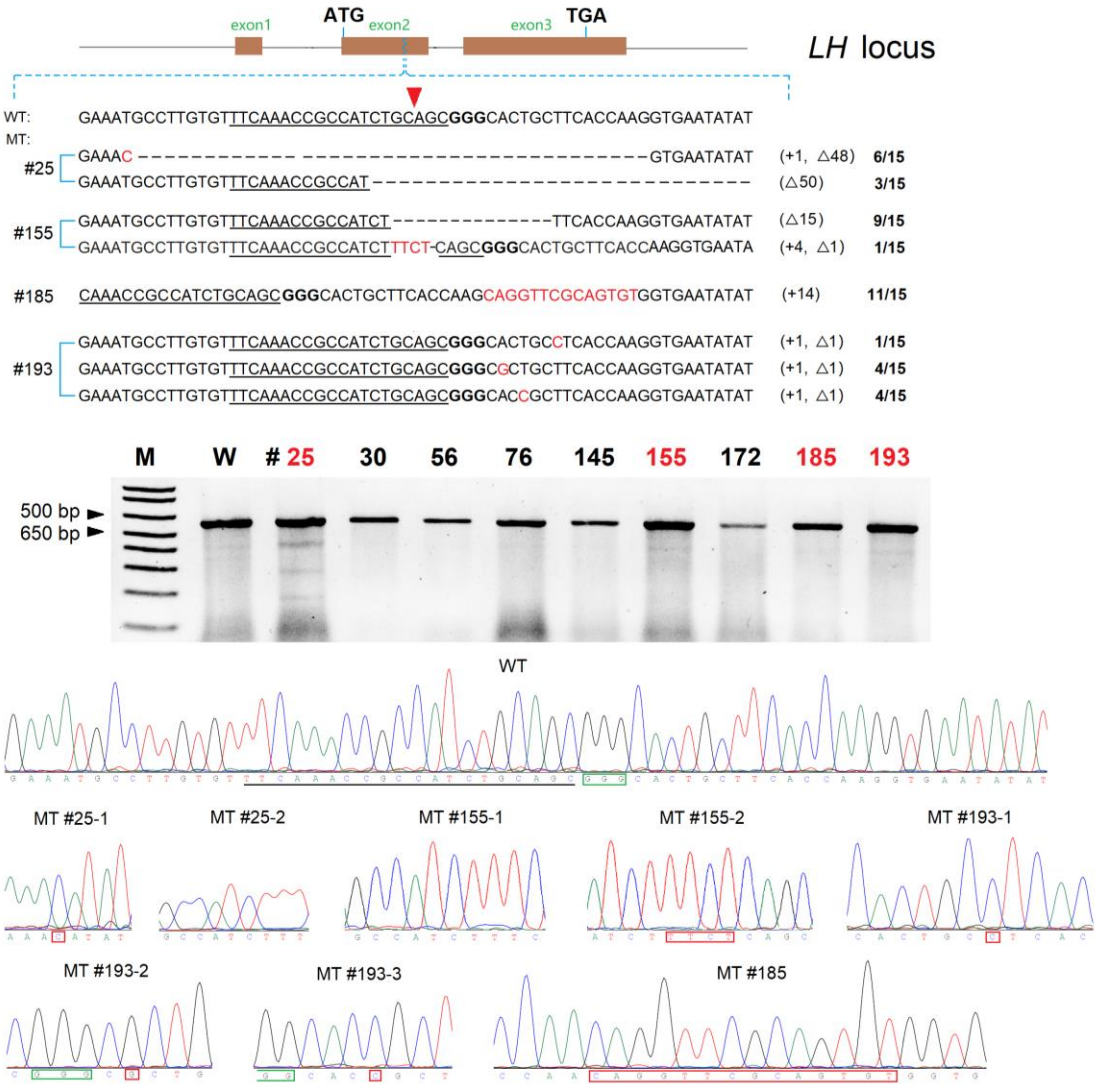
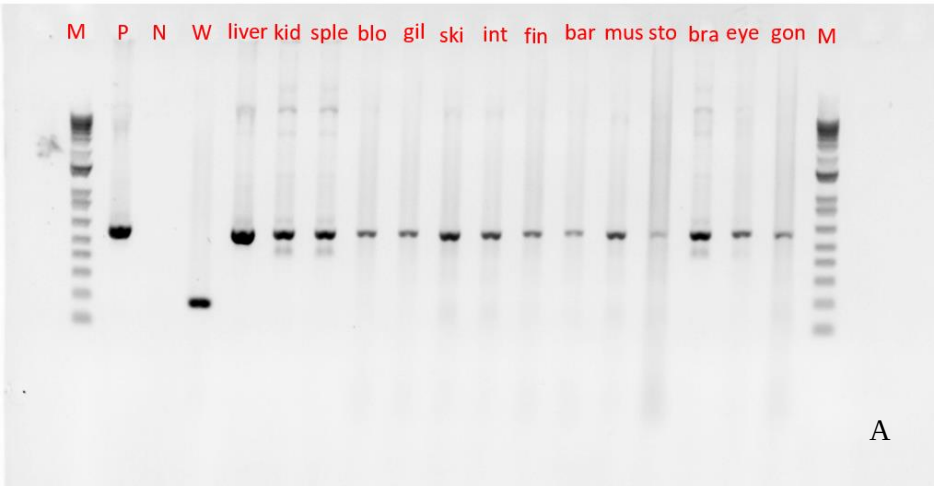


Figure S5. The detection of *lh* mutation from 4 non-*As-Cath*-integrated channel catfish. The number above each lane indicates the IDs of each individual. The gel picture and corresponding sequencing results were inserted.



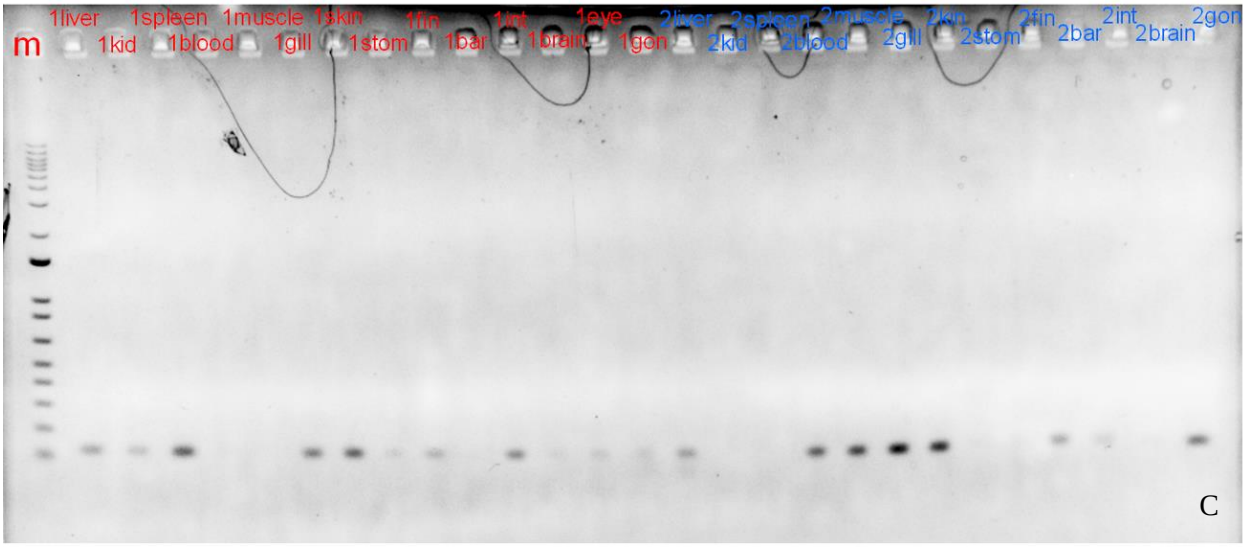


Figure S6. PCR and qRT-PCR analysis for mosaicism. (**A**) PCR detection of *As-Cath* gene for 14 different tissues. (**B**) A representative LH^-_As-$Cath^+$ fish showed *As-Cath* expression in 14 tissues. (**C**) Two LH^-_As-$Cath^+$ individuals have the *As-Cath* expression in 11 and 8 tissues, respectively, indicating two mosaics.

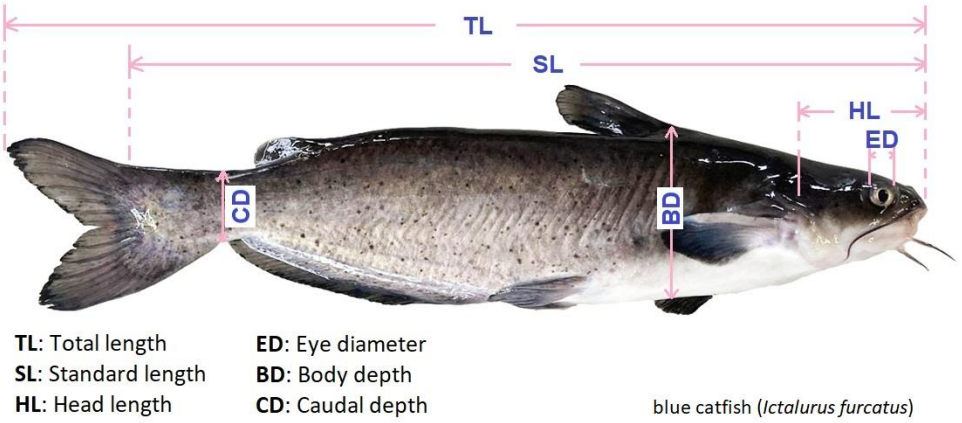


Figure S7. Schematic diagram of body size parameter measurements to determine body shape in blue catfish (*Ictalurus furcatus*).

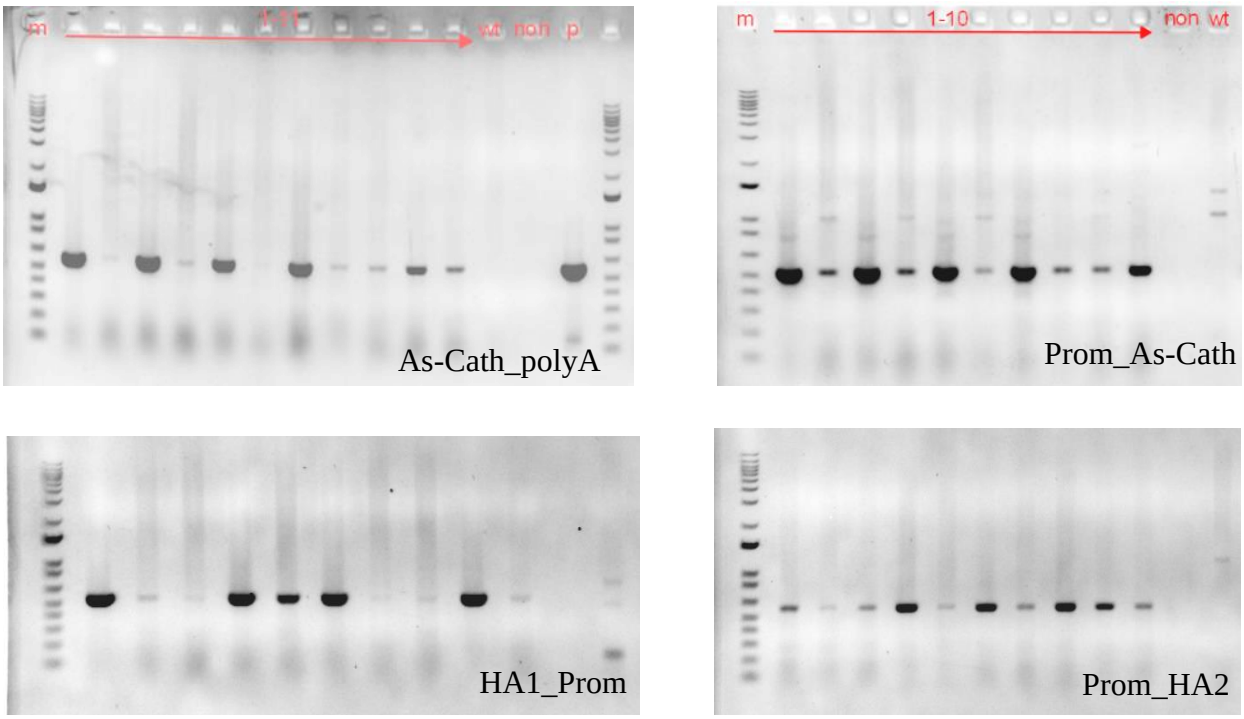


Figure S8. Full gel electrophoresis to identify the *As-Cath* transgenic blue catfish (*Ictalurus furcatus*).

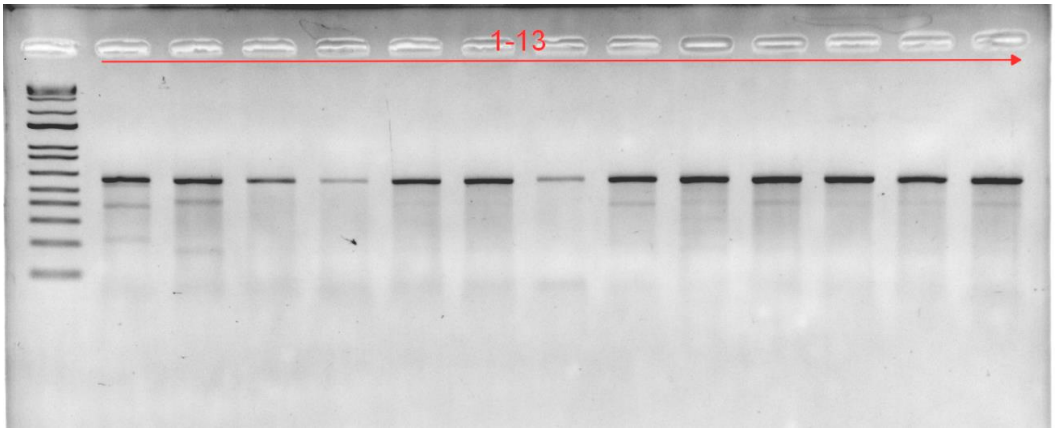


Figure S9. Full gel electrophoresis to identify the *As-Cath* transgenic blue catfish (*Ictalurus furcatus*).

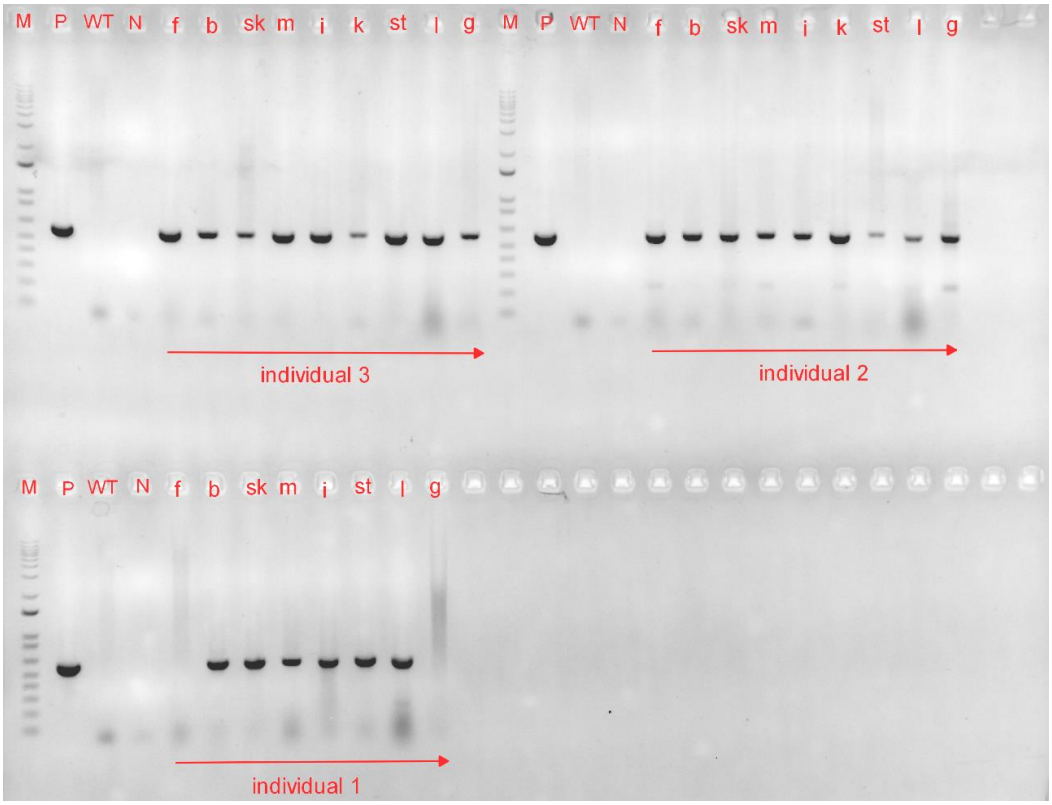


Figure S10. Detection of the *As-Cath* transgene in various from three positive blue catfish (*Ictalurus furcatus*).

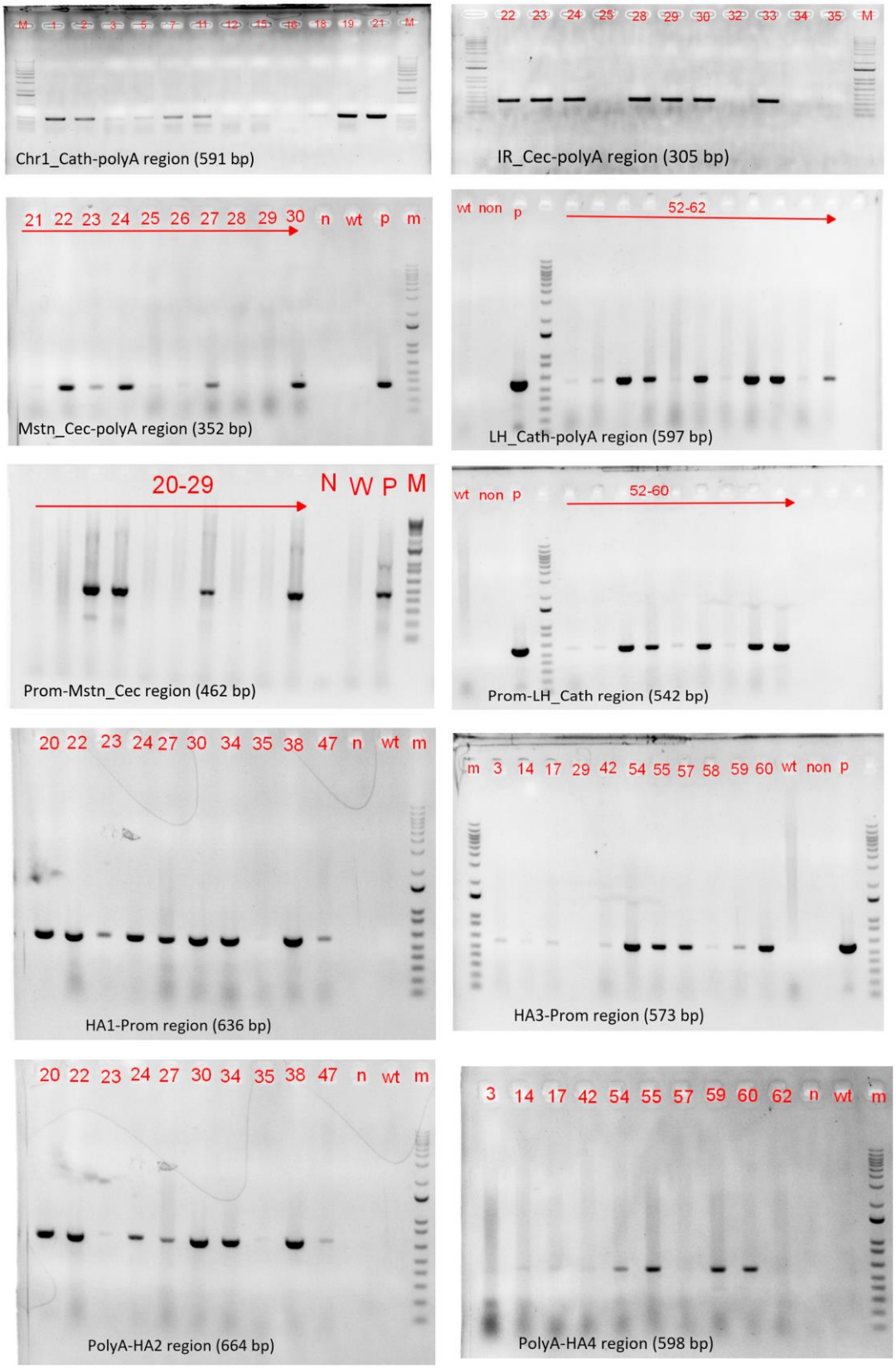


Figure S11. Full gel pictures for the transgenic channel catfish determination.

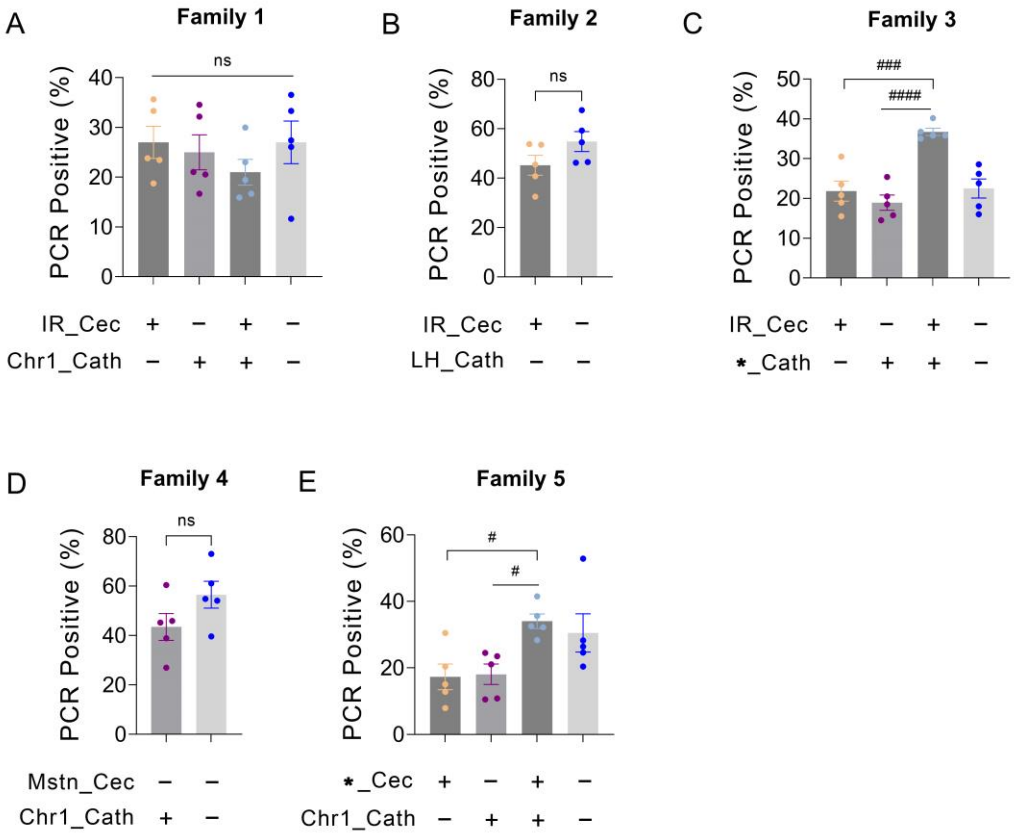


Figure S12. The resulting inheritance and integration rate of transgenes was aided by the CRISPR/Cas9-mediated microinjection using transgenic embryos. The inheritance and integration rate of the cecropin (*Cec*) and cathelicidin (*Cath*) transgenes by PCR detection in different genotypes from five families. The embryos from family 3 and family 5 received a microinjection using pUC57_Cath and pUC57_Cec plasmids, respectively. IR_Cec, the *Cec* transgene was integrated randomly; Chr1_Cath, the *Cath* transgene was integrated at the noncoding region of chromosome 1; LH_Cath, the *Cath* transgene was integrated at the *lh* locus; Mstn_Cec, the *Cec* transgene was integrated at the *mstn* locus; *_Cath, the *Cath* transgene was integrated at the *lh* locus (on-target positives) or non-*lh* locus (off-target positives) after microinjecting the pUC57_Cath plasmid donor; *_Cec, the *Cec* transgene was integrated at the *mstn* locus (on-target positives) or non-*mstn* locus (off-target positives) after microinjecting the pUC57_Cec plasmid donor; PCR Positive, the transgene were detected by corresponding primers; Cec, cecropin; Cath, cathelicidin; # = $P < 0.05$; ## = $P < 0.01$; ### = $P < 0.001$; #### = $P < 0.0001$; ns = not significant, by unpaired student's *t*-test or one-way ANOVA.

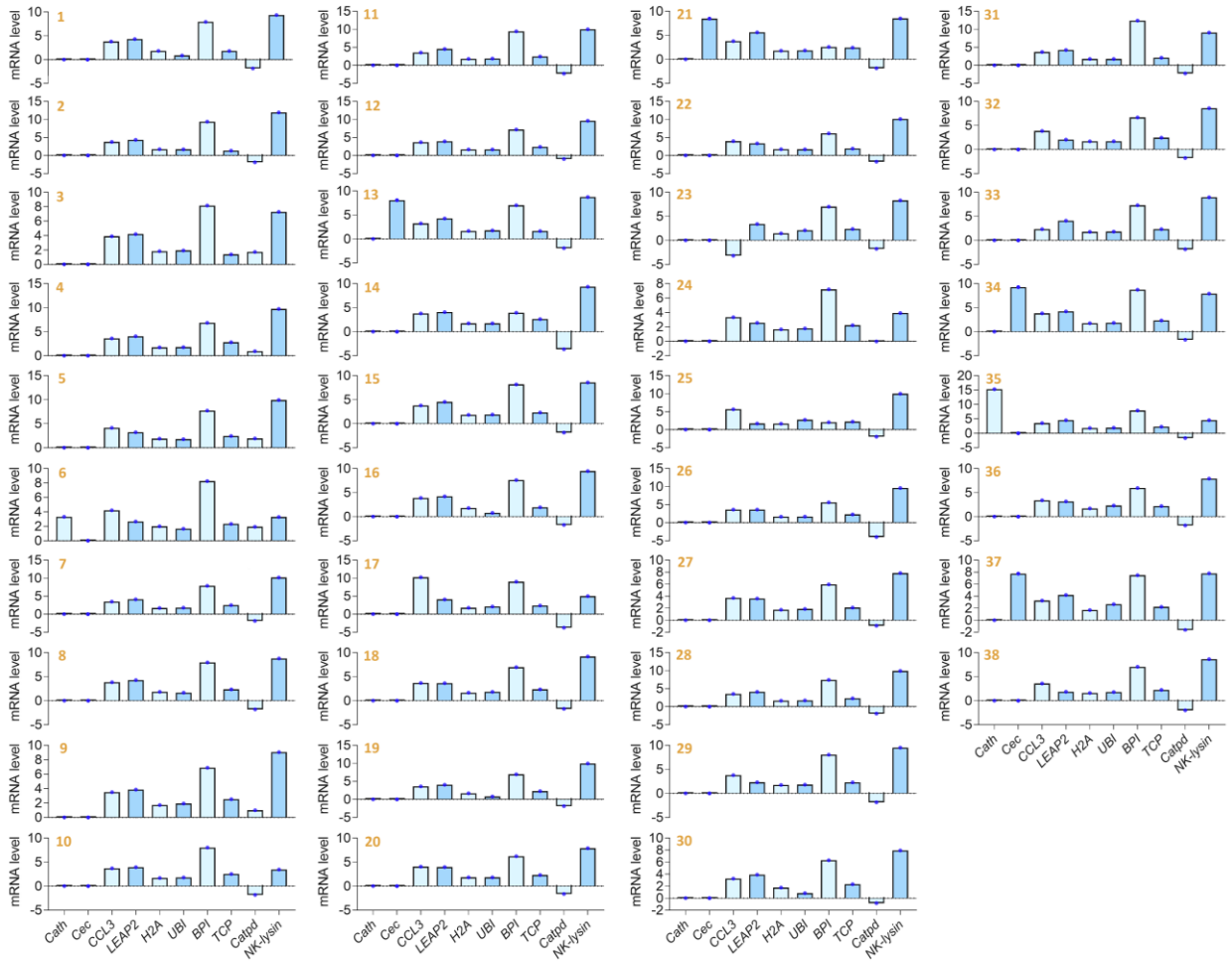


Figure S13. Relative expression of transgenes (*Cec* and *Cath*) and innate AMGs (*CCL3*, *LEAP2*, *H2A*, *UBI*, *BPI*, *TCP*, *Catpd* and *NK-lysin*) in the head kidney of each moribund fish after *Edwardsiella ictaluri* challenge in channel catfish. Bars below the 0 line indicate reduced mRNA levels compared to the control while bars above the 0 line indicate increased expression compared to the control. The inserted number of each histogram indicates the number of fish.

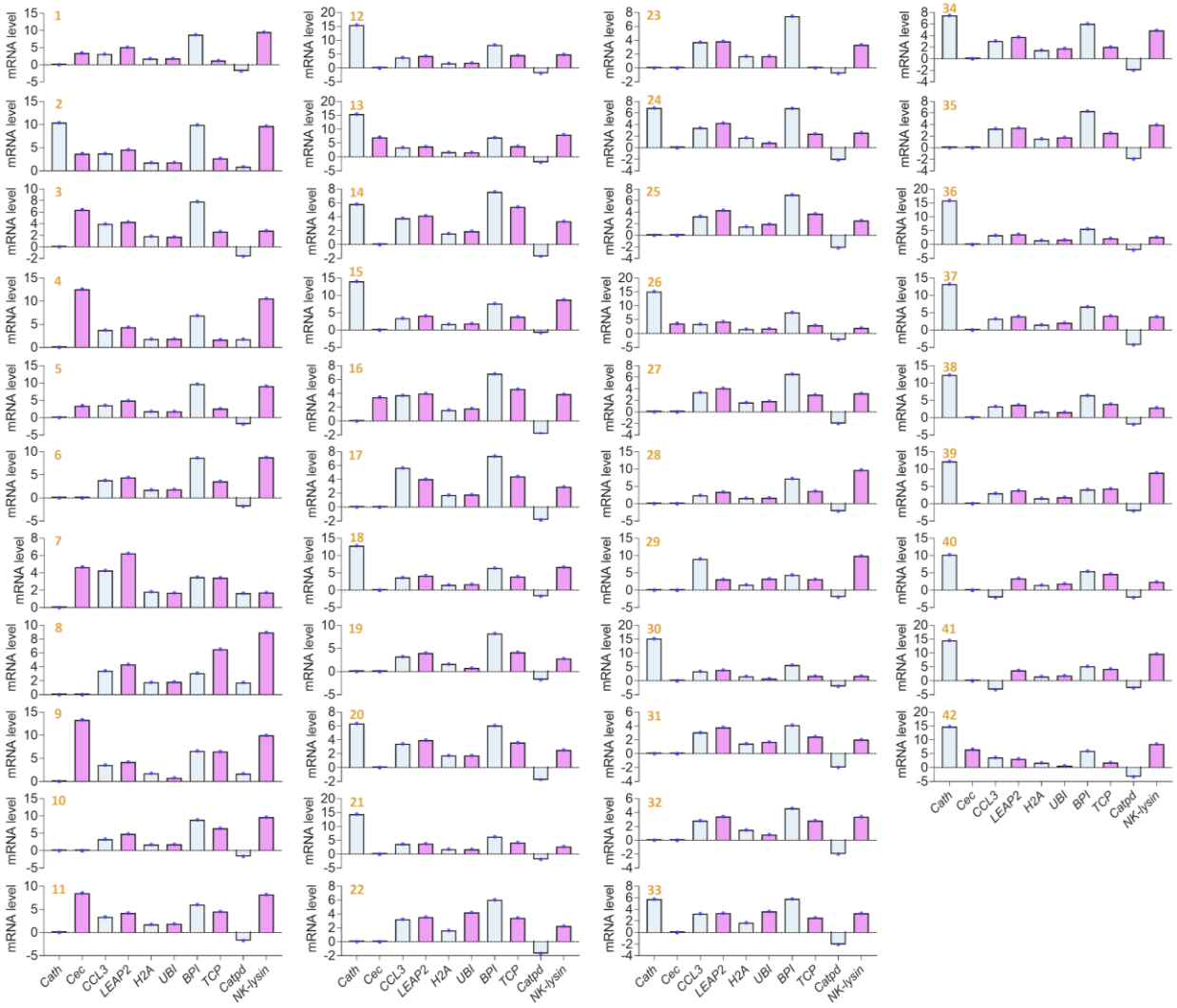


Figure S14. Relative expression of transgenes (*Cec* and *Cath*) and innate AMGs (*CCL3*, *LEAP2*, *H2A*, *UBI*, *BPI*, *TCP*, *Catpd* and *NK-lysin*) in the head kidney of each sacrificed fish after *Edwardsiella ictaluri* challenge in channel catfish. Bars below the 0 line indicate reduced mRNA levels compared to the control while bars above the 0 line indicate increased expression compared to the control. The inserted number of each histogram indicates the number of fish.

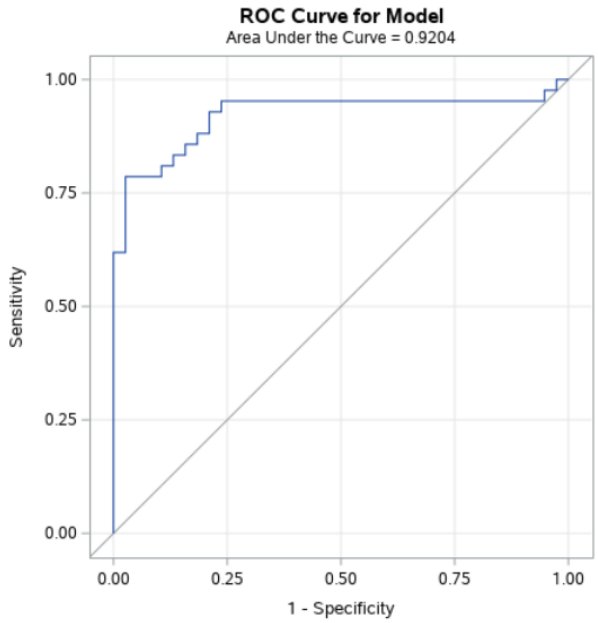


Figure S15. The ROC curve showed that we had 92.04% confidence that higher mRNA levels of *Cath*, *Cec*, *TCP* and lower mRNA levels of *NK-lysin* significantly improved survival rate post- *Edwardsiella ictaluri* infection.

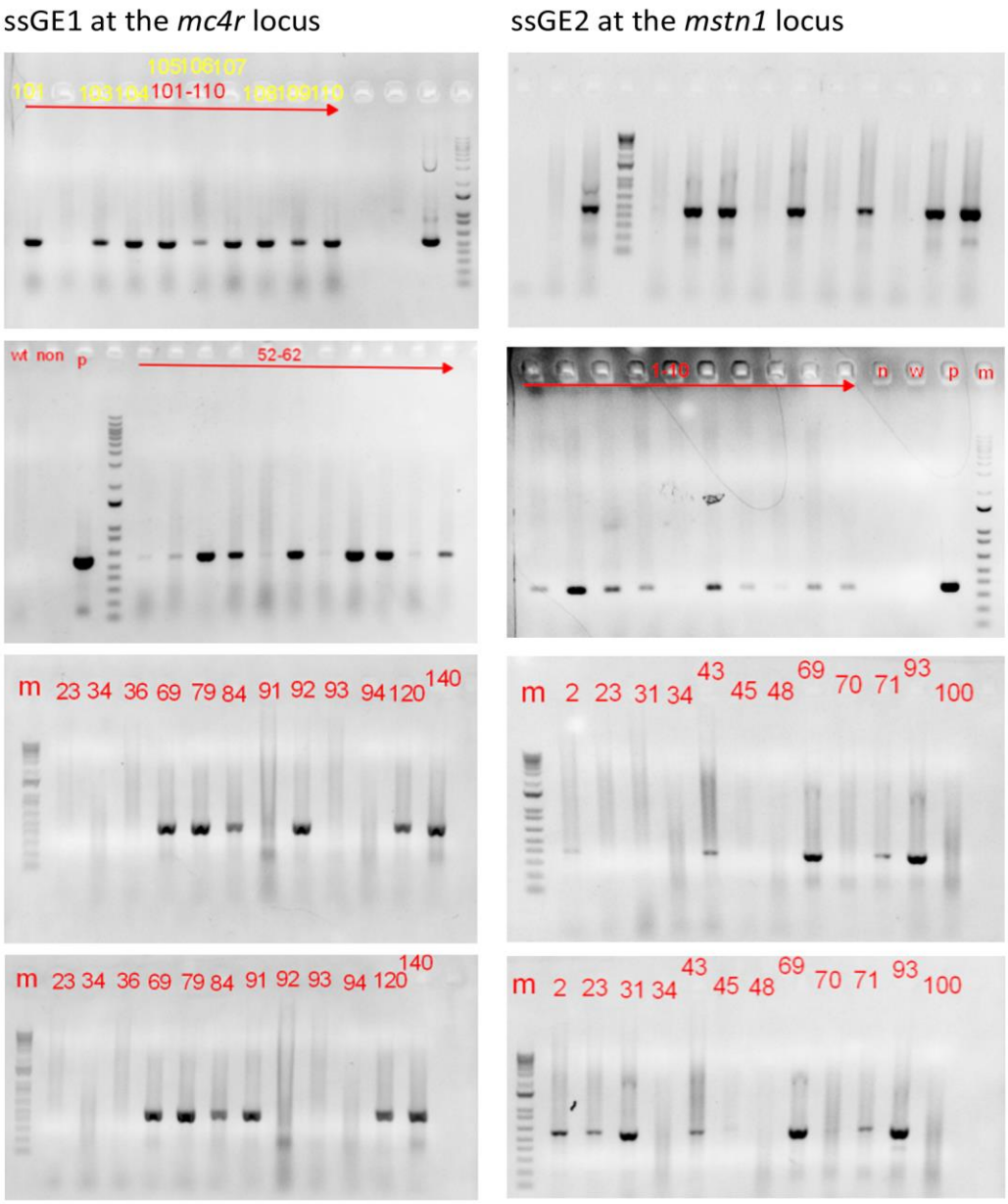


Figure S16. Full-length gel electrophoresis images to determine the *As-Cath*/*Cec* transgenic positive channel catfish (*Ictalurus punctatus*) in the ssGE1/ssGE2 strategy. *Cec*, the cecropin gene from moth; *As-Cath*, the cathelicidin gene from alligator; ssGE1, single sgRNA-based genome editing coupled with the pUC57_*As-Cath* donor targeting the *mc4r* locus; ssGE2, single sgRNA-based genome editing coupled with the pUC57_*Cec* donor targeting the *mstn1* site.

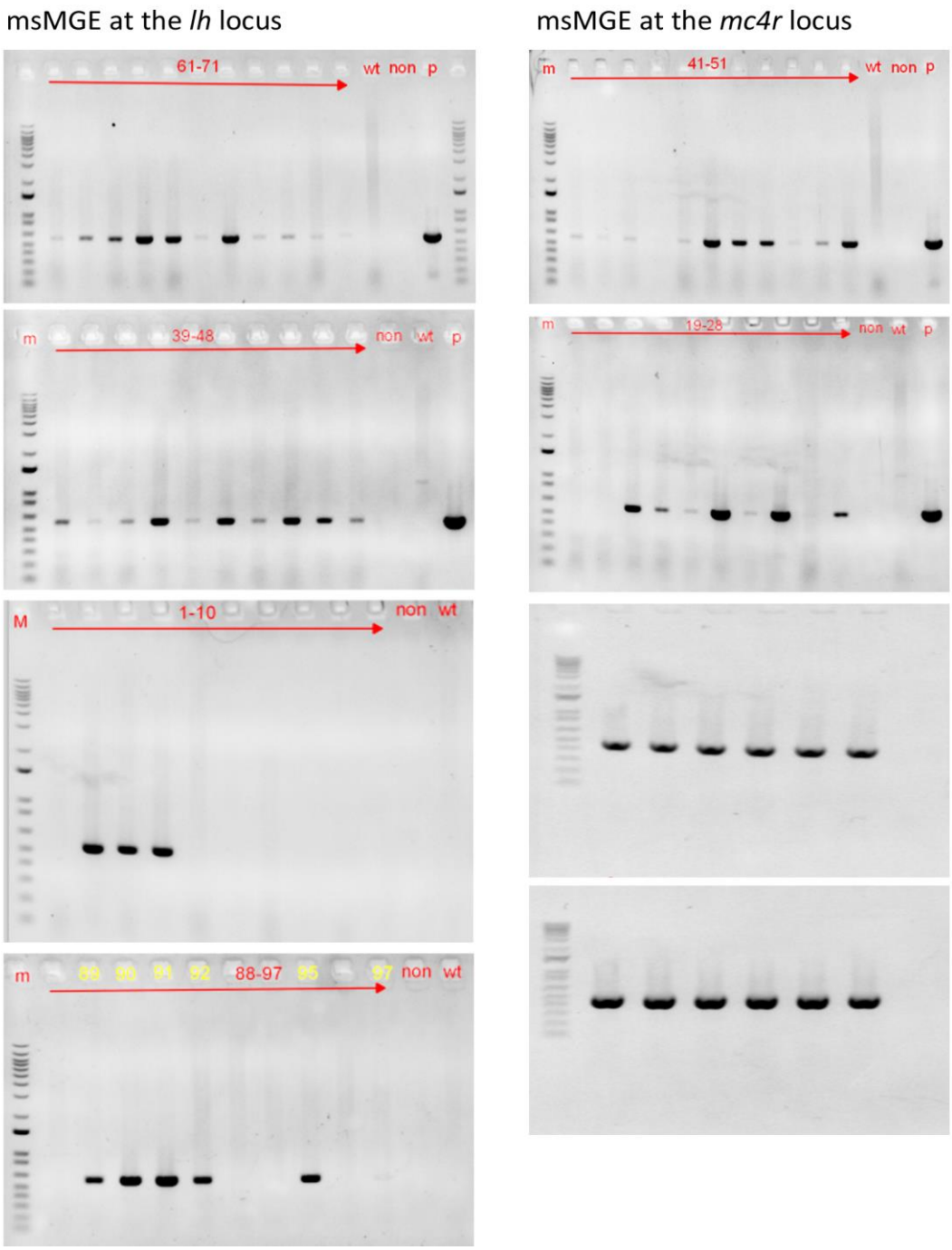


Figure S17. Full-length gel electrophoresis images to determine the *As-Cath* transgenic positive channel catfish (*Ictalurus punctatus*) in the msMGE strategy. *As-Cath*, the cathelicidin gene from alligator; msMGE, multi-sgRNA-based multiplex genome editing coupled with the mixture of dsDNA_*As-Cath*, pUC57_*As-Cath* and pUC57_*Cec* donors targeting the *lh*, *mc4r*, *mstn1* and *mstn2* sites in parallel.

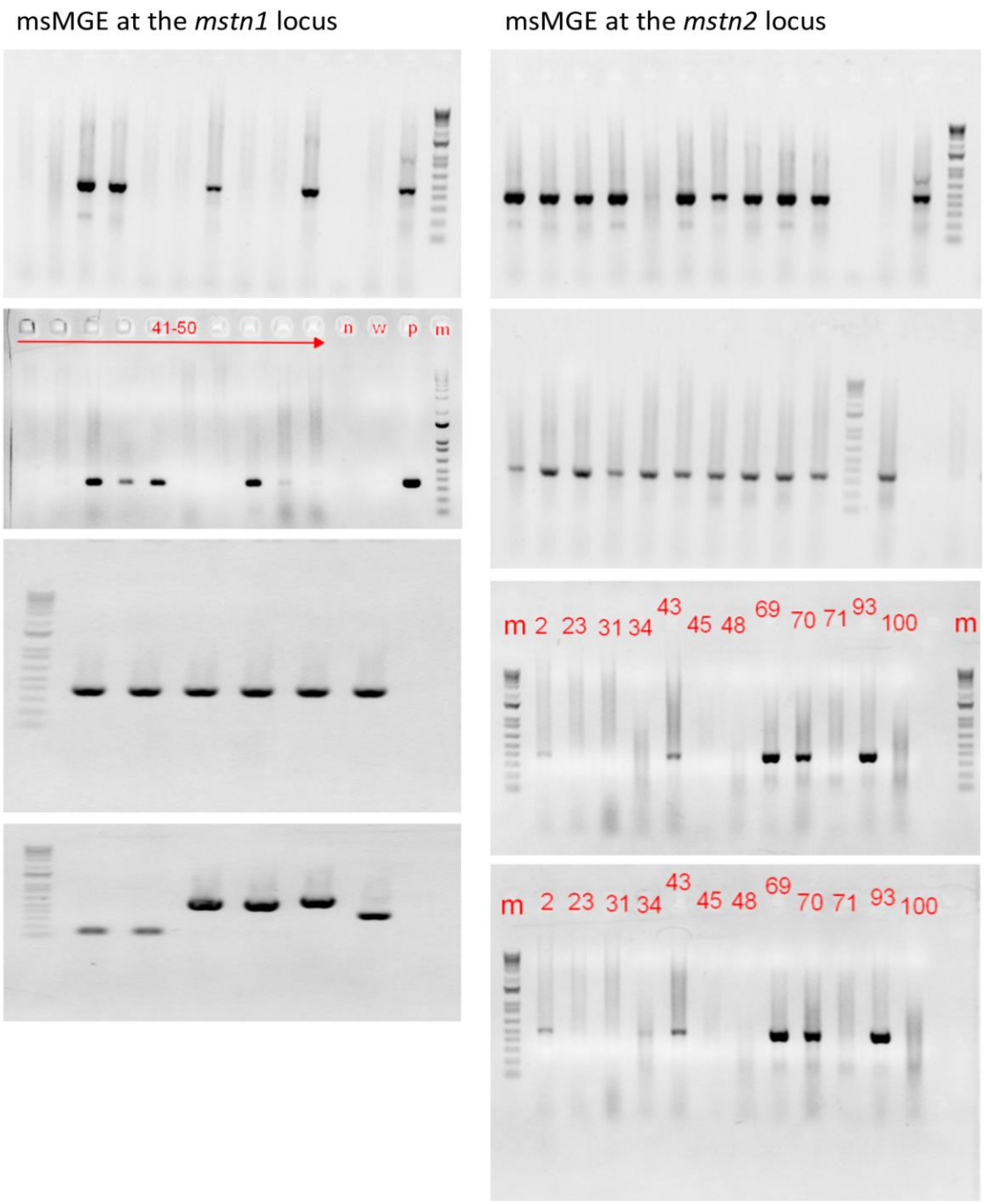


Figure S18. Full-length gel electrophoresis images to determine the *Cec* transgenic positive channel catfish (*Ictalurus punctatus*) in the msMGE strategy. *Cec*, the cecropin gene from moth; msMGE, multi-sgRNA-based multiplex genome editing coupled with the mixture of dsDNA_*As-Cath*, pUC57_*As-Cath* and pUC57_*Cec* donors targeting the *lh*, *mc4r*, *mstn1* and *mstn2* sites in parallel.

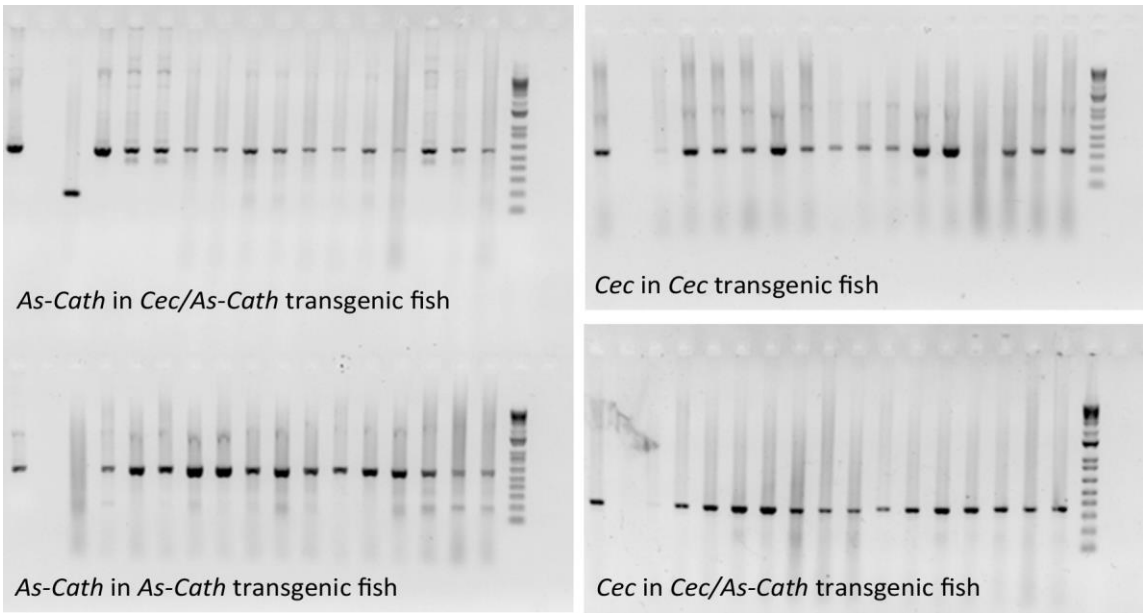


Figure S19. Full-length gel electrophoresis images to determine the mosaicism in transgenic channel catfish (*Ictalurus punctatus*) by detecting the transgene in fourteen tissues from three individuals (#2, #4 and #21).

Appendix 3: R/SAS codes for logistic model construction

R codes for multi-collinearity test:

```
####multi-collinearity test of gene expression#####
#######read the data########
alldata <- read.table("E:/AuburnUniversity/Myproject/Ph.D. research papers/2-cecF1+cath_cathF1_cec in
channels/data_gene_expression.csv",sep=',', header = T)
########Data Structure and Peculiarity######
gene <- alldata[,c(4:13)]
######correlation plot of traits and gene####
library(corrplot)
b <-cor(gene);b
corrplot(b,method ="color",addCoef.col="black",  order = "AOE",number.cex=0.7)
```

SAS codes for logistic regression analysis:

```
/***SAS codes for logistic regression analysis *Start Here***/
/****data input****/
filename gene "/home/u42944627/STAT6110/data_gene_expression.csv";

proc import datafile = gene out = gene_expression
    dbms = csv replace;
run;
proc print data = gene_expression; run;

/****A full multiple logit model****/
proc logistic data = gene_expression descending;
  model outcome = BW Sex Cath Cec CCL3 LEAP2 H2A UBI BPI TCP Catpd NK_lysin;
run;

/****Removed non-significant predictors****/
/****Refined Model: A reduced multiple logit model****/
proc logistic data = gene_expression descending;
  class BW Sex / param = ref;
  model outcome = Cath Cec LEAP2 TCP NK_lysin;
run;

/****Removed non-significant predictors****/
/****Refined Model: A reduced multiple logit model****/
proc logistic data = gene_expression descending;
  class BW Sex / param = ref;
  model outcome = Cath Cec TCP NK_lysin;
run;

/****Logit Interactions were added among predictors****/
proc logistic data = gene_expression descending;
```

```
  class BW Sex / param = ref;
  model outcome = Cath|Cec|TCP|NK_lysin;
run;

/****Assess Model Adequacy***/
/***Hosmer-Lemeshow test***/
proc logistic data = gene_expression descending;
  class BW Sex / param = ref;
  model outcome = Cath Cec TCP NK_lysin / lackfit;
run;

/****ROC curve***/
proc logistic data = gene_expression descending plots = roc;
  class BW Sex / param = ref;
  model outcome = Cath Cec TCP NK_lysin;
run;
```

Appendix 4: Different constructs of CRISPR/Cas9 system (different methods)
Knock in *As-Cath* at the *lh* locus in channel catfish (2H2OP *vs*. dsDNA)

> *lh* gene sequence (5'-3')

exon1 61 bp — intron1 136 bp

GACAGCAATCCACTGAGCGATCACAGCAAAATCTCTAAAGTAAGGACAGTAATGTGATAAGGTGTGATTTAATTA

LH-F →

AATGTATAAATATTACATCCAAAATGTACATAAAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTGCTT

exon2 184 bp

GATTATTTCAGGTATGTGTCGTATTTTTCATGTGTTTTATTACAGG**ATG**TCAGTGCCAGCTTCCTCTTTTCTTCTCCTG

start codon

TGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCT

GTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGCCATCTGCAGC**GGG**CACTGCTTCACCAAGGTG

intron2 85 bp — Target sequence of sgRNA1 — PAM

AATATATATTTTACTTATTTTACTTATTTTACAGATATCATCTGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTA

exon3 365 bp

CAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGTACCTACAGGGACGTTCGCTATGAAA

CCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATCCTGTCGCACTAAGCTGCGAGTGCAG

← LH-R

CCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAATCCGGATTTCTGTATGACACAGAAAGAGTAC

ATCCTGGATTAC**TGA**ACCTCTGCCAACGTGCATGCAGAAGTCTCAGTCTAACTACAACGAGTCCATTTACTCCAAGA

stop codon

ATGATCAGAGTCTGACCTTGTTGTAAGAGCTACTTGCAAAGTACAATATACAACATGACTTAA

> full As-Cath_CDS_sequence (5'-3')

CDS1 198 bp

ATGCAGACCTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGGGGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCG

start code

ACCCACCACAGCTCACGCCGACCTACGCCCAGGCCCTGGCCACGGCCGTCGACGTCTACAACCAAGGGCCCGGCG

CDS2 108 bp

TGGACTTCGCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGACGACTGGGACGCGAGCACGGATCCCCTGCGGCAG

CTGGAGTTCACCCTGAAGGAGACCGAGTGCCCCGTGGGCGAGGACCAGCCCCTGGACCAGTGCGACTTCAAGGA

CDS3 84

TGGCGGGGCGGTGCTGGACTGCACAGGGACCTTCTCCTGCTCCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCA

CDS4 129 bp

ACCCGCCGAGCCCCTGCCTGACCGCGTC**CGCAGGGGTCTATTCAAGAAGCTGAGGAGGAAAATCAAGAAGGGC**

CDS sequence for mature peptide (114 bp)

TTTAAGAAAATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGTGTCTCCATCCCACTCGCAGGAAGGCGGTGA

stop codon

#1 ssODN1-As-Cath-ssODN2 construct (2H2OP):

> UBI-full As-Cath_CDS_pUC57-Mini_sequence (donor plasmid)

GACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGT

Backbone of pUC57_mini vector 1827

ATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTT

CCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAA

AGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAG

TTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTG

ACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAG

AAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGG
CCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAAC
TCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGC
AATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGG
ATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTG
GAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTA
TCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTA
AGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGA
TCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGAC
CCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACC
ACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGA
GCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTA
CATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTC
AAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGC
GAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAG
GCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCT
Target sequence of sgRNA2 PAM
GGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCG
GAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCT
TTCCGATACCAGCAAAGTTCTAGAATTTGTCGAAACATTTATGTTATATATTTCCTGAAAAAAATTCTGAGTAAGTT
CTTAAGTGTTATTGCCAGCAACATAAACAACAGACGGCAAAATGAATAAATGATAACAAAGCAGTAGGCTTAAAT
UBI promoter 1438 bp
AAACCTAATTTTTATAGGCTGTTCTCTACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCCCTTTCATTC
ATATTTCAAGAAAATTGGTTCAGAAGATCTGGATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCAGTGACCTG
ATTCCAAATGACTAGGCCTAATCCAGAAATTAGATGACTGTCAACATAAAAAGGCACAGCACTCACTAGCTGCCCT
ATATATTTTATTATATTTTACATATATTATTTTATTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAGAAAACTGC
TTACAACAGCTAACCTGTACTACCTCAGGCTCAGGGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTTAACCATGC
ATGCTTGTTTTCAACTATGGCAACACAGTCACATGGGACATTACAGAAATGATTTGTCGATGACATGCGACTTTTCT
TTAATAAAGCGCAAAGATCCCAAAAAGCAAACTTTTAACAAAAATCATATAATTATATTTTCAATCCAGCTTTGTAG
CAACTTTGTGCTGCTGTTCACTCAGCAACAGATAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTGCCATCTGTTT
CACACATTGCGTTCTATATATAAGTGTGCTGGTTGACACGACACTGTATAAGGCCTAGGCTAAAACACAAACAATG
TAGAATGACACTGTGTTTTTTTTGTAAACAAATGTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCTCCTGTCATG
TATTTGCTATATTCAAATGTTAAACCCGTGCAGAATAGAACATATACAAAAAAAAACAACACAACACATTTTTAAAC
ATTATTAAATATCAAGTATTGCTGGCAGTTCTGTTTCTGTTTTACAGTACCCTTTGCCACAGTTCTCCGCTTTTCCTG

GTCCAGATTCCACAAGTCTGATTCACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGAC

TGTCCTGCGAGACCGGCCCATTCCAGCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATT

GATACAAAACATGTAATTCCTAGAACATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTT

GACTGCGGAAAGATCTGGTCATGTGACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAGTAGGACTCT

CGACCATCGGCAGATTTTTCGAAGAAGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGCAGACCTGCTGGGT

start codon

CATCCTGCTCCTGCCCCTGCTCGGGGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCGACCCACCACAGCTCACG

Full CDS of As-CATH2 519 bp

CCGACCTACGCCCAGGCCCTGGCCACGGCCGTCGACGTCTACAACCAAGGGCCCGGCGTGGACTTCGCCTTCCGG

CTCCTGGAGGCAGAGTCCCGGGACGACTGGGACGCGAGCACGGATCCCCTGCGGCAGCTGGAGTTCACCCTGAA

GGAGACCGAGTGCCCCGTGGGCGAGGACCAGCCCCTGGACCAGTGCGACTTCAAGGATGGCGGGGCGGTGCTG

GACTGCACAGGGACCTTCTCCTGCTCCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCAACCCGCCGAGCCCCTGC

CTGACCGCGTC**CGCAGGGGTCTATTCAAGAAGCTGAGGAGGAAAATCAAGAAGGGCTTTAAGAAAATCTTCAA**

GCGCCTGCCTCCCGTCGGTGTCGGTGTCTCCATCCCACTCGCAGGAAGGCGGTGAgatccagacatgataagatacattga

stop codon

tgagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgcaata

PolyA tail 135 bp

aacaagttATCATTGGGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAA

TACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTAT

Backbone of pUC57_mini vector 441

GAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACG

CTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGG

TAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCG

GTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGAC

> putative positive fish sequence after insertion (5’-3’)

TTAAGTCATGTTGTATATTGTACTTTGCAAGTAGCTCTTACAACAAGGTCAGACTCTGATCATTCTTGGAGTAAATG

GACTCGTTGTAGTTAGACTGAGACTTCTGCATGCACGTTGGCAGAGGTTCAGTAATCCAGGATGTACTCTTTCTGT

GTCATACAGAAATCCGGATTCAGGCTCTCGATGGTACAGTCCGAGGTGTCCATGGTGCACAGGCTGCACTCGCAG

CTTAGTGCGACAGGATATGTGACGTGAGGATCCACACCGGGCCGACAGTCAGGCAGGCGTACGGTTTCATAGCG

AACGTCCCTGTAGGTACACACATGCTGATAGATGGAAGAGAACGGGCTCTTGTACACAGGTTCCTGTAGAAAACA

ssODN1-F

TTTATTTGTAAACCATTTTTCATACTCAGATGATATCTGTAAAATAAGTAAAATAAGTAAAATATATATTCACCTTGG

ssODN1 80 bp

TGAAGCAGTGCCCGCTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTG

GTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGG

ssODN1-R

AGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTT

TCCGATACCAGCAAAGTTCTAGAATTTGTCGAAACATTTATGTTATATATTTCCTGAAAAAAATTCTGAGTAAGTTC

TTAAGTGTTATTGCCAGCAACATAAACAACAGACGGCAAAATGAATAAATGATAACAAAGCAGTAGGCTTAAATA

AACCTAATTTTTATAGGCTGTTCTCTACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCCCTTTCATTCA

TATTTCAAGAAAATTGGTTCAGAAGATCTGGATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCAGTGACCTGA

TTCCAAATGACTAGGCCTAATCCAGAAATTAGATGACTGTCAACATAAAAAGGCACAGCACTCACTAGCTGCCCTA

TATATTTTATTATATTTTACATATATTATTTTATTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAGAAAACTGCT

TACAACAGCTAACCTGTACTACCTCAGGCTCAGGGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTTAACCATGCA

TGCTTGTTTTCAACTATGGCAACACAGTCACATGGGACATTACAGAAATGATTTGTCGATGACATGCGACTTTTCTT

TAATAAAGCGCAAAGATCCCAAAAAGCAAACTTTTAACAAAAATCATATAATTATATTTTCAATCCAGCTTTGTAGC

AACTTTGTGCTGCTGTTCACTCAGCAACAGATAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTGCCATCTGTTTC

ACACATTGCGTTCTATATATAAGTGTGCTGGTTGACACGACACTGTATAAGGCCTAGGCTAAAACACAAACAATGT

AGAATGACACTGTGTTTTTTTTGTAAACAAATGTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCTCCTGTCATGT

ATTTGCTATATTCAAATGTTAAACCCGTGCAGAATAGAACATATACAAAAAAAAACAACACAACACATTTTTAAACA

TTATTAAATATCAAGTATTGCTGGCAGTTCTGTTTCTGTTTTACAGTACCCTTTGCCACAGTTCTCCGCTTTTCCTGGT

Prom1-F →

CCAGATTCCACAAGTCTGATTCACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGACTG

TCCTGCGAGACCGGCCCATTCCAGCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATTGA

TACAAAACATGTAATTCCTAGAACATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGA

CTGCGGAAAGATCTGGTCATGTGACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAGTAGGACTCTCG

ACCATCGGCAGATTTTTCGAAGAAGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGCAGACCTGCTGGGTCA

Cath1-F →

start codon

TCCTGCTCCTGCCCCTGCTCGGGGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCGACCCACCACAGCTCACGCC

GACCTACGCCCAGGCCCTGGCCACGGCCGTCGACGTCTACAACCAAGGGCCCGGCGTGGACTTCGCCTTCCGGCT

← Prom1-R

CCTGGAGGCAGAGTCCCGGGACGACTGGGACGCGAGCACGGATCCCCTGCGGCAGCTGGAGTTCACCCTGAAGG

AGACCGAGTGCCCCGTGGGCGAGGACCAGCCCCTGGACCAGTGCGACTTCAAGGATGGCGGGGCGGTGCTGGA

CTGCACAGGGACCTTCTCCTGCTCCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCAACCCGCCGAGCCCCTGCCT

GACCGCGTCCGCAGGGGTCTATTCAAGAAGCTGAGGAGGAAAATCAAGAAGGGCTTTAAGAAAATCTTCAAGC

qPCR-F →

stop codon

GCCTGCCTCCCGTCGGTGTCGGTGTCTCCATCCCACTCGCAGGAAGGCGGTGAgatccagacatgataagatacattgatga

← qPCR-R

← Cath1-R

gtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgcaataaa

caagttATCATTGGGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATA

Backbone of pUC57_mini vector 1579

CATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGA

GTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCT

GGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTA

AGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGT

ATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTAC

TCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGT

GATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATG

GGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACC

ACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAAC

AATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTA

TTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCT

CCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAG

GTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATT

TTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCC

ACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTG

CAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAA

CTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTC

TGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTA

ssODN2-F →

CCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAG

CCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCC

GAAGGGAGAAAGGCGGACAGGTATCCGGTGCAGATGGCGGTTTGAAACACAAGGCATTTCGGGCAGCCATCTTT

ssODN2 80 bp

CTCCACA]GAAACAGTCTCATTAACAGGTTCGCAGTGTGGCAGAATGTAGCTTTGAGCGGGGGAGAAGGAGTTCA

← ssODN2-R

TCAAGAAACACAGGAGAAGAAAAGAGGAAGCTGGCACTGACATCCTGTAATAAAACACATGAAAAATACGACAC

ATACCTGAAATAATCAAGCAAGGACAAATTATCTTGAGTTCAGACTACACTTTGCACTTTTATGTACATTTTGGATG

TAATATTTATACATTTAATTAAATCACACCTTATCACATTACTGTCCTTACTTTAGAGATTTTGCTGTGATCGCTCAGT

GGATTGCTGTC

#2 HA1-As-Cath-HA2 construct (dsDNA):

> *lh* gene sequence (5'-3')

ACCTGTCCAAACAGGAAGCTCATTAATCCTTTTGGCTGAGGGCAAGATGAACTGCAGGTTTCTAGTGTCATGGTGT

Upstream of *lh* gene 200 bp

ACCTTTGAAAAAGTAGCCGTGATTATGTCAGAGCCCTGAGCTGTAATTTATTTGAGATTAGTGCGATAAGGCCACG

exon1 61 bp LHqPCR-F →

TTTCGATTCTGCGACACTATATAAACATGTTAAACTCTTGTAGGAACAGACAGCAATCCACTGAGCGATCACAGCA

intron1 136 bp LH-F →

AAATCTCTAAAGTAAGGACAGTAATGTGATAAGGTGTGATTTAATTAAATGTATAAATATTACATCCAAAATGTAC

ATAAAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTGCTTGATTATTTCAGGTATGTGTCGTATTTTTC

exon2 184 bp

ATGTGTTTTATTACAGGATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGC

TCAAAGCTACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGC

← LHqPCR-R intron2 85 bp

CTTGTGTTTCAAACCGCCATCTGCAGCGGGCACTGCTTCACCAAGGTGAATATATATTTTACTTATTTTACTTATTTT

Target sequence of sgRNA1 PAM

ACAGATATCATCTGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTACAGGAACCTGTGTACAAGAGCCCGTTC

TCTTCCATCTATCAGCATGTGTGTACCTACAGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCG

GTGTGGATCCTCACGTCACATATCCTGTCGCACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTG

exon3 365 bp

← LH-R

TACCATCGAGAGCCTGAATCCGGATTTCTGTATGACACAGAAAGAGTACATCCTGGATTAC**TGA**ACCTCTGCCAAC

GTGCATGCAGAAGTCTCAGTCTAACTACAACGAGTCCATTTACTCCAAGAATGATCAGAGTCTGACCTTGTTGTAA

GAGCTACTTGCAAAGTACAATATACAACATGACTTAAGACTTGCTTGACTAAGAATTTTCTTTTCTTGTATGAACAG

Downstream of *lh* gene 200 bp

GGAAACATTTTCAATCAATCACACATAAGGGTAAAAGAGAGTCCTCGAAAAACACAGAGATCATAATCGCTGCAC

TCTCGTATTTTAAAATCCACTGATACGCATCATTTAACATAATTTACATCAATTTACTTGATTTGTAAAGTCCTGGTT

GCTAGTA

> HA1_UBI-full As-CATH2_CDS_HA2 sequence (donor linear dsDNA)

GGTGTGATTTAATTAAATGTATAAATATTACATCCAAAATGTACATAAAAGTGCAAAGTGTAGTCTGAACTCAAGATAAT

Left homologous arm (HA1) 300 bp

TTGTCCTTGCTTGATTATTTCAGGTATGTGTCGTATTTTTCATGTGTTTTATTACAGGATGTCAGTGCCAGCTTCCTCTTT

TCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGCCACACTGCGAACCTGTTAATGAG

ACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGCCATCTGCACCAGCAAAGTTCTAGA

UBI promoter 1438 bp

ATTTGTCGAAACATTTATGTTATATATTTCCTGAAAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATA

AACAACAGACGGCAAAATGAATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATAGGCTGTTCTC

TACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCCCTTTCATTCATATTTCAAGAAAATTGGTTCAGAA

GATCTGGATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCAGTGACCTGATTCCAAATGACTAGGCCTAATCCA

GAAATTAGATGACTGTCAACATAAAAAGGCACAGCACTCACTAGCTGCCCTATATATTTTATTATATTTTACATATA

TTATTTTATTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAGAAAACTGCTTACAACAGCTAACCTGTACTACCT

CAGGCTCAGGGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTTAACCATGCATGCTTGTTTTCAACTATGGCAACA

CAGTCACATGGGACATTACAGAAATGATTTGTCGATGACATGCGACTTTTCTTTAATAAAGCGCAAAGATCCCAAA

AAGCAAACTTTTAACAAAAATCATATAATTATATTTTCAATCCAGCTTTGTAGCAACTTTGTGCTGCTGTTCACTCAG

CAACAGATAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTGCCATCTGTTTCACACATTGCGTTCTATATATAAGT

GTGCTGGTTGACACGACACTGTATAAGGCCTAGGCTAAAACACAAACAATGTAGAATGACACTGTGTTTTTTTTGT

AAACAAATGTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCTCCTGTCATGTATTTGCTATATTCAAATGTTAAAC

CCGTGCAGAATAGAACATATACAAAAAAAAACAACACAACACATTTTTAAACATTATTAAATATCAAGTATTGCTG

GCAGTTCTGTTTCTGTTTTACAGTACCCTTTGCCACAGTTCTCCGCTTTTCCTGGTCCAGATTCCACAAGTCTGATTC

ACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGACTGTCCTGCGAGACCGGCCCATTCC

AGCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATTGATACAAAACATGTAATTCCTAGA

ACATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGAAAGATCTGGTCATGT

GACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAGTAGGACTCTCGACCATCGGCAGATTTTTCGAAG

AAGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGCAGACCTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGG

start codon

GGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCGACCCACCACAGCTCACGCCGACCTACGCCCAGGCCCTGGC

full AS-Cath CDS 519 bp

CACGGCCGTCGACGTCTACAACCAAGGGCCCGGCGTGGACTTCGCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGA

CGACTGGGACGCGAGCACGGATCCCCTGCGGCAGCTGGAGTTCACCCTGAAGGAGACCGAGTGCCCCGTGGGCG

AGGACCAGCCCCTGGACCAGTGCGACTTCAAGGATGGCGGGGCGGTGCTGGACTGCACAGGGACCTTCTCCTGCT

CCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCAACCCGCCGAGCCCCTGCCTGACCGCGTC**CGCAGGGGTCTATT**

CAAGAAGCTGAGGAGGAAAATCAAGAAGGGCTTTAAGAAAATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGT

GTCTCCATCCCACTCGCAGGAAGGCGGTGAgatccagacatgataagatacattgatgagtttggacaaaccacaactagaatgcagt

stop codon

PolyA tail 135 bp

gaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgcaataaacaagttAGC**GGG**CACTGCTTCACC

AAGGTGAATATATATTTTACTTATTTTACTTATTTTACAGATATCATCTGAGTATGAAAAATGGTTTACAAATAAATGTTTT

Right homologous arm (HA2) 300

CTACAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGTACCTACAGGGACGTTCGCTATGAA

ACCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATCCTGTCGCACTAAGCTGCGAGTGCAGC

CTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAA

> putative positive fish sequence after insertion (5'-3')

Upstream of *lh* gene 200 bp

ACCTGTCCAAACAGGAAGCTCATTAATCCTTTTGGCTGAGGGCAAGATGAACTGCAGGTTTCTAGTGTCATGGTGT

HA1-F

ACCTTTGAAAAAGTAGCCGTGATTATGTCAGAGCCCTGAGCTGTAATTTATTTGAGATTAGTGCGATAAGGCCACG

Partial sequence of *lh* extron1 60 bp

TTTCGATTCTGCGACACTATATAAACATGTTAAACTCTTGTAGGAACAGACAGCAATCCACTGAGCGATCACAGCA

AAATCTCTAAAGTAAGGACAGTAATGTGATAAGGTGTGATTTAATTAAATGTATAAATATTACATCCAAAATGTACATA

HA1 300 bp

AAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTGCTTGATTATTTCAGGTATGTGTCGTATTTTTCATGTGT

TTTATTACAGGATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGC

TACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTT

CAAACCGCCATCTGCACCAGCAAAGTTCTAGAATTTGTCGAAACATTTATGTTATATATTTCCTGAAAAAAATTCTG

UBI promoter 1438

HA1-R

AGTAAGTTCTTAAGTGTTATTGCCAGCAACATAAACAACAGACGGCAAAATGAATAAATGATAACAAAGCAGTAG

GCTTAAATAAACCTAATTTTTATAGGCTGTTCTCTACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCC

CTTTCATTCATATTTCAAGAAAATTGGTTCAGAAGATCTGGATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCA

GTGACCTGATTCCAAATGACTAGGCCTAATCCAGAAATTAGATGACTGTCAACATAAAAAGGCACAGCACTCACTA

GCTGCCCTATATATTTTATTATATTTTACATATATTATTTTATTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAG

AAAACTGCTTACAACAGCTAACCTGTACTACCTCAGGCTCAGGGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTT

AACCATGCATGCTTGTTTTCAACTATGGCAACACAGTCACATGGGACATTACAGAAATGATTTGTCGATGACATGC

GACTTTTCTTTAATAAAGCGCAAAGATCCCAAAAAGCAAACTTTTAACAAAAATCATATAATTATATTTTCAATCCA

GCTTTGTAGCAACTTTGTGCTGCTGTTCACTCAGCAACAGATAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTG

CCATCTGTTTCACACATTGCGTTCTATATATAAGTGTGCTGGTTGACACGACACTGTATAAGGCCTAGGCTAAAACA

CAAACAATGTAGAATGACACTGTGTTTTTTTTGTAAACAAATGTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCT

CCTGTCATGTATTTGCTATATTCAAATGTTAAACCCGTGCAGAATAGAACATATACAAAAAAAAACAACACAACAC

ATTTTTAAACATTATTAAATATCAAGTATTGCTGGCAGTTCTGTTTCTGTTTTACAGTACCCTTTGCCACAGTTCTCCG

Prom2-F →

CTTTTCCTGGTCCAGATTCCACAAGTCTGATTCACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTG

CTTGTAGACTGTCCTGCGAGACCGGCCCATTCCAGCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACAT

TTAAATTATTGATACAAAACATGTAATTCCTAGAACATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTT

TTAGGATTTGACTGCGGAAAGATCTGGTCATGTGACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAG

TAGGACTCTCGACCATCGGCAGATTTTTCGAAGAAGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGCAGAC

Cath2-F →

CTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGGGGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCGACCCACCA

full AS-Cath CDS 519 bp

CAGCTCACGCCGACCTACGCCCAGGCCCTGGCCACGGCCGTCGACGTCTACAACCAAGGGCCCGGCGTGGACTTC

← Prom2-R

GCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGACGACTGGGACGCGAGCACGGATCCCCTGCGGCAGCTGGAGTT

CACCCTGAAGGAGACCGAGTGCCCCGTGGGCGAGGACCAGCCCCTGGACCAGTGCGACTTCAAGGATGGCGGGG

CGGTGCTGGACTGCACAGGGACCTTCTCCTGCTCCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCAACCCGCCGA

GCCCCTGCCTGACCGCGTC**CGCAGGGGTCTATTCAAGAAGCTGAGGAGGAAAATCAAGAAGGGCTTTAAGAAA**

ATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGTGTCTCCATCCCACTCGCAGGAAGGCGGTGAgatccagacatgataa

gatacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattata

HA2-F → ← Cath2-R

agctgcaataaacaagttAGC**GGG**CACTGCTTCACCAAGGTGAATATATATTTTACTTATTTTACTTATTTTACAGATATCATC

TGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTACAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAG

HA2 300 bp

CATGTGTGTACCTACAGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTC

ACATATCCTGTCGCACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAA

TACTGAACCTCTGCCAACGTGCATGCAGAAGTCTCAGTCTAACTACAACGAGTCCATTTACTCCAAGAATGATCAG

Partial sequence of *lh* extron3 131 bp

AGTCTGACCTTGTTGTAAGAGCTACTTGCAAAGTACAATATACAACATGACTTAAGACTTGCTTGACTAAGAATTTT

CTTTTCTTGTATGAACAGGGAAACATTTTCAATCAATCACACATAAGGGTAAAAGAGAGTCCTCGAAAAACACAGA

← HA2-R Downstream of *lh* gene 200 bp

GATCATAATCGCTGCACTCTCGTATTTTAAAATCCACTGATACGCATCATTTAACATAATTTACATCAATTTACTTGA

TTTGTAAAGTCCTGGTTGCTAGTA

Appendix 5: Different constructs of CRISPR/Cas9 system (different methods)
Knock in *As-Cath* at the *lh* locus in blue catfish (dcPlasmid *vs*. dsDNA)

> *lh* gene sequence (5'-3')

exon1 61 bp — intron1 136 bp

GACAGCAATCCACTGAGCGATCACAGCAAAATCTCTAAAGTAAGGACAGTAATGTGATAAGGTGTGATTTAATTA

LH-F →

AATGTATAAATATTACATCCAAAATGTACATAAAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTGCTT

exon2 184 bp

GATTATTTCAGGTATGTGTCGTATTTTTCATGTGTTTTATTACAGG**ATG**TCAGTGCCAGCTTCCTCTTTTCTTCTCCTG

start codon

TGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCT

GTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGCCATCTGCAGC**GGG**CACTGCTTCACCAAGGTG

intron2 85 bp — Target sequence of sgRNA — PAM

AATATATATTTTACTTATTTTACTTATTTTACAGATATCATCTGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTA

exon3 365 bp

CAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGTACCTACAGGGACGTTCGCTATGAAA

CCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATCCTGTCGCACTAAGCTGCGAGTGCAG

← LH-R

CCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAATCCGGATTTCTGTATGACACAGAAAGAGTAC

ATCCTGGATTAC**TGA**ACCTCTGCCAACGTGCATGCAGAAGTCTCAGTCTAACTACAACGAGTCCATTTACTCCAAGA

stop codon

ATGATCAGAGTCTGACCTTGTTGTAAGAGCTACTTGCAAAGTACAATATACAACATGACTTAA

> full As-Cath_CDS_sequence (5'-3')

CDS1 198 bp

ATGCAGACCTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGGGGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCG

start codon

ACCCACCACAGCTCACGCCGACCTACGCCCAGGCCCTGGCCACGGCCGTCGACGTCTACAACCAAGGGCCCGGCG

CDS2 108 bp

TGGACTTCGCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGACGACTGGGACGCGAGCACGGATCCCCTGCGGCAG

CTGGAGTTCACCCTGAAGGAGACCGAGTGCCCCGTGGGCGAGGACCAGCCCCTGGACCAGTGCGACTTCAAGGA

CDS3 84

TGGCGGGGCGGTGCTGGACTGCACAGGGACCTTCTCCTGCTCCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCA

CDS4 129 bp

ACCCGCCGAGCCCCTGCCTGACCGCGTC**CGCAGGGGTCTATTCAAGAAGCTGAGGAGGAAAATCAAGAAGGGC**

CDS sequence for mature peptide (114 bp)

TTTAAGAAAATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGTGTCTCCATCCCACTCGCAGGAAGGCGGTGA

stop codon

Different CRISPR/Cas9-mediated systems at *lh* locus:

> *lh* gene + upstream/downstream sequence(5'-3')

ACCTGTCCAAACAGGAAGCTCATTAATCCTTTTGGCTGAGGGCAAGATGAACTGCAGGTTTCTAGTGTCATGGTGT
Upstream of *lh* gene 200 bp
ACCTTTGAAAAAGTAGCCGTGATTATGTCAGAGCCCTGAGCTGTAATTTATTTGAGATTAGTGCGATAAGGCCACG
exon1 61 bp LHqPCR-F →
TTTCGATTCTGCGACACTATATAAACATGTTAAACTCTTGTAGGAACAGACAGCAATCCACTGAGCGATCACAGCA
intron1 136 bp LH-F →
AAATCTCTAAAGTAAGGACAGTAATGTGATAAGGTGTGATTTAATTAAATGTATAAATATTACATCCAAAATGTAC
ATAAAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTGCTTGATTATTTCAGGTATGTGTCGTATTTTTC
exon2 184 bp
ATGTGTTTTATTACAGG**ATG**TCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGC
TCAAAGCTACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGC
← LHqPCR-R intron2 85 bp
CTTGTGTTTCAAACCGCCATCTGCAGC**GGG**CACTGCTTCACCAAGGTGAATATATATTTTACTTATTTTACTTATTTT
Target sequence of sgRNA PAM
ACAGATATCATCTGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTACAGGAACCTGTGTACAAGAGCCCGTTC
exon3 365 bp
TCTTCCATCTATCAGCATGTGTGTACCTACAGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCG
← LH-R
GTGTGGATCCTCACGTCACATATCCTGTCGCACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTG
TACCATCGAGAGCCTGAATCCGGATTTCTGTATGACACAGAAAGAGTACATCCTGGATTAC**TGA**ACCTCTGCCAAC
GTGCATGCAGAAGTCTCAGTCTAACTACAACGAGTCCATTTACTCCAAGAATGATCAGAGTCTGACCTTGTTGTAA
Downstream of *lh* gene 200 bp
GAGCTACTTGCAAAGTACAATATACAACATGACTTAAGACTTGCTTGACTAAGAATTTTCTTTTCTTGTATGAACAG
GGAAACATTTTCAATCAATCACACATAAGGGTAAAAGAGAGTCCTCGAAAAACACAGAGATCATAATCGCTGCAC
TCTCGTATTTTAAAATCCACTGATACGCATCATTTAACATAATTTACATCAATTTACTTGATTTGTAAAGTCCTGGTT
GCTAGTA

#1 > HA1_UBI-full As-Cath_CDS_HA2 (linear dsDNA donor: dsDNA_As-Cath)

GGTGTGATTTAATTAAATGTATAAATATTACATCCAAAATGTACATAAAAGTGCAAAGTGTAGTCTGAACTCAAGATAAT
Left homologous arm (HA1) 300 bp
TTGTCCTTGCTTGATTATTTCAGGTATGTGTCGTATTTTTCATGTGTTTTATTACAGGATGTCAGTGCCAGCTTCCTCTTT
TCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGCCACACTGCGAACCTGTTAATGAG
ACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGCCATCTGCACCAGCAAAGTTCTAGA
UBI promoter 1438 bp
ATTTGTCGAAACATTTATGTTATATATTTCCTGAAAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATA
AACAACAGACGGCAAAATGAATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATAGGCTGTTCTC
TACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCCCTTTCATTCATATTTCAAGAAAATTGGTTCAGAA
GATCTGGATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCAGTGACCTGATTCCAAATGACTAGGCCTAATCCA
GAAATTAGATGACTGTCAACATAAAAAGGCACAGCACTCACTAGCTGCCCTATATATTTTATTATATTTTACATATA
TTATTTTATTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAGAAAACTGCTTACAACAGCTAACCTGTACTACCT
CAGGCTCAGGGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTTAACCATGCATGCTTGTTTTCAACTATGGCAACA

CAGTCACATGGGACATTACAGAAATGATTTGTCGATGACATGCGACTTTTCTTTAATAAAGCGCAAAGATCCCAAA

AAGCAAACTTTTAACAAAAATCATATAATTATATTTTCAATCCAGCTTTGTAGCAACTTTGTGCTGCTGTTCACTCAG

CAACAGATAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTGCCATCTGTTTCACACATTGCGTTCTATATATAAGT

GTGCTGGTTGACACGACACTGTATAAGGCCTAGGCTAAAACACAAACAATGTAGAATGACACTGTGTTTTTTTTGT

AAACAAATGTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCTCCTGTCATGTATTTGCTATATTCAAATGTTAAAC

CCGTGCAGAATAGAACATATACAAAAAAAAACAACACAACACATTTTTAAACATTATTAAATATCAAGTATTGCTG

GCAGTTCTGTTTCTGTTTTACAGTACCCTTTGCCACAGTTCTCCGCTTTTCCTGGTCCAGATTCCACAAGTCTGATTC

ACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGACTGTCCTGCGAGACCGGCCCATTCC

AGCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATTGATACAAAACATGTAATTCCTAGA

ACATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGAAAGATCTGGTCATGT

GACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAGTAGGACTCTCGACCATCGGCAGATTTTTCGAAG

AAGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGCAGACCTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGG

start codon

GGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCGACCCACCACAGCTCACGCCGACCTACGCCCAGGCCCTGGC

full As-Cath CDS 519 bp

CACGGCCGTCGACGTCTACAACCAAGGGCCCGGCGTGGACTTCGCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGA

CGACTGGGACGCGAGCACGGATCCCCTGCGGCAGCTGGAGTTCACCCTGAAGGAGACCGAGTGCCCCGTGGGCG

AGGACCAGCCCCTGGACCAGTGCGACTTCAAGGATGGCGGGGCGGTGCTGGACTGCACAGGGACCTTCTCCTGCT

CCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCAACCCGCCGAGCCCCTGCCTGACCGCGTC**CGCAGGGGTCTATT**

CAAGAAGCTGAGGAGGAAAATCAAGAAGGGCTTTAAGAAAATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGT

GTCTCCATCCCACTCGCAGGAAGGCGGTGAgatccagacatgataagatacattgatgagtttggacaaaccacaactagaatgcagt

stop codon **PolyA tail 135 bp**

gaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgcaataaacaagttAGCGGGCACTGCTTCACC

AAGGTGAATATATATTTTACTTATTTTACTTATTTTACAGATATCATCTGAGTATGAAAAATGGTTTACAAATAAATGTTTT

Right homologous arm (HA2) 300

CTACAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGTACCTACAGGGACGTTCGCTATGAA

ACCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATCCTGTCGCACTAAGCTGCGAGTGCAGC

CTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAA

#2 > HA1_UBI-full As-Cath_CDS_HA2_pUC57 (double-cut plasmid donor: pUC57_As-Cath)

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACTGTCACAGCTTGTCTGTAAGC
Backbone of pUC57 vector 415
GGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTAT

GCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAA

ATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCT

ATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCAC

GACGTTGTAAAACGACGGCCAGTGAATTGACGCGTATTGGGATTTCAAACCGCCATCTGCAGCGGGGGTGTGATT
Left homologous arm (HA1) 300 bp **Target sequence of sgRNA** **PAM**
TAATTAAATGTATAAATATTACATCCAAAATGTACATAAAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTG

CTTGATTATTTCAGGTATGTGTCGTATTTTTCATGTGTTTTATTACAGGATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCT

GTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCT

GTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGCCATCTGCACCAGCAAAGTTCTAGAATTTGTCGA

AACATTTATGTTATATATTTCCTGAAAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATAAACAACAG
UBI promoter 1438 bp
ACGGCAAAATGAATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATAGGCTGTTCTCTACAACCC

TCAAACAGTGATTAGTTTTGTACTTATAAACTTGCCCTTTCATTCATATTTCAAGAAAATTGGTTCAGAAGATCTGG

ATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCAGTGACCTGATTCCAAATGACTAGGCCTAATCCAGAAATTA

GATGACTGTCAACATAAAAAGGCACAGCACTCACTAGCTGCCCTATATATTTTATTATATTTTACATATATTATTTTA

TTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAGAAAACTGCTTACAACAGCTAACCTGTACTACCTCAGGCTC

AGGGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTTAACCATGCATGCTTGTTTTCAACTATGGCAACACAGTCAC

ATGGGACATTACAGAAATGATTTGTCGATGACATGCGACTTTTCTTTAATAAAGCGCAAAGATCCCAAAAAGCAAA

CTTTTAACAAAAATCATATAATTATATTTTCAATCCAGCTTTGTAGCAACTTTGTGCTGCTGTTCACTCAGCAACAGA

TAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTGCCATCTGTTTCACACATTGCGTTCTATATATAAGTGTGCTGG

TTGACACGACACTGTATAAGGCCTAGGCTAAAACACAAACAATGTAGAATGACACTGTGTTTTTTTTGTAAACAAA

TGTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCTCCTGTCATGTATTTGCTATATTCAAATGTTAAACCCGTGCA

GAATAGAACATATACAAAAAAAAACAACACAACACATTTTTAAACATTATTAAATATCAAGTATTGCTGGCAGTTCT

GTTTCTGTTTTACAGTACCCTTTGCCACAGTTCTCCGCTTTTCCTGGTCCAGATTCCACAAGTCTGATTCACCAATAG

CAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGACTGTCCTGCGAGACCGGCCCATTCCAGCACATT

CTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATTGATACAAAACATGTAATTCCTAGAACATAACC

ATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGAAAGATCTGGTCATGTGACGTCTC

ATGAACGTCACGGCCCTGGGTTTCTATAAATACAGTAGGACTCTCGACCATCGGCAGATTTTTCGAAGAAGAAGAT

CAGTTTCAGGAGCCGTACTGTTCCGTTATGCAGACCTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGGGGCAGCCA
start codon **full As-Cath CDS (519 bp)**
GCACCGAGCTGCCCACCCCTGGCACCGACCCACCACAGCTCACGCCGACCTACGCCCAGGCCCTGGCCACGGCCG

TCGACGTCTACAACCAAGGGCCCGGCGTGGACTTCGCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGACGACTGGG

ACGCGAGCACGGATCCCCTGCGGCAGCTGGAGTTCACCCTGAAGGAGACCGAGTGCCCCGTGGGCGAGGACCAG

CCCCTGGACCAGTGCGACTTCAAGGATGGCGGGGCGGTGCTGGACTGCACAGGGACCTTCTCCTGCTCCGAGGCC

TCGCTCATGGTGCTGGTCACCTGCCAACCCGCCGAGCCCCTGCCTGACCGCGTC**CGCAGGGGTCTATTCAAGAAG**

CDS sequence for mature peptide (114 bp)

CTGAGGAGGAAAATCAAGAAGGGCTTTAAGAAAATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGTGTCTCCAT

CCCACTCGCAGGAAGGCGGTGAgatccagacatgataagatacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaat

stop code **PolyA tail 135 bp**

gctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgcaataaacaagttAGC**GGG**CACTGCTTCACCAAGGTGAA

TATATATTTTACTTATTTTACTTATTTTACAGATATCATCTGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTACAGGA

Right homologous arm (HA2) 300

ACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGTACCTACAGGGACGTTCGCTATGAAACCGTACGC

CTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATCCTGTCGCACTAAGCTGCGAGTGCAGCCTGTGCACC

PAM

ATGGACACCTCGGACTGTACCATCGAGAGCCTGAA**CCC**GCTGCAGATGGCGGTTTGAAATCCCAATGGCGCGCCGA

Recognition site of sgRNA at anti-sense strand 23 bp

GCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCC

Backbone of pUC57 vector 2256 bp

GGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCC

GCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGT

ATTGGGCGCTGTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC

ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAG

CAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCAC

AAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGC

TCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC

GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAA

CCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC

GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT

GGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAA

AAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATT

ACGCGCA//CACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG

GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAAT

ATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATA

GGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAA

AAATAGGCGTATCACGAGGCCCTTTTGTC

> putative positive fish sequence after insertion (5'-3')

Upstream of *lh* gene 200 bp

ACCTGTCCAAACAGGAAGCTCATTAATCCTTTTGGCTGAGGGCAAGATGAACTGCAGGTTTCTAGTGTCATGGTGT

UBI promoter 1438 HA1-F

ACCTTTGAAAAAGTAGCCGTGATTATGTCAGAGCCCTGAGCTGTAATTTATTTGAGATTAGTGCGATAAGGCCACG

TTTCGATTCTGCGACACTATATAAACATGTTAAACTCTTGTAGGAACAGACAGCAATCCACTGAGCGATCACAGCA

Partial sequence of *lh* extron1 60 bp

AAATCTCTAAAGTAAGGACAGTAATGTGATAAGGTGTGATTTAATTAAATGTATAAATATTACATCCAAAATGTACATA

AAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTGCTTGATTATTTCAGGTATGTGTCGTATTTTTCATGTGT

HA1 300 bp

TTTATTACAGGATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGC

TACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTT

CAAACCGCCATCTGCACCAGCAAAGTTCTAGAATTTGTCGAAACATTTATGTTATATATTTCCTGAAAAAAATTCTG

HA1-R

AGTAAGTTCTTAAGTGTTATTGCCAGCAACATAAACAACAGACGGCAAAATGAATAAATGATAACAAAGCAGTAG

GCTTAAATAAACCTAATTTTTATAGGCTGTTCTCTACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCC

CTTTCATTCATATTTCAAGAAAATTGGTTCAGAAGATCTGGATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCA

GTGACCTGATTCCAAATGACTAGGCCTAATCCAGAAATTAGATGACTGTCAACATAAAAAGGCACAGCACTCACTA

GCTGCCCTATATATTTTATTATATTTTACATATATTATTTTATTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAG

AAAACTGCTTACAACAGCTAACCTGTACTACCTCAGGCTCAGGGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTT

AACCATGCATGCTTGTTTTCAACTATGGCAACACAGTCACATGGGACATTACAGAAATGATTTGTCGATGACATGC

GACTTTTCTTTAATAAAGCGCAAAGATCCCAAAAAGCAAACTTTTAACAAAAATCATATAATTATATTTTCAATCCA

GCTTTGTAGCAACTTTGTGCTGCTGTTCACTCAGCAACAGATAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTG

CCATCTGTTTCACACATTGCGTTCTATATATAAGTGTGCTGGTTGACACGACACTGTATAAGGCCTAGGCTAAAACA

CAAACAATGTAGAATGACACTGTGTTTTTTTTGTAAACAAATGTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCT

CCTGTCATGTATTTGCTATATTCAAATGTTAAACCCGTGCAGAATAGAACATATACAAAAAAAAACAACACAACAC

Prom-F

ATTTTTAAACATTATTAAATATCAAGTATTGCTGGCAGTTCTGTTTCTGTTTTACAGTACCCTTTGCCACAGTTCTCCG

CTTTTCCTGGTCCAGATTCCACAAGTCTGATTCACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTG

CTTGTAGACTGTCCTGCGAGACCGGCCCATTCCAGCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACAT

TTAAATTATTGATACAAAACATGTAATTCCTAGAACATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTT

TTAGGATTTGACTGCGGAAAGATCTGGTCATGTGACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAG

Cath-F

TAGGACTCTCGACCATCGGCAGATTTTTCGAAGAAGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGCAGAC

CTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGGGGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCGACCCACCA

full As-Cath CDS 519 bp

CAGCTCACGCCGACCTACGCCCAGGCCCTGGCCACGGCCGTCGACGTCTACAACCAAGGGCCCGGCGTGGACTTC

Prom-R

GCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGACGACTGGGACGCGAGCACGGATCCCCTGCGGCAGCTGGAGTT

CACCCTGAAGGAGACCGAGTGCCCCGTGGGCGAGGACCAGCCCCTGGACCAGTGCGACTTCAAGGATGGCGGGG

CGGTGCTGGACTGCACAGGGACCTTCTCCTGCTCCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCAACCCGCCGA

GCCCCTGCCTGACCGCGTC**CGCAGGGGTCTATTCAAGAAGCTGAGGAGGAAAATCAAGAAGGGCTTTAAGAAA**

ATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGTGTCTCCATCCCACTCGCAGGAAGGCGGTGAgatccagacatgataa

HA2-F→

gatacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattata

←Cath-R

agctgcaataaacaagttAGC**GGG**CACTGCTTCACCAAGGTGAATATATATTTTACTTATTTTACTTATTTTACAGATATCATC

TGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTACAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAG

HA2 300 bp

CATGTGTGTACCTACAGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTC

ACATATCCTGTCGCACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAA

TACTGAACCTCTGCCAACGTGCATGCAGAAGTCTCAGTCTAACTACAACGAGTCCATTTACTCCAAGAATGATCAG

Partial sequence of *lh* extron3 131 bp

AGTCTGACCTTGTTGTAAGAGCTACTTGCAAAGTACAATATACAACATGACTTAAGACTTGCTTGACTAAGAATTTT

CTTTTCTTGTATGAACAGGGAAACATTTTCAATCAATCACACATAAGGGTAAAAGAGAGTCCTCGAAAAACACAGA

←HA2-R

Downstream of *lh* gene 200 bp

GATCATAATCGCTGCACTCTCGTATTTTAAAATCCACTGATACGCATCATTTAACATAATTTACATCAATTTACTTGA

TTTGTAAAGTCCTGGTTGCTAGTA

Appendix 6: Different constructs of CRISPR/Cas9 system (same method)
Knock in *Cec* at the *mstn* locus, and *As-Cath* at the *lh* locus in channel catfish (dcPlasmid)

#1 Knock in *Cec* at the *mstn* locus

> *mstn* gene sequence (5'-3')

intron1 338 bp

GTAGTGGTAGTGGAGTGGTAGTGTAATGTAGTGTAGTGGAGAAAGTTGTGGGTCTGTCTCTTTAAGGTTTCAGCG

CTGGAAAGGGAGGAAAAAAATCCGGACTGAAGTCCACCTCTGATTTATTGTTGCTCCGAGTAGCCAATCATAGATT

TCGACGCCAGAGCCTAAATAAGAGCGGCGGAATAATTTGGCGGTATAAAAAGGCTTTTGGGCGAATTGAAGCAT

mstn-F →

GACATCTCGCGCTACCTGTCCGGTGTGCATGGCGCACGGTGTTCCTGTTACTGCTGCCACACAGAAAACACAACCG

exon1 421 bp **Target sequence of sgRNA1 at sense strand**

CGCGCGCACTCCTCTCTGAGACCTGACCTGGCTGATC**ATG**CATTTAGCGCAGGTTCTGATTTCGCTGGGCTTCG**TG**

start codon **PAM**

GTGGCGTTCGGTCCGATGGCGCGCACTGACACCGGAGCACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGC

CGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGCCAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCC

mstnqPCR-F →

GTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTGCGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGG

TCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAGCTGCTCGACCTGTACGACGTGCTCGGGGACGACGGC

← mstn-R

AAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGACGACGAGGAGCACGCCACCACCGAGACCGTCATGAGCA

intron2 1750 bp **← mstnqPCR-R**

TGGCCGCCGAGCCCGCCGAGCGTGAGTCCCTTTACTACTGCT//TTCAGTAAGTTGTTATAGAGTATTGTGAGGAG

TGTGAGACTTAAACTGACAGATCGAGGAGTTTAAGGGGTTAATTTGTGCTCTGTGTGCAGCTTAAGCAAGTGTCAC

ACACCGATCCCACAATGCATCAGTTCGTGTACCGTCTTCATTCGATTACTAGTGCAGCTAGCTAACCGTGTTCGGTT

GTGTGTAGTTAGAAGAATAAGGAAGGCGAGTCTGAATACAGGGCTTTACAGTGCTGTAACATTTAACCCCATGTT

GTCTCGGATACCTTTTAAATATATAATCTACTCCTGTTTTCTATTCGCTGAATAATTCTCCTCCTGGTCTCTCCCCCTC

exon2 371 bp

TCTCCTTTAGCCAACCCCGACGTTCAAGTCGACCAAAAACCGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCA

AGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGTGCACTTGCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGC

AGATATCGCGACTCATGCCCATCAAAGACGGGAGAAGGCACGTACGAATACGTTCGCTGAAGATCGACGTGGAC

GCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGAAGCAGGTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAA

CTGGGGGATTGAGATCAACGCGTTCGACTCCAAAAGCAACGATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGA

intron3 306 bp

GGGACTGGTGAGTGTGGATTATTGATATGTATTTGAC//TCTCAAAGAGTCATGCTTTTGTTTTTCCAATCCCAGCTC

exon3 381 bp

CCGTTCTTGGAGGTGAAAATTTCTGAAGTTCCAAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAAT

TCGTCCGAGTCCCGCTGCTGCCGCTACCCCCTTACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCC

CAAAACGCTACAAGGCCAACTACTGCTCGGGCGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACT

TGGTGAACAAGGCCAACCCACGTGGCACTGCCGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCT

CTACTTCAACGGAAAAGAGCAGATCATCTACGGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TG**

stop codon

A

> full Cec_CDS_sequence (5'-3')

CDS1 99 bp

ATGAACTTCAACAAGATCTTCGTCTTTGTGGCACTCATCCTGGCCATCAGCCTGGGAAACTCAGAGGCTGGTTGGC

start codon CDS2 93 bp

TTAGGAAGCTGGGAAAAAAAATCGAACGCATTGGTCAGCATACCAGGGATGCCTCAATCCAGGTCCTCGGAATCG

CCCAACAGGCCGCCAATGTTGCAGCCACCGCTCGAGGTTGA

stop codon

HA1_Cec_CDS_HA2 construct at *mstn* locus:

> *mstn* gene + upstream/downstream sequence(5'-3')

TGTAGTGGAGTGGTAGTGTAGTGGTAGTGGAGCGGTAGTGGAGCGGTAGTGTAATGTAGTGGAGCGGTAGTGT

Upstream of mstn 200 bp

AATGTAGTGTAGTGGTAGTGGAGCGGTAGTGTAGTGGTAGTGGAGCGGTAGTGTAATGTAGTGTAGTGGTAGTG

GTAGTGTAGTGGAGTGGTAGTGGAGTGGTAGTGTAATGTAGTGGAGTGGTAGTGTAGTGGAGAAAGTTGTGGG

TCTGTCTCTTTAAGGTTTCAGCGCTGGAAAGGGAGGAAAAAAATCCGGACTGAAGTCCACCTCTGATTTATTGTTG

CTCCGAGTAGCCAATCATAGATTTCGACGCCAGAGCCTAAATAAGAGCGGCGGAATAATTTGGCGGTATAAAAAG

mstn-F

GCTTTTGGGCGAATTGAAGCATGACATCTCGCGCTACCTGTCCGGTGTGCATGGCGCACGGTGTTCCTGTTACTGC

TGCCACACAGAAAACACAACCGCGCGCGCACTCCTCTCTGAGACCTGACCTGGCTGATCATGCATTTAGCGCAGGT

Target sequence of sgRNA1 at sense strand

TCTGATTTCGCTGGGCTTCGTGGTGGCGTTCGGTCCGATGGCGCGCACTGACACCGGAGCACCGGAGCAGCAGCA

PAM exon1 421 bp

GCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGCCAGCGCGTGCGCTTTCC

mstnqPCR-F

GCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTGCGCCTCAAACAAGCTCC

CAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAGCTGCTCGACCTGTACG

mstn-R

ACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGACGACGAGGAGCACGCCAC

CACCGAGACCGTCATGAGCATGGCCGCCGAGCCCGCCGAGCGTGAGTCCCTTTACTACTGCT//TTCAGTAAGTTG

mstnqPCR-R intron2 1750 bp

TTATAGAGTATTGTGAGGAGTGTGAGACTTAAACTGACAGATCGAGGAGTTTAAGGGGTTAATTTGTGCTCTGTG

TGCAGCTTAAGCAAGTGTCACACACCGATCCCACAATGCATCAGTTCGTGTACCGTCTTCATTCGATTACTAGTGCA

GCTAGCTAACCGTGTTCGGTTGTGTGTAGTTAGAAGAATAAGGAAGGCGAGTCTGAATACAGGGCTTTACAGTGC

TGTAACATTTAACCCCATGTTGTCTCGGATACCTTTTAAATATATAATCTACTCCTGTTTTCTATTCGCTGAATAATTC

TCCTCCTGGTCTCTCCCCCTCTCTCCTTTAGCCAACCCCGACGTTCAAGTCGACCAAAAACCGAAGTGCTGTTTTTTC

exon2 371 bp

TCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGTGCACTTGCGCCCGGCGGATGAG

GCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAGAAGGCACGTACGAATACGTTCG

CTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGAAGCAGGTGCTTGCGGTGTGGCT

GAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAAAGCAACGATCTCGCGATCACTTC

intron3 306 bp

TGCGGAGCCTGGAGAAGAGGGACTGGTGAGTGTGGATTATTGATATGTATTTGAC//TCTCAAAGAGTCATGCTTT

TGTTTTTCCAATCCCAGCTCCCGTTCTTGGAGGTGAAAATTTCTGAAGTTCCAAAGCGAACCAGGAGAGAATCAGG

ACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGCTACCCCCTTACGGTGGACTTTGAAGACTTCGGC

exon3 381 bp

TGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACTGCTCGGGCGAGTGCGACTACGTGCACTTGCAG

AAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTGGCACTGCCGGCCCCTGCTGCACGCCCACCAAG

ATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGATCATCTACGGCAAGATCCCCTCCATGGTAGTGG

ATCGCTGTGGCTGCTCG**TGA**

> HA1_UBI-full Cec_CDS_HA2_pUC57_sequence at exon 1 (dcplasmid: pUC57_Cec)

Backbone of pUC57 vector 415

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACTGTCACAGCTTGTCTGTAAGC

GGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTAT

GCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAA

ATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCT

ATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCAC

Recognition site of sgRNA1 at sense strand 23 bp

GACGTTGTAAAACGACGGCCAGTGAATTGACGCGTATTGGGATTCTGATTTCGCTGGGCTTCG**TGG**GCGCTGGAA

Left homologous arm (HA1) 300 bp **PAM**

AGGGAGGAAAAAAATCCGGACTGAAGTCCACCTCTGATTTATTGTTGCTCCGAGTAGCCAATCATAGATTTCGACGC

CAGAGCCTAAATAAGAGCGGCGGAATAATTTGGCGGTATAAAAAGGCTTTTGGGCGAATTGAAGCATGACATCTCGC

GCTACCTGTCCGGTGTGCATGGCGCACGGTGTTCCTGTTACTGCTGCCACACAGAAAACACAACCGCGCGCGCACTC

CTCTCTGAGACCTGACCTGGCTGATCATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTACCAGCAAAGTTCTAGA

UBI promoter 1438 bp

ATTTGTCGAAACATTTATGTTATATATTTCCTGAAAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATA

AACAACAGACGGCAAAATGAATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATAGGCTGTTCTC

TACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCCCTTTCATTCATATTTCAAGAAAATTGGTTCAGAA

GATCTGGATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCAGTGACCTGATTCCAAATGACTAGGCCTAATCCA

GAAATTAGATGACTGTCAACATAAAAAGGCACAGCACTCACTAGCTGCCCTATATATTTTATTATATTTTACATATA

TTATTTTATTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAGAAAACTGCTTACAACAGCTAACCTGTACTACCT

CAGGCTCAGGGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTTAACCATGCATGCTTGTTTTCAACTATGGCAACA

CAGTCACATGGGACATTACAGAAATGATTTGTCGATGACATGCGACTTTTCTTTAATAAAGCGCAAAGATCCCAAA

AAGCAAACTTTTAACAAAAATCATATAATTATATTTTCAATCCAGCTTTGTAGCAACTTTGTGCTGCTGTTCACTCAG

CAACAGATAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTGCCATCTGTTTCACACATTGCGTTCTATATATAAGT

GTGCTGGTTGACACGACACTGTATAAGGCCTAGGCTAAAACACAAACAATGTAGAATGACACTGTGTTTTTTTTGT

AAACAAATGTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCTCCTGTCATGTATTTGCTATATTCAAATGTTAAAC

CCGTGCAGAATAGAACATATACAAAAAAAAACAACACAACACATTTTTAAACATTATTAAATATCAAGTATTGCTG

GCAGTTCTGTTTCTGTTTTACAGTACCCTTTGCCACAGTTCTCCGCTTTTCCTGGTCCAGATTCCACAAGTCTGATTC

ACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGACTGTCCTGCGAGACCGGCCCATTCC

AGCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATTGATACAAAACATGTAATTCCTAGA

ACATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGAAAGATCTGGTCATGT

GACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAGTAGGACTCTCGACCATCGGCAGATTTTTCGAAG

Cec_CDS 192 bp cecqPCR-→

AAGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGAACTTCAACAAGATCTTCGTCTTTGTGGCACTCATCCTG

GCCATCAGCCTGGGAAACTCAGAGGCTGGTTGGCTTAGGAAGCTGGGAAAAAAAATCGAACGCATTGGTCAGCA

←cecqPCR-R

TACCAGGGATGCCTCAATCCAGGTCCTCGGAATCGCCCAACAGGCCGCCAATGTTGCAGCCACCGCTCGAGGTTG

PolyA tail 288

Agatctttttccctctgccaaaaattatggggacatcatgaagccccttgagcatctgacttctggctaataaaggaaatttatttttcattgcaatagtgt

gttggaatttttttgtgtctctcactcggaaggacatatgggagggcaaatcatttaaaacatcagaatgagtatttggtttagagtttggcaacatatgc

catatgctggctgccatgaacaaaggtggctataaagaggtcatcagtatatgaaacagccccctgctgtccattccttattccatagTCGTGGTG

GCGTTCGGTCCGATGGCGCGCACTGACACCGGAGCACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGT

GACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGCCAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTC

Right homologous arm (HA2) 300

TGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTGCGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCA

AGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAGCTGCTCGACCTGTACGACGTGCTCGGGGACGACGCCACG

PAM

AAGCCCAGCGAAATCAGAATCCCAATGGCGCGCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAA

Recognition site of sgRNA1 at anti-sense strand 23

ATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGT

Backbone of pUC57 vector 2256 bp

GAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAAT

GAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTGTTCCGCTTCCTCGCTCACTGACTCGCTGC

GCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG

ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGC

GTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA

GGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGG

ATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGT

AGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTA

TCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGC

GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGG

TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCT

GGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCA//CACTCGTGCACCCAACTGATCTTCAGCATCTT

TTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACA

CGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGA

TACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACG

TCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTTGTC

> putative positive fish sequence after insertion at exon 1 (5'-3')

TGTAGTGGAGTGGTAGTGTAGTGGTAGTGGAGCGGTAGTGGAGCGGTAGTGTAATGTAGTGGAGCGGTAGTGT

Upstream of mstn 200 bp

AATGTAGTGTAGTGGTAGTGGAGCGGTAGTGTAGTGGTAGTGGAGCGGTAGTGTAATGTAGTGTAGTGGTAGTG

Partial sequence of mstn intron1 72 bp

GTAGTGTAGTGGAGTGGTAGTGGAGTGGTAGTGTAATGTAGTGGAGTGGTAGTGTAGTGGTAGTGGAGTGGTA

HA1-F →

GTGTAATGTAGTGTAGTGGAGAAAGTTGTGGGTCTGTCTCTTTAAGGTTTCAGCGCTGGAAAGGGAGGAAAAAAA

TCCGGACTGAAGTCCACCTCTGATTTATTGTTGCTCCGAGTAGCCAATCATAGATTTCGACGCCAGAGCCTAAATAAG

HA1 300 bp

AGCGGCGGAATAATTTGGCGGTATAAAAAGGCTTTTGGGCGAATTGAAGCATGACATCTCGCGCTACCTGTCCGGTG

TGCATGGCGCACGGTGTTCCTGTTACTGCTGCCACACAGAAAACACAACCGCGCGCGCACTCCTCTCTGAGACCTGA

CCTGGCTGATCATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTACCAGCAAAGTTCTAGAATTTGTCGAAACATTT

ATGTTATATATTTCCTGAAAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATAAACAACAGACGGCAA

UBI promoter 1438 bp

AATGAATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATAG]GCTGTTCTCTACAACCCTCAAACA

GTGATTAGTTTTGTACTTATAAACTTGCCCTTTCATTCATATTTCAAGAAAATTGGTTCAGAAGATCTGGATATTCTA

GCAGTTGTTCAAGCTCATGGAGGGATCAGTGACCTGATTCCAAATGACTAGGCCTAATCCAGAAATTAGATGACT

HA1-R ←

GTCAACATAAAAAGGCACAGCACTCACTAGCTGCCCTATATATTTTATTATATTTTACATATATTATTTTATTTATTTA

GCTCTGAGTGCTGTACTTTCTGGTTAAAGAAAACT//TTCTCCGCTTTTCCTGGTCCAGATTCCACAAGTCTGATTCA

Prom-F →

CCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGACTGTCCTGCGAGACCGGCCCATTCCA

GCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATTGATACAAAACATGTAATTCCTAGAA

CATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGAAAGATCTGGTCATGTG

ACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAGTAGGACTCTCGACCATCGGCAGATTTTTCGAAGA

Cec-F →

Cec_CDS 192 bp

AGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGAACTTCAACAAGATCTTCGTCTTTGTGGCACTCATCCTGG

CCATCAGCCTGGGAAACTCAGAGGCTGGTTGGCTTAGGAAGCTGGGAAAAAAAATCGAACGCATTGGTCAGCAT

← Prom-R

ACCAGGGATGCCTCAATCCAGGTCCTCGGAATCGCCCAACAGGCCGCCAATGTTGCAGCCACCGCTCGAGGTTGA

HA2- →

gatcttttccctctgccaaaaattatggggacatcatgaagcccctgagcatctgacttctggctaataaaggaaatttattttcattgcaatagtgtgt

PolyA tail 288

← Cec-R

tggaatttttgtgtctctcactcggaaggacatatgggagggcaaatcatttaaaacatcagaatgagtatttggtttagagtttggcaacatatgcca

tatgctggctgccatgaacaaaggtggctataaagaggtcatcagtatatgaaacagccccctgctgtccattccttattccatagTCGTGGTGGC

GTTCGGTCCGATGGCGCGCACTGACACCGGAGCACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGA

HA2 300

CGGAGGAGCGCGAGGCGCAGTGTTCAGCGGCCAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTG

CAAGCCATCAAGTCCCAGATTCTGAGCAAACTGCGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAG

CAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAGCTGCTCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCC

Partial sequence of mstn exon1 87 bp

GGGCACAGCGCTCCAGGACGAGGAGGAGGACGACGAGGAGCACGCCACCACCGAGACCGTCATGAGCATGGCC

GCCGAGCCCGCCGAGCGTGAGTCCCTTTACTACTGCTTCATAGCCTAACTTTATTTATTCATTTCTTTATTTTGGTAA
HA2-R
Partial sequence of mstn intron2 130 bp
TGTGTTTTTTTTTTTTTTTTTTTGTTTGTTTTTTCTGTCTAATGCATAGCAACATCTG

#2 Knock in *As-Cath* at the *lh* locus

> *lh* gene sequence (5'-3')

exon1 61 bp
intron1 136 bp
GACAGCAATCCACTGAGCGATCACAGCAAAATCTCTAAAGTAAGGACAGTAATGTGATAAGGTGTGATTTAATTA
LH-F
AATGTATAAATATTACATCCAAAATGTACATAAAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTGCTT
exon2 184 bp
GATTATTTCAGGTATGTGTCGTATTTTTCATGTGTTTTATTACAGGATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCTG
start codon
TGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCT
GTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGCCATCTGCAGCGGGCACTGCTTCACCAAGGTG
intron2 85 bp
Target sequence of sgRNA1
PAM
AATATATATTTTACTTATTTTACTTATTTTACAGATATCATCTGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTA
exon3 365 bp
CAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGTACCTACAGGGACGTTCGCTATGAAA
CCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATCCTGTCGCACTAAGCTGCGAGTGCAG
LH-R
CCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAATCCGGATTTCTGTATGACACAGAAAGAGTAC
ATCCTGGATTACTGAACCTCTGCCAACGTGCATGCAGAAGTCTCAGTCTAACTACAACGAGTCCATTTACTCCAAGA
stop codon
ATGATCAGAGTCTGACCTTGTTGTAAGAGCTACTTGCAAAGTACAATATACAACATGACTTAA

> full As-Cath_CDS_sequence (5'-3')

CDS1 198 bp
ATGCAGACCTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGGGGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCG
start codon
ACCCACCACAGCTCACGCCGACCTACGCCCAGGCCCTGGCCACGGCCGTCGACGTCTACAACCAAGGGCCCGGCG
CDS2 108 bp
TGGACTTCGCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGACGACTGGGACGCGAGCACGGATCCCCTGCGGCAG
CTGGAGTTCACCCTGAAGGAGACCGAGTGCCCCGTGGGCGAGGACCAGCCCCTGGACCAGTGCGACTTCAAGGA
CDS3 84
TGGCGGGGCGGTGCTGGACTGCACAGGGACCTTCTCCTGCTCCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCA
CDS4 129 bp
ACCCGCCGAGCCCCTGCCTGACCGCGTCCGCAGGGGTCTATTCAAGAAGCTGAGGAGGAAAATCAAGAAGGGC
CDS sequence for mature peptide (114 bp)
TTTAAGAAAATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGTGTCTCCATCCCACTCGCAGGAAGGCGGTGA
stop codon

HA3_As_Cath_HA4 construct at *lh* locus:

> *lh* gene + upstream/downstream sequence(5'-3')

ACCTGTCCAAACAGGAAGCTCATTAATCCTTTTGGCTGAGGGCAAGATGAACTGCAGGTTTCTAGTGTCATGGTGT
Upstream of *lh* gene 200 bp
ACCTTTGAAAAAGTAGCCGTGATTATGTCAGAGCCCTGAGCTGTAATTTATTTGAGATTAGTGCGATAAGGCCACG
exon1 61 bp **LHqPCR-F →**
TTTCGATTCTGCGACACTATATAAACATGTTAAACTCTTGTAGGAACAGACAGCAATCCACTGAGCGATCACAGCA
intron1 136 bp **LH-F→**
AAATCTCTAAAGTAAGGACAGTAATGTGATAAGGTGTGATTTAATTAAATGTATAAATATTACATCCAAAATGTAC
ATAAAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTGCTTGATTATTTCAGGTATGTGTCGTATTTTTC
exon2 184 bp
ATGTGTTTTATTACAGG**ATG**TCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGC
TCAAAGCTACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGC
← LHqPCR-R **intron2 85 bp**
CTTGTGTTTCAAACCGCCATCTGCAGC**GGG**CACTGCTTCACCAAGGTGAATATATATTTTACTTATTTTACTTATTTT
Target sequence of sgRNA2 **PAM**
ACAGATATCATCTGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTACAGGAACCTGTGTACAAGAGCCCGTTC
exon3 365 bp
TCTTCCATCTATCAGCATGTGTGTACCTACAGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCG
← LH-R
GTGTGGATCCTCACGTCACATATCCTGTCGCACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTG
TACCATCGAGAGCCTGAATCCGGATTTCTGTATGACACAGAAAGAGTACATCCTGGATTAC**TGA**ACCTCTGCCAAC
GTGCATGCAGAAGTCTCAGTCTAACTACAACGAGTCCATTTACTCCAAGAATGATCAGAGTCTGACCTTGTTGTAA
Downstream of *lh* gene 200 bp
GAGCTACTTGCAAAGTACAATATACAACATGACTTAAGACTTGCTTGACTAAGAATTTTCTTTTCTTGTATGAACAG
GGAAACATTTTCAATCAATCACACATAAGGGTAAAAGAGAGTCCTCGAAAAACACAGAGATCATAATCGCTGCAC
TCTCGTATTTTAAAATCCACTGATACGCATCATTTAACATAATTTACATCAATTTACTTGATTTGTAAAGTCCTGGTT
GCTAGTA

> HA3_UBI-full As-Cath_CDS_HA4_pUC57_sequence at exon 2 (dcPlasmid: pUC_57 Cath)

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACTGTCACAGCTTGTCTGTAAGC
Backbone of pUC57 vector 415
GGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTAT
GCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAA
ATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCT
ATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCAC
GACGTTGTAAAACGACGGCCAGTGAATTGACGCGTATTGGGATTTCAAACCGCCATCTGCAGC**GGG**GGTGTGATT
Left homologous arm (HA3) 300 bp **Target sequence of sgRNA2 PAM**
TAATTAAATGTATAAATATTACATCCAAAATGTACATAAAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTG
CTTGATTATTTCAGGTATGTGTCGTATTTTTCATGTGTTTTATTACAGGATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCT
GTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCT
GTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGCCATCTGCACCAGCAAAGTTCTAGAATTTGTCGA
UBI promoter 1438 bp
AACATTTATGTTATATATTTCCTGAAAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATAAACAACAG

ACGGCAAAATGAATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATAGGCTGTTCTCTACAACCC

TCAAACAGTGATTAGTTTTGTACTTATAAACTTGCCCTTTCATTCATATTTCAAGAAAATTGGTTCAGAAGATCTGG

ATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCAGTGACCTGATTCCAAATGACTAGGCCTAATCCAGAAATTA

GATGACTGTCAACATAAAAAGGCACAGCACTCACTAGCTGCCCTATATATTTTATTATATTTTACATATATTATTTTA

TTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAGAAAACTGCTTACAACAGCTAACCTGTACTACCTCAGGCTC

AGGGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTTAACCATGCATGCTTGTTTTCAACTATGGCAACACAGTCAC

ATGGGACATTACAGAAATGATTTGTCGATGACATGCGACTTTTCTTTAATAAAGCGCAAAGATCCCAAAAAGCAAA

CTTTTAACAAAAATCATATAATTATATTTTCAATCCAGCTTTGTAGCAACTTTGTGCTGCTGTTCACTCAGCAACAGA

TAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTGCCATCTGTTTCACACATTGCGTTCTATATATAAGTGTGCTGG

TTGACACGACACTGTATAAGGCCTAGGCTAAAACACAAACAATGTAGAATGACACTGTGTTTTTTTTGTAAACAAA

TGTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCTCCTGTCATGTATTTGCTATATTCAAATGTTAAACCCGTGCA

GAATAGAACATATACAAAAAAAAACAACACAACACATTTTTAAACATTATTAAATATCAAGTATTGCTGGCAGTTCT

GTTTCTGTTTTACAGTACCCTTTGCCACAGTTCTCCGCTTTTCCTGGTCCAGATTCCACAAGTCTGATTCACCAATAG

CAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGACTGTCCTGCGAGACCGGCCCATTCCAGCACATT

CTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATTGATACAAAACATGTAATTCCTAGAACATAACC

ATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGAAAGATCTGGTCATGTGACGTCTC

ATGAACGTCACGGCCCTGGGTTTCTATAAATACAGTAGGACTCTCGACCATCGGCAGATTTTTCGAAGAAGAAGAT

full As-Cath CDS (519 bp)

CAGTTTCAGGAGCCGTACTGTTCCGTTATGCAGACCTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGGGGCAGCCA

start codon

GCACCGAGCTGCCCACCCCTGGCACCGACCCACCACAGCTCACGCCGACCTACGCCCAGGCCCTGGCCACGGCCG

TCGACGTCTACAACCAAGGGCCCGGCGTGGACTTCGCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGACGACTGGG

ACGCGAGCACGGATCCCCTGCGGCAGCTGGAGTTCACCCTGAAGGAGACCGAGTGCCCCGTGGGCGAGGACCAG

CCCCTGGACCAGTGCGACTTCAAGGATGGCGGGGCGGTGCTGGACTGCACAGGGACCTTCTCCTGCTCCGAGGCC

TCGCTCATGGTGCTGGTCACCTGCCAACCCGCCGAGCCCCTGCCTGACCGCGTC**CGCAGGGGTCTATTCAAGAAG**

CDS sequence for mature peptide (114 bp)

CTGAGGAGGAAAATCAAGAAGGGCTTTAAGAAAATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGTGTCTCCAT

CCCACTCGCAGGAAGGCGGTGAgatccagacatgataagatacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaat

PolyA tail 135 bp **stop codon**

gctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgcaataaacaagttAGC**GGG**CACTGCTTCACCAAGGTGAA

Right homologous arm (HA4) 300 bp

TATATATTTTACTTATTTTACTTATTTTACAGATATCATCTGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTACAGGA

ACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGTACCTACAGGGACGTTCGCTATGAAACCGTACGC

CTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATCCTGTCGCACTAAGCTGCGAGTGCAGCCTGTGCACC

Recognition site of sgRNA1 at anti-sense strand 23

ATGGACACCTCGGACTGTACCATCGAGAGCCTGAA**CCC**GCTGCAGATGGCGGTTTGAAATCCCAATGGCGCGCCGA

PAM

GCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCC

Backbone of pUC57 vector 2256 bp

GGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCC

GCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGT

ATTGGGCGCTGTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC

ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAG

CAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCAC

AAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGC

TCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC

GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAA

CCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC

GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT

GGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAA

AAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATT

ACGCGCA//CACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG

GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAAT

ATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATA

GGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAA

AAATAGGCGTATCACGAGGCCCTTTTGTC

> putative positive fish sequence after insertion (5'-3')

Upstream of *lh* gene 200 bp

ACCTGTCCAAACAGGAAGCTCATTAATCCTTTTGGCTGAGGGCAAGATGAACTGCAGGTTTCTAGTGTCATGGTGT

UBI promoter 1438

ACCTTTGAAAAAGTAGCCGTGATTATGTCAGAGCCCTGAGCTGTAATTTATTTGAGATTAGTGCGA**TAAGGCCACG**

TTTCGATTCTGCGACACTATATAAACATGTTAAACTCTTGTAGGAACA**GACAGCAATCCACTGAGCGATCACAGCA**

Partial sequence of *lh* extron1 60 bp

AAATCTCTAAAGTAAGGACAGTAATGTGATAAGGTGTGATTTAATTAAATGTATAAATATTACATCCAAAATGTACATA

AAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTGCTTGATTATTTCAGGTATGTGTCGTATTTTTCATGTGT

HA3 300 bp

TTTATTACAGGATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGC

TACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTT

CAAACCGCCATCTGCACCAGCAAAGTTCTAGAATTTGTCGAAACATTTATGTTATATATTTCCTGAAAAAAATTCTG

HA3-R

AGTAAGTTCTTAAGTGTTATTGCCAGCAACATA**AACAACAGACGGCAAAATGA**ATAAATGATAACAAAGCAGTAG

GCTTAAATAAACCTAATTTTTATAGGCTGTTCTCTACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCC

CTTTCATTCATATTTCAAGAAAATTGGTTCAGAAGATCTGGATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCA

GTGACCTGATTCCAAATGACTAGGCCTAATCCAGAAATTAGATGACTGTCAACATAAAAAGGCACAGCACTCACTA

GCTGCCCTATATATTTTATTATATTTTACATATATTATTTTATTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAG

AAAACTGCTTACAACAGCTAACCTGTACTACCTCAGGCTCAGGGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTT

AACCATGCATGCTTGTTTTCAACTATGGCAACACAGTCACATGGGACATTACAGAAATGATTTGTCGATGACATGC

GACTTTTCTTTAATAAAGCGCAAAGATCCCAAAAAGCAAACTTTTAACAAAAATCATATAATTATATTTTCAATCCA

GCTTTGTAGCAACTTTGTGCTGCTGTTCACTCAGCAACAGATAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTG

CCATCTGTTTCACACATTGCGTTCTATATATAAGTGTGCTGGTTGACACGACACTGTATAAGGCCTAGGCTAAAACA

CAAACAATGTAGAATGACACTGTGTTTTTTTTGTAAACAAATGTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCT

CCTGTCATGTATTTGCTATATTCAAATGTTAAACCCGTGCAGAATAGAACATATACAAAAAAAAACAACACAACAC

Prom2-F →

ATTTTTAAACATTATTAAATATCAAGTATTGCTGGCAGTTCTGTTTCTGTTTTACAGTACCCTTTGCCACAGTTCTCCG

CTTTTCCTGGTCCAGATTCCACAAGTCTGATTCACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTG

CTTGTAGACTGTCCTGCGAGACCGGCCCATTCCAGCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACAT

TTAAATTATTGATACAAAACATGTAATTCCTAGAACATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTT

TTAGGATTTGACTGCGGAAAGATCTGGTCATGTGACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAG

Cath-F →

TAGGACTCTCGACCATCGGCAGATTTTTCGAAGAAGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGCAGAC

full AS-Cath CDS 519 bp

CTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGGGGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCGACCCACCA

CAGCTCACGCCGACCTACGCCCAGGCCCTGGCCACGGCCGTCGACGTCTACAACCAAGGGCCCGGCGTGGACTTC

← Prom2-R

GCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGACGACTGGGACGCGAGCACGGATCCCCTGCGGCAGCTGGAGTT

CACCCTGAAGGAGACCGAGTGCCCCGTGGGCGAGGACCAGCCCCTGGACCAGTGCGACTTCAAGGATGGCGGGG

CGGTGCTGGACTGCACAGGGACCTTCTCCTGCTCCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCAACCCGCCGA

GCCCCTGCCTGACCGCGTC**CGCAGGGGTCTATTCAAGAAGCTGAGGAGGAAAATCAAGAAGGGCTTTAAGAAA**

ATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGTGTCTCCATCCCACTCGCAGGAAGGCGGTGAgatccagacatgataa

HA4-F →

gatacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattata

← Cath-R

agctgcaataaacaagttAGC**GGG**CACTGCTTCACCAAGGTGAATATATATTTTACTTATTTTACTTATTTTACAGATATCATC

HA4 300 bp

TGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTACAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAG

CATGTGTGTACCTACAGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTC

ACATATCCTGTCGCACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAA

TACTGAACCTCTGCCAACGTGCATGCAGAAGTCTCAGTCTAACTACAACGAGTCCATTTACTCCAAGAATGATCAG

Partial sequence of *lh* extron3 131 bp

AGTCTGACCTTGTTGTAAGAGCTACTTGCAAAGTACAATATACAACATGACTTAAGACTTGCTTGACTAAGAATTTT

Downstream of *lh* gene 200 bp

CTTTTCTTGTATGAACAGGGAAACATTTTCAATCAATCACACATAAGGGTAAAAGAGAGTCCTCGAAAAACACAGA

← HA4-R

GATCATAATCGCTGCACTCTCGTATTTTAAAATCCACTGATACGCATCATTTAACATAATTTACATCAATTTACTTGA

TTTGTAAAGTCCTGGTTGCTAGTA

Appendix 7: Different constructs of CRISPR/Cas9 system (different methods)
Knock in *As-Cath* at the *lh* and *mc4r* loci, *Cec* at the *mstn* locus (two sites: *mstn1* and *mstn2*)

#1 Knock in the *As-Cath* transgene at the *lh* locus (linear dsDNA)

> *lh* gene sequence (5'-3')

exon1 61 bp — intron1 136 bp

GACAGCAATCCACTGAGCGATCACAGCAAAATCTCTAAAGTAAGGACAGTAATGTGATAAGGTGTGATTTAATTA

LH-F →

AATGTATAAATATTACATCCAAAATGTACATAAAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTGCTT

exon2 184 bp

GATTATTTCAGGTATGTGTCGTATTTTTCATGTGTTTTATTACAGG**ATG**TCAGTGCCAGCTTCCTCTTTTCTTCTCCTG

start codon

TGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCT

GTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGCCATCTGCAGC**GGG**CACTGCTTCACCAAGGTG

intron2 85 bp — Target sequence of sgRNA1 — PAM

AATATATATTTTACTTATTTTACTTATTTTACAGATATCATCTGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTA

exon3 365 bp

CAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGTACCTACAGGGACGTTCGCTATGAAA

CCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATCCTGTCGCACTAAGCTGCGAGTGCAG

← LH-R

CCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAATCCGGATTTCTGTATGACACAGAAAGAGTAC

ATCCTGGATTAC**TGA**ACCTCTGCCAACGTGCATGCAGAAGTCTCAGTCTAACTACAACGAGTCCATTTACTCCAAGA

stop codon

ATGATCAGAGTCTGACCTTGTTGTAAGAGCTACTTGCAAAGTACAATATACAACATGACTTAA

> full *As-Cath* CDS sequence (5'-3')

CDS1 198 bp

ATGCAGACCTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGGGGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCG

start codon

ACCCACCACAGCTCACGCCGACCTACGCCCAGGCCCTGGCCACGGCCGTCGACGTCTACAACCAAGGGCCCGGCG

CDS2 108 bp

TGGACTTCGCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGACGACTGGGACGCGAGCACGGATCCCCTGCGGCAG

CTGGAGTTCACCCTGAAGGAGACCGAGTGCCCCGTGGGCGAGGACCAGCCCCTGGACCAGTGCGACTTCAAGGA

CDS3 84

TGGCGGGGCGGTGCTGGACTGCACAGGGACCTTCTCCTGCTCCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCA

CDS4 129 bp

ACCCGCCGAGCCCCTGCCTGACCGCGTC**CGCAGGGGTCTATTCAAGAAGCTGAGGAGGAAAATCAAGAAGGGC**

CDS sequence for mature peptide (114 bp)

TTTAAGAAAATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGTGTCTCCATCCCACTCGCAGGAAGGCGGTGA

stop codon

HA1_UBI_*As-Cath*_pA_HA2 construct at the *lh* locus:

> *lh* gene + upstream/downstream sequence(5'-3')

ACCTGTCCAAACAGGAAGCTCATTAATCCTTTTGGCTGAGGGCAAGATGAACTGCAGGTTTCTAGTGTCATGGTGT
Upstream of *lh* gene 200 bp
ACCTTTGAAAAAGTAGCCGTGATTATGTCAGAGCCCTGAGCTGTAATTTATTTGAGATTAGTGCGATAAGGCCACG
exon1 61 bp LHqPCR-F
TTTCGATTCTGCGACACTATATAAACATGTTAAACTCTTGTAGGAACAGACAGCAATCCACTGAGCGATCACAGCA
intron1 136 bp LH-F
AAATCTCTAAAGTAAGGACAGTAATGTGATAAGGTGTGATTTAATTAAATGTATAAATATTACATCCAAAATGTAC
ATAAAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTGCTTGATTATTTCAGGTATGTGTCGTATTTTTC
exon2 184 bp
ATGTGTTTTATTACAGG**ATG**TCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGC
TCAAAGCTACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGC
LHqPCR-R intron2 85 bp
CTTGTGTTTCAAACCGCCATCTGCAGC**GGG**CACTGCTTCACCAAGGTGAATATATATTTTACTTATTTTACTTATTTT
Target sequence of sgRNA1 PAM
ACAGATATCATCTGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTACAGGAACCTGTGTACAAGAGCCCGTTC
exon3 365 bp
TCTTCCATCTATCAGCATGTGTGTACCTACAGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCG
LH-R
GTGTGGATCCTCACGTCACATATCCTGTCGCACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTG
TACCATCGAGAGCCTGAATCCGGATTTCTGTATGACACAGAAAGAGTACATCCTGGATTAC**TGA**ACCTCTGCCAAC
GTGCATGCAGAAGTCTCAGTCTAACTACAACGAGTCCATTTACTCCAAGAATGATCAGAGTCTGACCTTGTTGTAA
Downstream of *lh* gene 200 bp
GAGCTACTTGCAAAGTACAATATACAACATGACTTAAGACTTGCTTGACTAAGAATTTTCTTTTCTTGTATGAACAG
GGAAACATTTTCAATCAATCACACATAAGGGTAAAAGAGAGTCCTCGAAAAACACAGAGATCATAATCGCTGCAC
TCTCGTATTTTAAAATCCACTGATACGCATCATTTAACATAATTTACATCAATTTACTTGATTTGTAAAGTCCTGGTT
GCTAGTA

> HA1_UBI_*As-Cath*_pA_HA2 sequence (linear dsDNA)

GGTGTGATTTAATTAAATGTATAAATATTACATCCAAAATGTACATAAAAGTGCAAAGTGTAGTCTGAACTCAAGATAAT
Left homologous arm (HA1) 300 bp
TTGTCCTTGCTTGATTATTTCAGGTATGTGTCGTATTTTTCATGTGTTTTATTACAGGATGTCAGTGCCAGCTTCCTCTTT
TCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGCCACACTGCGAACCTGTTAATGAG
ACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGCCATCTGCACCAGCAAAGTTCTAGA
ATTTGTCGAAACATTTATGTTATATATTTCCTGAAAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATA
UBI promoter 1438 bp
AACAACAGACGGCAAAATGAATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATAGGCTGTTCTC
TACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCCCTTTCATTCATATTTCAAGAAAATTGGTTCAGAA
GATCTGGATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCAGTGACCTGATTCCAAATGACTAGGCCTAATCCA
GAAATTAGATGACTGTCAACATAAAAAGGCACAGCACTCACTAGCTGCCCTATATATTTTATTATATTTTACATATA
TTATTTTATTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAGAAAACTGCTTACAACAGCTAACCTGTACTACCT
CAGGCTCAGGGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTTAACCATGCATGCTTGTTTTCAACTATGGCAACA

CAGTCACATGGGACATTACAGAAATGATTTGTCGATGACATGCGACTTTTCTTTAATAAAGCGCAAAGATCCCAAA

AAGCAAACTTTTAACAAAAATCATATAATTATATTTTCAATCCAGCTTTGTAGCAACTTTGTGCTGCTGTTCACTCAG

CAACAGATAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTGCCATCTGTTTCACACATTGCGTTCTATATATAAGT

GTGCTGGTTGACACGACACTGTATAAGGCCTAGGCTAAAACACAAACAATGTAGAATGACACTGTGTTTTTTTTGT

AAACAAATGTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCTCCTGTCATGTATTTGCTATATTCAAATGTTAAAC

CCGTGCAGAATAGAACATATACAAAAAAAAACAACACAACACATTTTTAAACATTATTAAATATCAAGTATTGCTG

GCAGTTCTGTTTCTGTTTTACAGTACCCTTTGCCACAGTTCTCCGCTTTTCCTGGTCCAGATTCCACAAGTCTGATTC

ACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGACTGTCCTGCGAGACCGGCCCATTCC

AGCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATTGATACAAAACATGTAATTCCTAGA

ACATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGAAAGATCTGGTCATGT

GACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAGTAGGACTCTCGACCATCGGCAGATTTTTCGAAG

AAGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGCAGACCTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGG

start codon

GGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCGACCCACCACAGCTCACGCCGACCTACGCCCAGGCCCTGGC

full AS-Cath CDS 519

CACGGCCGTCGACGTCTACAACCAAGGGCCCGGCGTGGACTTCGCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGA

CGACTGGGACGCGAGCACGGATCCCCTGCGGCAGCTGGAGTTCACCCTGAAGGAGACCGAGTGCCCCGTGGGCG

AGGACCAGCCCCTGGACCAGTGCGACTTCAAGGATGGCGGGGCGGTGCTGGACTGCACAGGGACCTTCTCCTGCT

CCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCAACCCGCCGAGCCCCTGCCTGACCGCGTC**CGCAGGGGTCTATT**

CAAGAAGCTGAGGAGGAAAATCAAGAAGGGCTTTAAGAAAATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGT

GTCTCCATCCCACTCGCAGGAAGGCGGTGAgatccagacatgataagatacattgatgagtttggacaaaccacaactagaatgcagt

stop codon

gaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgcaataaacaagttAGC**GGG**CACTGCTTCACC

PolyA tail 135 bp

AAGGTGAATATATATTTTACTTATTTTACTTATTTTACAGATATCATCTGAGTATGAAAAATGGTTTACAAATAAATGTTTT

CTACAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGTACCTACAGGGACGTTCGCTATGAA

ACCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATCCTGTCGCACTAAGCTGCGAGTGCAGC

Right homologous arm (HA2) 300

CTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAA

> putative positive fish sequence after insertion at the *lh* locus(5'-3')

Upstream of *lh* gene 200 bp

ACCTGTCCAAACAGGAAGCTCATTAATCCTTTTGGCTGAGGGCAAGATGAACTGCAGGTTTCTAGTGTCATGGTGT

HA1-F

ACCTTTGAAAAAGTAGCCGTGATTATGTCAGAGCCCTGAGCTGTAATTTATTTGAGATTAGTGCGATAAGGCCACG

Partial sequence of *lh* extron1 60 bp

TTTCGATTCTGCGACACTATATAAACATGTTAAACTCTTGTAGGAACAGACAGCAATCCACTGAGCGATCACAGCA

AAATCTCTAAAGTAAGGACAGTAATGTGATAAGGTGTGATTTAATTAAATGTATAAATATTACATCCAAAATGTACATA

HA1 300 bp

AAAGTGCAAAGTGTAGTCTGAACTCAAGATAATTTGTCCTTGCTTGATTATTTCAGGTATGTGTCGTATTTTTCATGTGT

TTTATTACAGGATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGC

TACATTCTGCCACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTT

CAAACCGCCATCTGCACCAGCAAAGTTCTAGAATTTGTCGAAACATTTATGTTATATATTTCCTGAAAAAAATTCTG

AGTAAGTTCTTAAGTGTTATTGCCAGCAACATAAACAACAGACGGCAAAATGAATAAATGATAACAAAGCAGTAG

UBI promoter 1438 ← HA1-R

GCTTAAATAAACCTAATTTTTATAGGCTGTTCTCTACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCC

CTTTCATTCATATTTCAAGAAAATTGGTTCAGAAGATCTGGATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCA

GTGACCTGATTCCAAATGACTAGGCCTAATCCAGAAATTAGATGACTGTCAACATAAAAAGGCACAGCACTCACTA

GCTGCCCTATATATTTTATTATATTTTACATATATTATTTTATTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAG

AAAACTGCTTACAACAGCTAACCTGTACTACCTCAGGCTCAGGGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTT

AACCATGCATGCTTGTTTTCAACTATGGCAACACAGTCACATGGGACATTACAGAAATGATTTGTCGATGACATGC

GACTTTTCTTTAATAAAGCGCAAAGATCCCAAAAAGCAAACTTTTAACAAAAATCATATAATTATATTTTCAATCCA

GCTTTGTAGCAACTTTGTGCTGCTGTTCACTCAGCAACAGATAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTG

CCATCTGTTTCACACATTGCGTTCTATATATAAGTGTGCTGGTTGACACGACACTGTATAAGGCCTAGGCTAAAACA

CAAACAATGTAGAATGACACTGTGTTTTTTTTGTAAACAAATGTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCT

CCTGTCATGTATTTGCTATATTCAAATGTTAAACCCGTGCAGAATAGAACATATACAAAAAAAAACAACACAACAC

Prom2-F →

ATTTTTAAACATTATTAAATATCAAGTATTGCTGGCAGTTCTGTTTCTGTTTTACAGTACCCTTTGCCACAGTTCTCCG

CTTTTCCTGGTCCAGATTCCACAAGTCTGATTCACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTG

CTTGTAGACTGTCCTGCGAGACCGGCCCATTCCAGCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACAT

TTAAATTATTGATACAAAACATGTAATTCCTAGAACATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTT

TTAGGATTTGACTGCGGAAAGATCTGGTCATGTGACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAG

TAGGACTCTCGACCATCGGCAGATTTTTCGAAGAAGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGCAGAC

← Prom2-R

CTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGGGGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCGACCCACCA

full As-Cath CDS 519 bp

CAGCTCACGCCGACCTACGCCCAGGCCCTGGCCACGGCCGTCGACGTCTACAACCAAGGGCCCGGCGTGGACTTC

Cath-F →

GCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGACGACTGGGACGCGAGCACGGATCCCCTGCGGCAGCTGGAGTT

CACCCTGAAGGAGACCGAGTGCCCCGTGGGCGAGGACCAGCCCCTGGACCAGTGCGACTTCAAGGATGGCGGGG

CGGTGCTGGACTGCACAGGGACCTTCTCCTGCTCCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCAACCCGCCGA

GCCCCTGCCTGACCGCGTC**CGCAGGGGTCTATTCAAGAAGCTGAGGAGGAAAATCAAGAAGGGCTTTAAGAAA**

ATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGTGTCTCCATCCCACTCGCAGGAAGGCGGTGAgatccagacatgataa

gatacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattata

HA2-F → ← Cath-R

agctgcaataaacaagttAGC**GGG**CACTGCTTCACCAAGGTGAATATATATTTTACTTATTTTACTTATTTTACAGATATCATC

TGAGTATGAAAAATGGTTTACAAATAAATGTTTTCTACAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAG

CATGTGTGTACCTACAGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTC

HA2 300 bp

ACATATCCTGTCGCACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAA

TACTGAACCTCTGCCAACGTGCATGCAGAAGTCTCAGTCTAACTACAACGAGTCCATTTACTCCAAGAATGATCAG

Partial sequence of *lh* extron3 131 bp

AGTCTGACCTTGTTGTAAGAGCTACTTGCAAAGTACAATATACAACATGACTTAAGACTTGCTTGACTAAGAATTTT

CTTTTCTTGTATGAACAGGGAAACATTTTCAATCAATCACACATAAGGGTAAAAGAGAGTCCTCGAAAAACACAGA

HA2-R

Downstream of *lh* gene 200 bp

GATCATAATCGCTGCACTCTCGTATTTTAAAATCCACTGATACGCATCATTTAACATAATTTACATCAATTTACTTGA

TTTGTAAAGTCCTGGTTGCTAGTA

#2 Knock in the *As-Cath* transgene at the *mc4r* locus (double-cut plasmid)

> *mc4r* gene sequence (5'-3')

Partial sequence of mc4r intron1421 bp

CAGTGATCACAGTGCAGAACAGGGGCTAAGAGGGAAAGTGCCAATCTGCCAAAGGAACGCTCCCGAGAGCAGGT

GCCTACCTCGGATTCCTGCATCTTCTTCTTTTATTCCTCCGGTCTCTTCTTCTCAAGCTCAGTCTCAGCACTGTCCTTTT

AAGGCTCTCGGCTTTTTTTCTTTATCTGCTCTTCCTCATCCTTCGAACGCTGCTCGAGGATGACAGTGTGTGAGGGT

Mc4r-F

TAAATAGTAGCTGCTGGAAAGCTCGCTGATGCTCGAGGATGACACGGGGACCTCCGTGCACTGGGAAAACGCATC

TCTTTGGGACTCTGTAGCTCAGATGGAGATGGAGGACACGGAAGAGACTCGCAGATTAGAATAAACGCAGATGA

AGACGGAAAGCGGAGGACTGTGGTGAGGAGGTCTTGCGGATATGAACGTGTCGGAGCACCACGGGATGCAGCA

mc4r CDS 1017 bp

start codon

TGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAAAGCCGGCTCGGGGGAAAGGAACTCGGAGTCG

Recognition site of sgRNA at anti-sense strand 23 bp

GGCTGCTACGAGCAGCTGTTGATCTCCACCGAGGTCTTCATCACGCTAGGGTTGGTCAGCCTTCTGGAGAACATCC

PAM

TGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGCCCATGTACTTCTTCATCTGCAGCCTGGCGGTGGC

CGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTGTGATGGCGCTGATCACCAGCGGCAACCTGACCAT

CTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACTCCATGATCTGCAGCTCACTCCTGGCCTCCATTTGG

Mc4r-R

AGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTCTACGCCCTGCGCTACCACAACATCATGACCCAACG

CCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCACGGCGTCCGGTGTGCTCTTCATCATCTACTCGGAG

AGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACCATGCTGGCCCTCATGGCCTCGCTTTACGTGCACAT

GTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTTACCGGGGAACGGCCCCGTGTGGCAGGCGGCCAA

CATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGTTTGTCGTGTGCTGGGCGCCGTTTTTTCTCCACCTC

ATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCTTCATGTCTCACTTCAACATGTACCTGATTCTGATC

ATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGGAGTCAGGAGATGAGGAAGACCTTCCGGGAGATC

TGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTGCGTCGGCTTCGACGAGAGGCTTAACAGCTATTG

stop codon

AACCGGACGTTTTGTTTTTGACCTCGAACGCACTGACGATTCGCTCCTCGTCCCGGACCGTTCGCTGGCTGAAATCG

Partial sequence of mc4r intron2 302 bp

CGAACCACGTCGGACGATTGCTTTTCTTCACCAACAAAAAACCAAGACCTTCAAAATAACCCGACTCAAAACCTCCA

GAAACACTCGTCTTGTGAAACAACCTTGAATTTCGTCTTTTGCTACACGACTGTGAAATCATTTCTGCCTTTCCATCA

TCATCATCATCACCACCAACAGGCGGGTTACCTACGCTCGAAAATCAGTTCTGGACAAACTCCTGCATGAC

> sg_HA3_UBI_*As-Cath*_pA_HA4_sg sequence (dcPlasmid: pUC57_*As-Cath*)

Upstream of pUC57 vector 415

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACTGTCACAGCTTGTCTGTAAGC

GGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTAT

GCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAA

ATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCT

ATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCAC

Recognition site of sgRNA at sense strand 23 bp (sgRNA2)

GACGTTGTAAAACGACGGCCAGTGAATTGACGCGTATTGGGATGCAGCTGTTGATCTCCACCG**AGG**CGCTGATGC

HA3 300 bp **PAM**

TCGAGGATGACACGGGGACCTCCGTGCACTGGGAAAACGCATCTCTTTGGGACTCTGTAGCTCAGATGGAGATGG

AGGACACGGAAGAGACTCGCAGATTAGAATAAACGCAGATGAAGACGGAAAGCGGAGGACTGTGGTGAGGAGG

TCTTGCGGATATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGA

TTGGAAACAAAGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCTCCAACCAG

CAAAGTTCTAGAATTTGTCGAAACATTTATGTTATATATTTCCTGAAAAAAATTCTGAGTAAGTTCTTAAGTGTTATT

UBI promoter 1438 bp

GCCAGCAACATAAACAACAGACGGCAAAATGAATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTT

ATAGGCTGTTCTCTACAACCCTCAAACAGTGATT//ACCAAAGCAGCCAATCACTGCTTGTAGACTGTCCTGCGAGA

CCGGCCCATTCCAGCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATTGATACAAAACAT

GTAATTCCTAGAACATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGAAAG

ATCTGGTCATGTGACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAGTAGGACTCTCGACCATCGGCA

GATTTTTCGAAGAAGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGCAGACCTGCTGGGTCATCCTGCTCCT

GCCCCTGCTCGGGGCAGCCAGCACCGAGCTGCCCACCCCTGGCACCGACCCACCACAGCTCACGCCGACCTACGC

CCAGGCCCTGGCCACGGCCGTCGACGTCTACAACCAAGGGCCCGGCGTGGACTTCGCCTTCCGGCTCCTGGAGGC

AGAGTCCCGGGACGACTGGGACGCGAGCACGGATCCCCTGCGGCAGCTGGAGTTCACCCTGAAGGAGACCGAGT

GCCCCGTGGGCGAGGACCAGCCCCTGGACCAGTGCGACTTCAAGGATGGCGGGGCGGTGCTGGACTGCACAGG

GACCTTCTCCTGCTCCGAGGCCTCGCTCATGGTGCTGGTCACCTGCCAACCCGCCGAGCCCCTGCCTGACCGCGTC

CGCAGGGGTCTATTCAAGAAGCTGAGGAGGAAAATCAAGAAGGGCTTTAAGAAAATCTTCAAGCGCCTGCCTC

CCGTCGGTGTCGGTGTCTCCATCCCACTCGCAGGAAGGCGGTGAgatccagacatgataagatacattgatgagtttggacaaa

ccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgcaataaacaagttCCG**A**

GGTCTTCATCACGCTAGGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAA

HA4 300 bp

CTTCCACTCGCCCATGTACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACA

GAAACGGCTGTGATGGCGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAAT

GTGTTCGACTCCATGATCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACC**CCT**CGGT

PAM

GGAGATCAACAGCTGCATCCCAATGGCGCGCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT

Recognition site of sgRNA at anti-sense strand 23 bp

GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAA//ATCATTGGAAAACGTTCTTCG

GGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTT

Downstream of pUC57 vector 2256 bp

CAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAA

GGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTC

ATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC

CACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTTGTC

> putative positive fish sequence after insertion at the *mc4r* locus (5'-3')

CAGTGATCACAGTGCAGAACAGGGGCTAAGAGGGAAAGTGCCAATCTGCCAAAGGAACGCTCCCGAGAGCAGGT

HA3-F→

GCCTACCTCGGATTCCTGCATCTTCTTCTTTTATTCCTCCGGTCTCTTCTTCTCAAGCTCAGTCTCAGCACTGTCCTTTT

AAGGCTCTCGGCTTTTTTTCTTTATCTGCTCTTCCTCATCCTTCGAACGCTGCTCGAGGATGACAGTGTGTGAGGGT

TAAATAGTAGCTGCTGGAAAGCTCGCTGATGCTCGAGGATGACACGGGGACCTCCGTGCACTGGGAAAACGCATC

TCTTTGGGACTCTGTAGCTCAGATGGAGATGGAGGACACGGAAGAGACTCGCAGATTAGAATAAACGCAGATGA

HA3 300 bp

AGACGGAAAGCGGAGGACTGTGGTGAGGAGGTCTTGCGGATATGAACGTGTCGGAGCACCACGGGATGCAGCA

TGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAAAGCCGGCTCGGGGGAAAGGAACTCGGAGTCG

←HA3-R

GGCTGCTACGAGCAGCTGTTGATCTCCAACCAGCAAAGTTCTAGAATTTGTCGAAACATTTATGTTATATATTTCCT

GAAAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATAAACAACAGACGGCAAAATGAATAAATGAT

AACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATAGGCTGTTCTCTACAACCCTCAAACAGTGATT//ACCAAAG

CAGCCAATCACTGCTTGTAGACTGTCCTGCGAGACCGGCCCATTCCAGCACATTCTGGAAACTTCCTTTATATGATA

ATTATAAATACATTTAAATTATTGATACAAAACATGTAATTCCTAGAACATAACCATAGCAATCATTAGTTTTCAGG

GTAATTATGTATTTTTAGGATTTGACTGCGGAAAGATCTGGTCATGTGACGTCTCATGAACGTCACGGCCCTGGGT

TTCTATAAATACAGTAGGACTCTCGACCATCGGCAGATTTTTCGAAGAAGAAGATCAGTTTCAGGAGCCGTACTGT

TCCGTTATGCAGACCTGCTGGGTCATCCTGCTCCTGCCCCTGCTCGGGGCAGCCAGCACCGAGCTGCCCACCCCTG

GCACCGACCCACCACAGCTCACGCCGACCTACGCCCAGGCCCTGGCCACGGCCGTCGACGTCTACAACCAAGGGC

CCGGCGTGGACTTCGCCTTCCGGCTCCTGGAGGCAGAGTCCCGGGACGACTGGGACGCGAGCACGGATCCCCTG

CGGCAGCTGGAGTTCACCCTGAAGGAGACCGAGTGCCCCGTGGGCGAGGACCAGCCCCTGGACCAGTGCGACTT

CAAGGATGGCGGGGCGGTGCTGGACTGCACAGGGACCTTCTCCTGCTCCGAGGCCTCGCTCATGGTGCTGGTCAC

CTGCCAACCCGCCGAGCCCCTGCCTGACCGCGTC**CGCAGGGGTCTATTCAAGAAGCTGAGGAGGAAAATCAAGA**

AGGGCTTTAAGAAAATCTTCAAGCGCCTGCCTCCCGTCGGTGTCGGTGTCTCCATCCCACTCGCAGGAAGGCGGT

GAgatccagacatgataagatacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctatt

gctttatttgtaaccattataagctgcaataaacaagttCCG**AGG**TCTTCATCACGCTAGGGTTGGTCAGCCTTCTGGAGAACATC

HA4-F→

CTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGCCCATGTACTTCTTCATCTGCAGCCTGGCGGTGG

HA4 300 bp

CCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTGTGATGGCGCTGATCACCAGCGGCAACCTGACC

ATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACTCCATGATCTGCAGCTCACTCCTGGCCTCCATTT

GGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTCTACGCCCTGCGCTACCACAACATCATGACCCA

ACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCACGGCGTCCGGTGTGCTCTTCATCATCTACTCG

GAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACCATGCTGGCCCTCATGGCCTCGCTTTACGTGCA

HA4-R

CATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTTACCGGGGAACGGCCCCGTGTGGCAGGCGGC

CAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGTTTGTCGTGTGCTGGGCGCCGTTTTTTCTCCAC

CTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCTTCATGTCTCAC

#3 Knock in the *Cec* at the *mstn* exon1 (dcPlasmid: pUC57_*Cec*)

> *mstn* gene sequence (5'-3')

intron1 338 bp

GTAGTGGTAGTGGAGTGGTAGTGTAATGTAGTGTAGTGGAGAAAGTTGTGGGTCTGTCTCTTTAAGGTTTCAGCG

CTGGAAAGGGAGGAAAAAAATCCGGACTGAAGTCCACCTCTGATTTATTGTTGCTCCGAGTAGCCAATCATAGATT

TCGACGCCAGAGCCTAAATAAGAGCGGCGGAATAATTTGGCGGTATAAAAAGGCTTTTGGGCGAATTGAAGCAT

mstn-F1

GACATCTCGCGCTACCTGTCCGGTGTGCATGGCGCACGGTGTTCCTGTTACTGCTGCCACACAGAAAACACAACCG

exon1 421 bp

Target sequence of sgRNA3 at sense strand

CGCGCGCACTCCTCTCTGAGACCTGACCTGGCTGATCATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTTCGTG

PAM

GTGGCGTTCGGTCCGATGGCGCGCACTGACACCGGAGCACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGC

CGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGCCAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCC

GTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTGCGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGG

mstn-R1

TCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAGCTGCTCGACCTGTACGACGTGCTCGGGGACGACGGC

AAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGACGACGAGGAGCACGCCACCACCGAGACCGTCATGAGCA

intron2 1750 bp

TGGCCGCCGAGCCCGCCGAGCGTGAGTCCCTTTACTACTGCT//TTCAGTAAGTTGTTATAGAGTATTGTGAGGAG

TGTGAGACTTAAACTGACAGATCGAGGAGTTTAAGGGGTTAATTTGTGCTCTGTGTGCAGCTTAAGCAAGTGTCAC

mstn-F2

ACACCGATCCCACAATGCATCAGTTCGTGTACCGTCTTCATTCGATTACTAGTGCAGCTAGCTAACCGTGTTCGGTT

GTGTGTAGTTAGAAGAATAAGGAAGGCGAGTCTGAATACAGGGCTTTACAGTGCTGTAACATTTAACCCCATGTT

GTCTCGGATACCTTTTAAATATATAATCTACTCCTGTTTTCTATTCGCTGAATAATTCTCCTCCTGGTCTCTCCCCCTC

Target sequence of sgRNA3 at anti-sense

exon2 371 bp

TCTCCTTTAGCCAACCCCGACGTTCAAGTCGACCAAAAACCGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCA

PAM

AGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGTGCACTTGCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGC

AGATATCGCGACTCATGCCCATCAAAGACGGGAGAAGGCACGTACGAATACGTTCGCTGAAGATCGACGTGGAC

GCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGAAGCAGGTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAA

mstn-R2

CTGGGGGATTGAGATCAACGCGTTCGACTCCAAAAGCAACGATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGA

intron3 306 bp

GGGACTGGTGAGTGTGGATTATTGATATGTATTTGAC//TCTCAAAGAGTCATGCTTTTGTTTTTCCAATCCCAGCTC

CCGTTCTTGGAGGTGAAAATTTCTGAAGTTCCAAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAAT

TCGTCCGAGTCCCGCTGCTGCCGCTACCCCCTTACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCC

exon3 381 bp

CAAAACGCTACAAGGCCAACTACTGCTCGGGCGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACT

TGGTGAACAAGGCCAACCCACGTGGCACTGCCGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCT

CTACTTCAACGGAAAAGAGCAGATCATCTACGGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCGTG

A

> full *Cec* CDS sequence (5'-3')

CDS1 99 bp

ATGAACTTCAACAAGATCTTCGTCTTTGTGGCACTCATCCTGGCCATCAGCCTGGGAAACTCAGAGGCTGGTTGGC

start codon

CDS2 93 bp

TTAGGAAGCTGGGAAAAAAAATCGAACGCATTGGTCAGCATACCAGGGATGCCTCAATCCAGGTCCTCGGAATCG

CCCAACAGGCCGCCAATGTTGCAGCCACCGCTCGAGGTTGA

stop codon

> sg_HA5_UBI_*Cec*_pA_HA6_sg construct at the *mstn* exon1 (pUC57_*Cec*)

Upstream of pUC57 vector 415

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACTGTCACAGCTTGTCTGTAAGC

GGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTAT

GCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAA

ATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCT

ATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCAC

Recognition site of sgRNA3 at sense strand 23 bp

GACGTTGTAAAACGACGGCCAGTGAATTGACGCGTATTGGGATTCTGATTTCGCTGGGCTTCG**TGG**GCGCTGGAA

Left homologous arm (HA5) 300 bp

PAM

AGGGAGGAAAAAAATCCGGACTGAAGTCCACCTCTGATTTATTGTTGCTCCGAGTAGCCAATCATAGATTTCGACGC

CAGAGCCTAAATAAGAGCGGCGGAATAATTTGGCGGTATAAAAAGGCTTTTGGGCGAATTGAAGCATGACATCTCGC

GCTACCTGTCCGGTGTGCATGGCGCACGGTGTTCCTGTTACTGCTGCCACACAGAAAACACAACCGCGCGCGCACTC

CTCTCTGAGACCTGACCTGGCTGATCATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTACCAGCAAAGTTCTAGA

ATTTGTCGAAACATTTATGTTATATATTTCCTGAAAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATA

UBI promoter 1438 bp

AACAACAGACGGCAAAATGAATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATAGGCTGTTCTC

TACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCCCTTTCATTCATATTTCAAGAAAATTGGTTCAGAA

GATCTGGATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCAGTGACCTGATTCCAAATGACTAGGCCTAATCCA

GAAATTAGATGACTGTCAACATAAAAAGGCACAGCACTCACTAGCTGCCCTATATATTTTATTATATTTTACATATA

TTATTTTATTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAGAAAACTGCTTACAACAGCTAACCTGTACTACCT

CAGGCTCAGGGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTTAACCATGCATGCTTGTTTTCAACTATGGCAACA

CAGTCACATGGGACATTACAGAAATGATTTGTCGATGACATGCGACTTTTCTTTAATAAAGCGCAAAGATCCCAAA

AAGCAAACTTTTAACAAAAATCATATAATTATATTTTCAATCCAGCTTTGTAGCAACTTTGTGCTGCTGTTCACTCAG

CAACAGATAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTGCCATCTGTTTCACACATTGCGTTCTATATATAAGT

GTGCTGGTTGACACGACACTGTATAAGGCCTAGGCTAAAACACAAACAATGTAGAATGACACTGTGTTTTTTTTGT

AAACAAATGTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCTCCTGTCATGTATTTGCTATATTCAAATGTTAAAC

CCGTGCAGAATAGAACATATACAAAAAAAAACAACACAACACATTTTTAAACATTATTAAATATCAAGTATTGCTG

GCAGTTCTGTTTCTGTTTTACAGTACCCTTTGCCACAGTTCTCCGCTTTTCCTGGTCCAGATTCCACAAGTCTGATTC

ACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGACTGTCCTGCGAGACCGGCCCATTCC

AGCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATTGATACAAAACATGTAATTCCTAGA

ACATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGAAAGATCTGGTCATGT

GACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAGTAGGACTCTCGACCATCGGCAGATTTTTCGAAG

AAGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGAACTTCAACAAGATCTTCGTCTTTGTGGCACTCATCCTG

Cec_CDS 192 bp

GCCATCAGCCTGGGAAACTCAGAGGCTGGTTGGCTTAGGAAGCTGGGAAAAAAAATCGAACGCATTGGTCAGCA

TACCAGGGATGCCTCAATCCAGGTCCTCGGAATCGCCCAACAGGCCGCCAATGTTGCAGCCACCGCTCGAGGTTG

Agatcttttccctctgccaaaaattatggggacatcatgaagccccttgagcatctgacttctggctaataaaggaaatttattttcattgcaatagtgt

PolyA tail 288

gttggaatttttttgtgtctctcactcggaaggacatatgggagggcaaatcatttaaaacatcagaatgagtatttggtttagagtttggcaacatatgc

catatgctggctgccatgaacaaaggtggctataaagaggtcatcagtatatgaaacagccccctgctgtccattccttattccatagTCGTGGTG

GCGTTCGGTCCGATGGCGCGCACTGACACCGGAGCACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGT

GACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGCCAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTC

Right homologous arm (HA6) 300

TGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTGCGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCA

AGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAGCTGCTCGACCTGTACGACGTGCTCGGGGACGACGCCACG

Recognition site of sgRNA3 at anti-sense strand 23

PAM

AAGCCCAGCGAAATCAGAATCCCAATGGCGCGCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAA

ATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGT

GAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAAT

Downstream of pUC57 vector 2256 bp

GAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTGTTCCGCTTCCTCGCTCACTGACTCGCTGC

GCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG

ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGC//TTGAGATCCAGTTCGATGTAACCCACTCGT

GCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCG

CAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTAT

CAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACAT

TTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACG

AGGCCCTTTTGTC

> positive fish sequence after insertion at the *mstn* exon1 (5'-3')

TGTAGTGGAGTGGTAGTGTAGTGGTAGTGGAGCGGTAGTGGAGCGGTAGTGTAATGTAGTGGAGCGGTAGTGT

Upstream of mstn 200 bp

AATGTAGTGTAGTGGTAGTGGAGCGGTAGTGTAGTGGTAGTGGAGCGGTAGTGTAATGTAGTGTAGTGGTAGTG

GTAGTGTAGTGGAGTGGTAGTGGAGTGGTAGTGTAATGTAGTGGAGTGGTAGTGTAGTGGTAGTGGAGTGGTA

Partial sequence of mstn intron1 72 bp

GTGTAATGTAGTGTAGTGGAGAAAGTTGTGGGTCTGTCTCTTTAAGGTTTCAGCGCTGGAAAGGGAGGAAAAAAA

HA5-F →

TCCGGACTGAAGTCCACCTCTGATTTATTGTTGCTCCGAGTAGCCAATCATAGATTTCGACGCCAGAGCCTAAATAAG

HA5 300 bp

AGCGGCGGAATAATTTGGCGGTATAAAAAGGCTTTTGGGCGAATTGAAGCATGACATCTCGCGCTACCTGTCCGGTG

TGCATGGCGCACGGTGTTCCTGTTACTGCTGCCACACAGAAAACACAACCGCGCGCGCACTCCTCTCTGAGACCTGA

CCTGGCTGATCATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTACCAGCAAAGTTCTAGAATTTGTCGAAACATTT

← HA5-R

ATGTTATATATTTCCTGAAAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATAAACAACAGACGGCAA

UBI promoter 1438 bp

AATGAATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATAGGCTGTTCTCTACAACCCTCAAACAG

TGATTAGTTTTGTACTTATAAACTTGCCCTTTCATTCATATTTCAAGAAAATTGGTTCAGAAGATCTGGATATTCTAG

CAGTTGTTCAAGCTCATGGAGGGATCAGTGACCTGATTCCAAATGACTAGGCCTAATCCAGAAATTAGATGACTGT

CAACATAAAAAGGCACAGCACTCACTAGCTGCCCTATATATTTTATTATATTTTACATATATTATTTTATTTATTTAGC

TCTGAGTGCTGTACTTTCTGGTTAAAGAAAACT//TTCTCCGCTTTTCCTGGTCCAGATTCCACAAGTCTGATTCACC

AATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGACTGTCCTGCGAGACCGGCCCATTCCAG

CACATTCTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATTGATACAAAACATGTAATTCCTAGAAC

ATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGAAATGAACTTCAACAAGA

Cec CDS 192 bp

TCTTCGTCTTTGTGGCACTCATCCTGGCCATCAGCCTGGGAAACTCAGAGGCTGGTTGGCTTAGGAAGCTGGGAAA

AAAAATCGAACGCATTGGTCAGCATACCAGGGATGCCTCAATCCAGGTCCTCGG[AATCGCCCAACAGGCCGCCAA

PolyA tail 288 bp

TGTTGCAGCCACCGCTCGAGGTTGAgatctttttccctctgccaaaaattatggggacatcatgaagccccttgagcatctgacttctggcta

ataaaggaaatttatttttcattgcaatagtgtgttggaatttttttgtgtctctcactcggaaggacatatgggagggcaaatcatttaaaacatcagaat

HA6-F →

gagtatttggtttagagtttggcaacatatgccatatgctggctgccatgaacaaaggtggctataaagaggtcatcagtatatgaaacagccccctgc

tgtccattccttattccatagTCGTGGTGGCGTTCGGTCCGATGGCGCGCACTGACACCGGAGCACCGGAGCAGCAGCAG

HA6 300 bp

CAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGCCAGCGCGTGCGCTTTCCG

CCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTGCGCCTCAAACAAGCTCCC

AACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAGCTGCTCGACCTGTACGA

CGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGACGACGAGGAGCACGCCACC

Partial sequence of mstn exon1 87 bp

ACCGAGACCGTCATGAGCATGGCCGCCGAGCCCGCCGAGCGTGAGTCCCTTTACTACTGCTTCATAGCCTAACTTT

← HA6-R

ATTTATTCATTTCTTTATTTTGGTAATGTGTTTTTTTTTTTTTTTTTTTTGTTTGTTTTTTCTGTCTAATGCATAGCAACA

Partial sequence of mstn intron2 130 bp

TCTG

#4 Knock in the *Cec* at the *mstn* exon2 (dcPlasmid: pUC57_*Cec*)

> sg_HA7_UBI_*Cec*_pA_HA8_sg construct at the *mstn* exon2 (pUC57_*Cec*)

Upstream of pUC57 vector 415

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACTGTCACAGCTTGTCTGTAAGC

GGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTAT

GCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAA

ATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCT

ATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCAC

Recognition site of sgRNA4 at anti-sense strand 23 bp

GACGTTGTAAAACGACGGCCAGTGAATTGACGCGTATTGGGAT**CCC**CGACGTTCAAGTCGACCAAAAGTTTAAGG

Left homologous arm (HA7) 300 bp **PAM**

GGTTAATTTGTGCTCTGTGTGCAGCTTAAGCAAGTGTCACACACCGATCCCACAATGCATCAGTTCGTGTACCGTCT

TCATTCGATTACTAGTGCAGCTAGCTAACCGTGTTCGGTTGTGTGTAGTTAGAAGAATAAGGAAGGCGAGTCTGA

ATACAGGGCTTTACAGTGCTGTAACATTTAACCCCATGTTGTCTCGGATACCTTTTAAATATATAATCTACTCCTGTT

TTCTATTCGCTGAATAATTCTCCTCCTGGTCTCTCCCCCTCTCTCCTTTAGCCAA**CCC**CGAACCAGCAAAGTTCTAGA

ATTTGTCGAAACATTTATGTTATATATTTCCTGAAAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATA

UBI promoter 1438 bp

AACAACAGACGGCAAAATGAATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATAGGCTGTTCTC

TACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCCCTTTCATTCATATTTCAAGAAAATTGGTTCAGAA

GATCTGGATATTCTAGCAGTTGTTCAAGCTCATGGAGGGATCAGTGACCTGATTCCAAATGACTAGGCCTAATCCA

GAAATTAGATGACTGTCAACATAAAAAGGCACAGCACTCACTAGCTGCCCTATATATTTTATTATATTTTACATATA

TTATTTTATTTATTTAGCTCTGAGTGCTGTACTTTCTGGTTAAAGAAAACT//TTCTCCGCTTTTCCTGGTCCAGATTC

CACAAGTCTGATTCACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGACTGTCCTGCGA

GACCGGCCCATTCCAGCACATTCTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATTGATACAAAAC

ATGTAATTCCTAGAACATAACCATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGAA

AGATCTGGTCATGTGACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAATACAGTAGGACTCTCGACCATCGG

CAGATTTTTCGAAGAAGAAGATCAGTTTCAGGAGCCGTACTGTTCCGTTATGAACTTCAACAAGATCTTCGTCTTTG

Cec CDS 192 bp

TGGCACTCATCCTGGCCATCAGCCTGGGAAACTCAGAGGCTGGTTGGCTTAGGAAGCTGGGAAAAAAAATCGAAC

GCATTGGTCAGCATACCAGGGATGCCTCAATCCAGGTCCTCGGAATCGCCCAACAGGCCGCCAATGTTGCAGCCA

PolyA tail 288 bp

CCGCTCGAGGTTGAgatcttttccctctgccaaaaattatggggacatcatgaagccccttgagcatctgacttctggctaataaaggaaatttta

ttttcattgcaatagtgtgttggaatttttttgtgtctctcactcggaaggacatatgggagggcaaatcatttaaaacatcagaatgagtatttggtttag

agtttggcaacatatgccatatgctggctgccatgaacaaaggtggctataaagaggtcatcagtatatgaaacagccccctgctgtccattccttattc

catagCGTTCAAGTCGACCAAAAACCGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGT

Right homologous arm (HA8) 300

AAGGGCGCAGCTCTGGGTGCACTTGCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCAT

GCCCATCAAAGACGGGAGAAGGCACGTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGT

GGCAGAGCATCGACGTGAAGCAGGTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGAT

PAM

CAACGCTTTGGTCGACTTGAACGTCG**GGG**ATCCCAATGGCGCGCCGAGCTTGGCGTAATCATGGTCATAGCTGTTT

Recognition site of sgRNA4 at sense strand 23 bp

CCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAA**//**ATCATTGG

AAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA

CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAA

AAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCA

Downstream of pUC57 vector 2256 bp

GGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTT

CCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGA

GGCCCTTTTGTC

> positive fish sequence after insertion at the *mstn* exon2 (5’-3’)

TCTCAGCTGTATCTGTTACGTTGTGAACCGTCCCGTGTTAATTCAATACATCTCTATATCTTTGTAAATAAACGCTAT

Partial sequence of *mstn* intron2 200 bp

TTAACCTGTAATTGGTAGGAGTTTTGAATCGTTTCAGAATGGACGATTCGTACCTGTTTATTATTCAGTAAGTTGTT

HA7-F→

HA7 300 bp

ATAGAGTATTGTGAGGAGTGTGAGACTTAAACTGACAGATCGAGGAGTTTAAGGGGTTAATTTGTGCTCTGTGTG

CAGCTTAAGCAAGTGTCACACACCGATCCCACAATGCATCAGTTCGTGTACCGTCTTCATTCGATTACTAGTGCAGC

TAGCTAACCGTGTTCGGTTGTGTGTAGTTAGAAGAATAAGGAAGGCGAGTCTGAATACAGGGCTTTACAGTGCTG

TAACATTTAACCCCATGTTGTCTCGGATACCTTTTAAATATATAATCTACTCCTGTTTTCTATTCGCTGAATAATTCTC

CTCCTGGTCTCTCCCCCTCTCTCCTTTAGCCAA**CCCC**GAACCAGCAAAGTTCTAGAATTTGTCGAAACATTTATGTTA

←HA7-R

TATATTTCCTGAAAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATAAACAACAGACGGCAAAATGA

UBI promoter 1438 bp

ATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATAGGCTGTTCTCTACAACCCTCAAACAGTGATT

AGTTTTGTACTTATAAACTTGCCCTTTCATTCATATTTCAAGAAAATTGGTTCAGAAGATCTGGATATTCTAGCAGTT

GTTCAAGCTCATGGAGGGATCAGTGACCTGATTCCAAATGACTAGGCCTAATCCAGAAATTAGATGACTGTCAACA

TAAAAAGGCACAGCACTCACTAGCTGCCCTATATATTTTATTATATTTTACATATATTATTTTATTTATTTAGCTCTGA

GTGCTGTACTTTCTGGTTAAAGAAAACT**//**TTCTCCGCTTTTCCTGGTCCAGATTCCACAAGTCTGATTCACCAATAG

CAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGACTGTCCTGCGAGACCGGCCCATTCCAGCACATT

CTGGAAACTTCCTTTATATGATAATTATAAATACATTTAAATTATTGATACAAAACATGTAATTCCTAGAACATAACC

ATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGAAATGAACTTCAACAAGATCTTCG

Cec CDS 192 bp

TCTTTGTGGCACTCATCCTGGCCATCAGCCTGGGAAACTCAGAGGCTGGTTGGCTTAGGAAGCTGGGAAAAAAAA

TCGAACGCATTGGTCAGCATACCAGGGATGCCTCAATCCAGGTCCTCGGAATCGCCCAACAGGCCGCCAATGTTG

CAGCCACCGCTCGAGGTTGAgatcttttccctctgccaaaaattatggggacatcatgaagccccttgagcatctgacttctggctaataaag

PolyA tail 288 bp

gaaatttattttcattgcaatagtgtgttggaatttttgtgtctctcactcggaaggacatatgggagggcaaatcatttaaaacatcagaatgagtatt

tggtttagagtttggcaacatatgccatatgctggctgccatgaacaaaggtggctataaagaggtcatcagtatatgaaacagccccctgctgtccat

tccttattccatagCGTTCAAGTCGACCAAAAACCGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCG

HA8-F →

CATCGTAAGGGCGCAGCTCTGGGTGCACTTGCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCG

HA8 300 bp

ACTCATGCCCATCAAAGACGGGAGAAGGCACGTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCA

GTTCGTGGCAGAGCATCGACGTGAAGCAGGTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATT

GAGATCAACGCGTTCGACTCCAAAAGCAACGATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGGT

Partial sequence of mstn exon2 61 bp

GAGTGTGGATTATTGATATGTATTTGACGTTTAACTGAACGTACATCCTAAAGCCCGTTCCAAGCGAACGAAACGT

Partial sequence of mstn intron3 200 bp

TAAAAAGAAGACTCTAGTATCCGTGTTCTCGGTGTGCCGTGACACTCTTTGCTGGTTGAAATGGGATTATTTGAAG

← HA8-R

TTAGCAAAAGGTAGTTAATACCAATAAAATCTCCTAGATAACTCGA

Appendix 8: Predicted the 3D structure of mutant proteins using AlphaFold

#1 Mutations from ssGE1 (Sequence, AA and predicted protein structure)

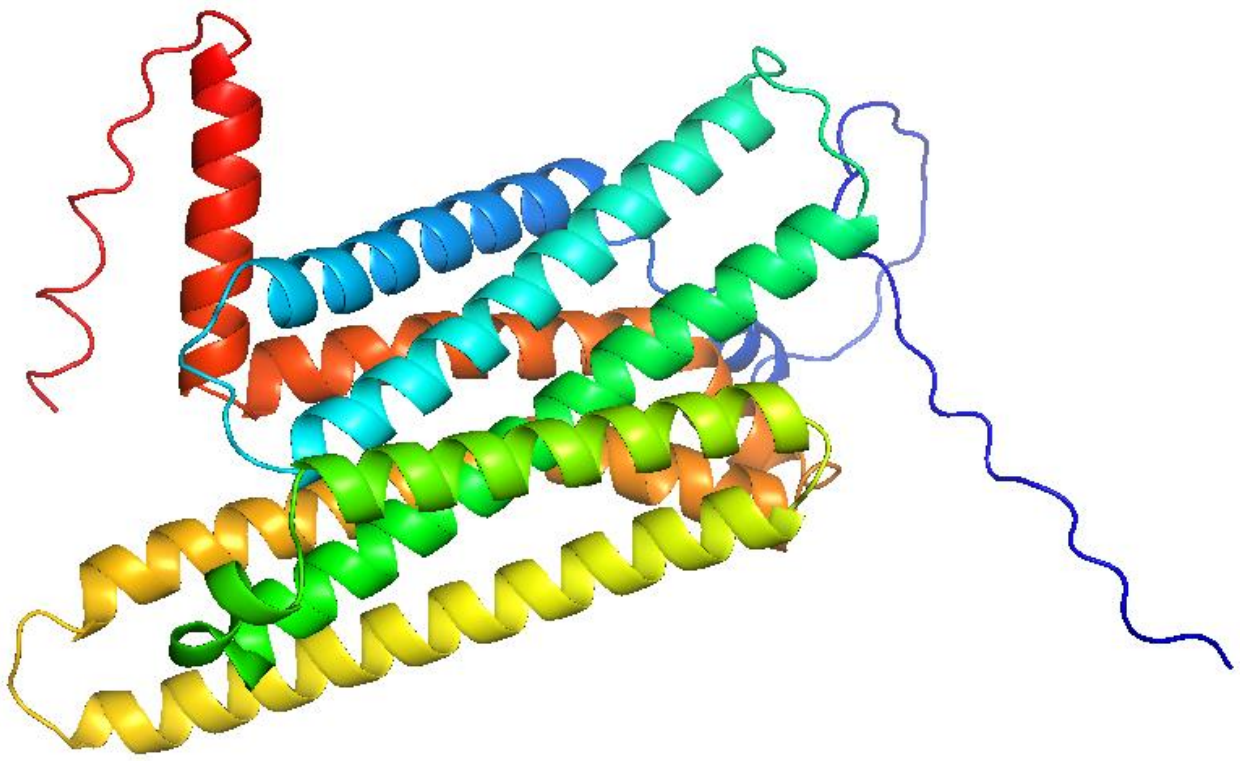
***mc4r* CDS sequence for AA:**

WT CDS:

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCTCCACCG**AGG**TCTTCATCAC
GCTAGGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGCCC
ATGTACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTGTG
ATGGCGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACTCC
ATGATCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTCTA
CGCCCTGCGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCACG
GCGTCCGGTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACCAT
GCTGGCCCTCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTTAC
CGGGGAACGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGTTT
GTCGTGTGCTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCTTC
ATGTCTCACTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGGAG
TCAGGAGATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTGCG
TCGGCTTCGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLISTEVFITLGLVSLLENILVIAAIVKNKNFHSPMYFFI
CSLAVADLLVSVSNATETAVMALITSGNLTISGDVVKSMDNVFDSMICSSLLASIWSLLAIAVDRYVTIFYALRYHNIMTQ
RRAALIIVCIWSFCTASGVLFIIYSESATVLICLISMFFTMLALMASLYVHMFLLARLHMKRIAALPGNGPVWQAANMKG
AVTLTILLGVFVVCWAPFFLHLILMISCPRNPYCVCFMSHFNMYLILIMCNSVIDPLIYAFRSQEMRKTFREICCGWASG
WSCGWSCVGFDERLNSY

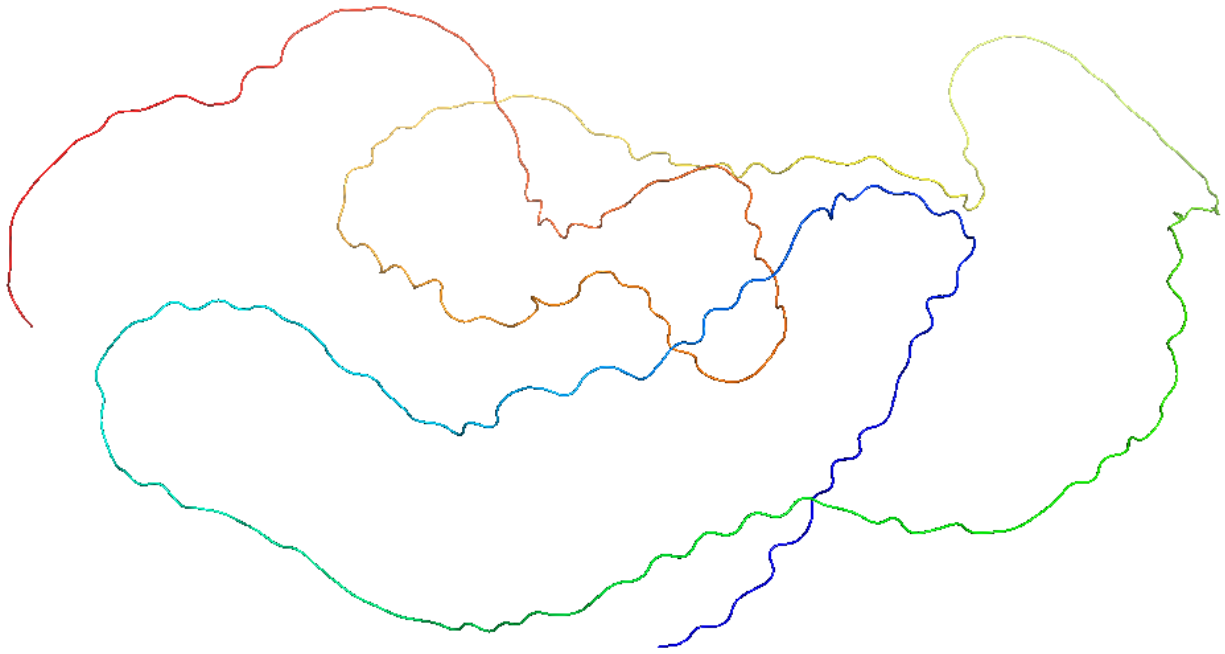


ssGE1-MUTATION#32, #38: (mc4r-mutation-ssGE1 #32, #38) +1G, -2TC

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCCAGCCG**AGG**TCTTCATCAC
GCTAGGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGCCC
ATGTACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTGTG
ATGGCGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACTCC
ATGATCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTCTA
CGCCCTGCGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCACG
GCGTCCGGTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACCAT
GCTGGCCCTCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTTAC
CGGGGAACGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGTTT
GTCGTGTGCTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCTTC
ATGTCTCACTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGGAG
TCAGGAGATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTGCG
TCGGCTTCGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLIQPRSSSRGWSAFWRTSWSRPSSRTRTS
TRPCTSSSAAWRWPTCWAYRTRQKRLWRSPAATPSLETSKAWTMCSTPSAAHSWPPFGVSWPSPWTATSP
SSTPCATTTSPNAARRSSSYAYGASARRPVCSSSSTRRALQSSSALSACSSPCWPSWPRFTCTCSSWRGF
TNASPPYRGTAPCGRRPTRAPRSPSCSECLSCAGRRFFSTSFSSLVRGTRIASASCLTSTCTFSCATRST
RSSTRSGVRRGRPSGRSAAAGLRDGAAAGVASASTRGLTAI

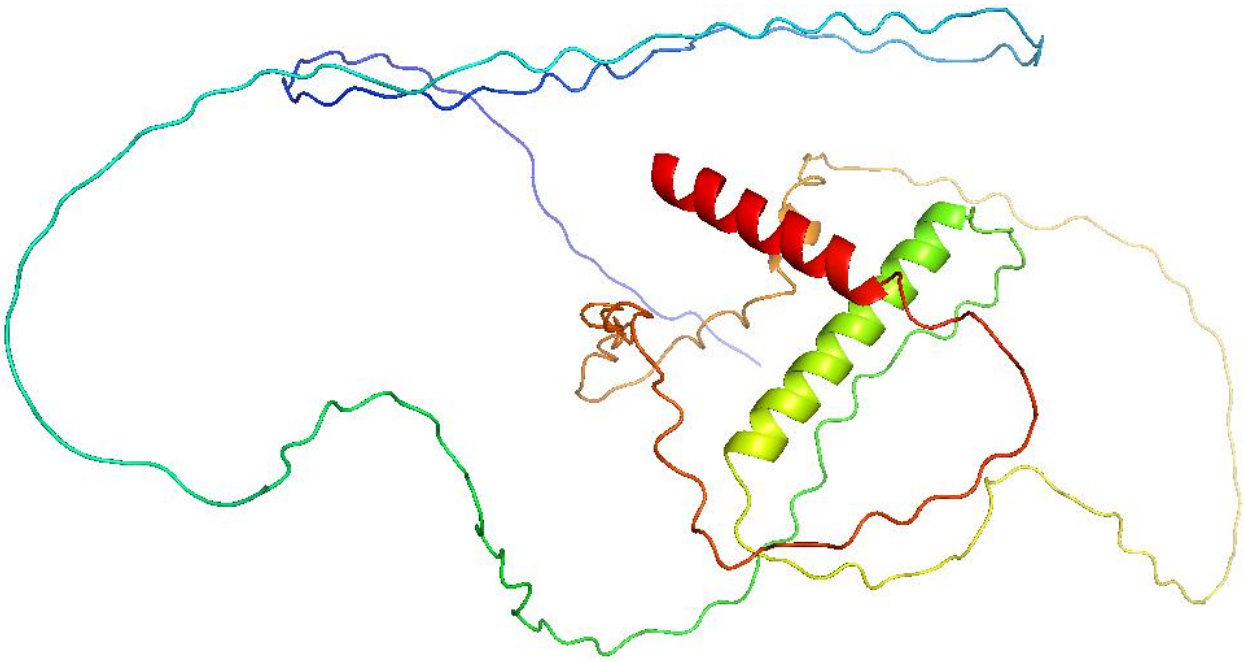


ssGE1-MUTATION#33, #41: (mc4r-mutation-ssGE1, #33, #41) +1G

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCTCCAGCCG**AGG**TCTTCATC
ACGCTAGGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGC
CCATGTACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTG
TGATGGCGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACT
CCATGATCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTC
TACGCCCTGCGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCA
CGGCGTCCGGTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACC
ATGCTGGCCCTCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTT
ACCGGGGAACGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGT
TTGTCGTGTGCTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCT
TCATGTCTCACTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGG
AGTCAGGAGATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTG
CGTCGGCTTCGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLISSRGLHHARVGQPSGEHPGNRGHRQEQ
ELPLAHVLLHLQPGGGRPAGERIERDRNGCDGADHQRQPDHLWRRREKHGQCVRLHDLQLTPGLHLESPG
HRRGPLRHHLLRPALPQHHDPTPRGAHHRMHMELLHGVRCALHHLLGERYSPHLPYQHVLHHAGPHGLAL
RAHVPLGAASHETHRRLTGERPRVAGGQHEGRRDAHHPARSVCRVLGAVFSPPHSHDLLSEEPVLRLLHV
SLQHVPDSDHVQLGDRPAHLRVQESGDEEDLPGDLLRLGFGMELRLELRRLRREAQLL

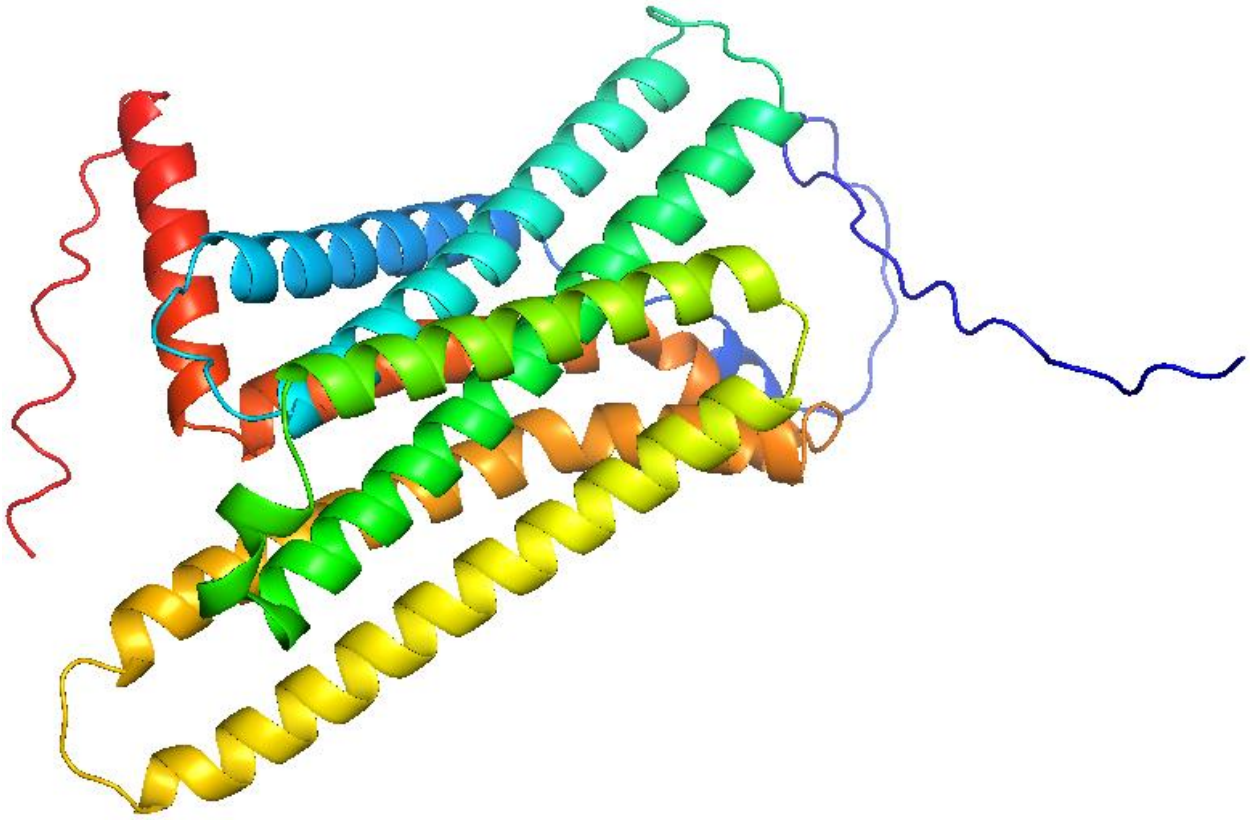


ssGE1-MUTATION#34: (mc4r-mutation-ssGE1, #34) -2/+2 AG/TT

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCTCCACCG**TTG**TCTTCATCAC
GCTAGGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGCCC
ATGTACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTGTG
ATGGCGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACTCC
ATGATCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTCTA
CGCCCTGCGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCACG
GCGTCCGGTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACCAT
GCTGGCCCTCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTTAC
CGGGGAACGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGTTT
GTCGTGTGCTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCTTC
ATGTCTCACTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGGAG
TCAGGAGATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTGCG
TCGGCTTCGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLISTVVFITLGLVSLLENILVIAAIVKNK
NFHSPMYFFICSLAVADLLVSVSNATETAVMALITSGNLTISGDVVKSMDNVFDSMICSSLLASIWSLLA
IAVDRYVTIFYALRYHNIMTQRRAALIIVCIWSFCTASGVLFIIYSESATVLICLISMFFTMLALMASLY
VHMFLLARLHMKRIAALPGNGPVWQAANMKGAVTLTILLGVFVVCWAPFFLHLILMISCPRNPYCVCFMS
HFNMYLILIMCNSVIDPLIYAFRSQEMRKTFREICCGWASGWSCGWSCVGFDERLNSY



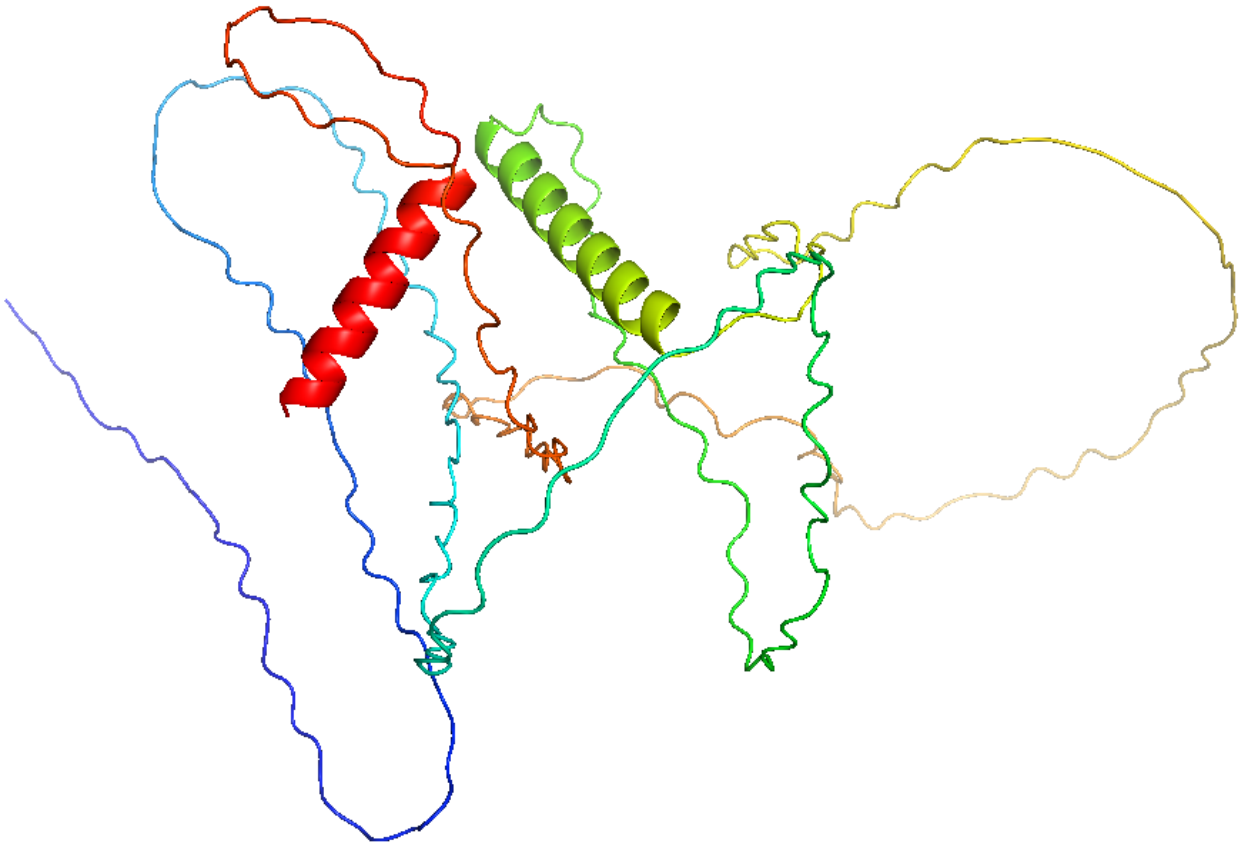
ssGE1-MUTATION#35, #37, #44: (mc4r-mutation-ssGE1 #35, #37, #44) -2 TC

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCCACCG**AGG**TCTTCATCACG
CTAGGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGCCCA
TGTACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTGTGA
TGGCGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACTCCA
TGATCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTCTAC
GCCCTGCGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCACGG
CGTCCGGTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACCATGC
TGGCCCTCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTTACCG
GGGAACGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGTTTGT
CGTGTGCTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCTTCAT
GTCTCACTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGGAGTC
AGGAGATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTGCGTC
GGCTTCGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLIHRGLHHARVGQPSGEHPGNRGHRQEQE
LPLAHVLLHLQPGGGRPAGERIERDRNGCDGADHQRQPDHLWRRREKHGQCVRLHDLQLTPGLHLESPGH
RRGPLRHHLLRPALPQHHDPTPRGAHHRMHMELLHGVRCALHHLLGERYSPHLPYQHVLHHAGPHGLALR

AHVPLGAASHETHRRLTGERPRVAGGQHEGRRDAHHPARSVCRVLGAVFSPPHSHDLLSEEPVLRLLHVS
LQHVPDSDHVQLGDRPAHLRVQESGDEEDLPGDLLRLGFGMELRLELRRLRREAQLL



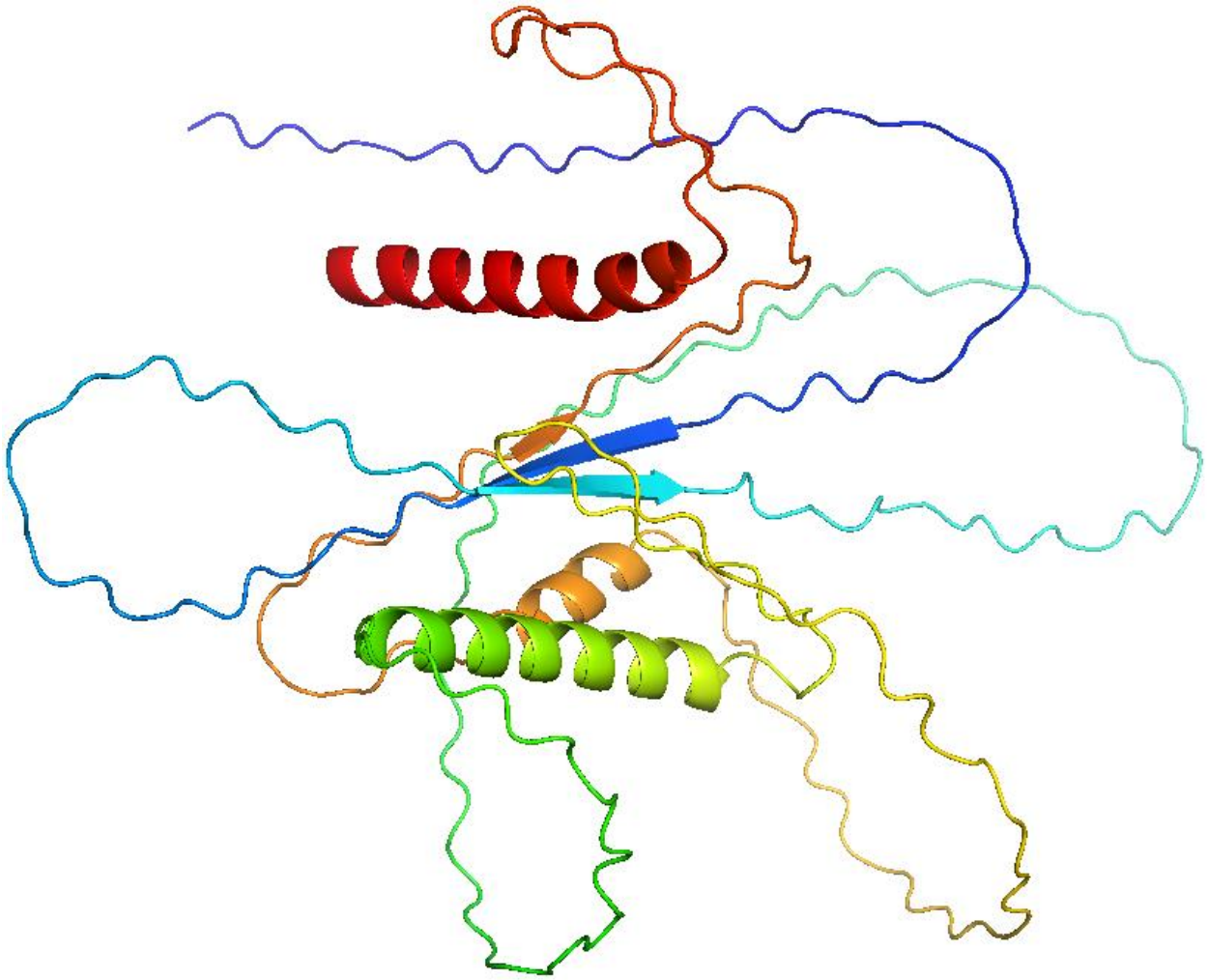
ssGE1-MUTATION#36: (mc4r-mutation-ssGE1, #36) -2CG

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCTCCAC**AGG**TCTTCATCACGC
TAGGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGCCCAT
GTACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTGTGAT
GGCGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACTCCAT
GATCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTCTACG
CCCTGCGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCACGGC
GTCCGGTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACCATGCT
GGCCCTCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTTACCGG
GGAACGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGTTTGTC
GTGTGCTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCTTCATG
TCTCACTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGGAGTCA
GGAGATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTGCGTCG
GCTTCGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLISTGLHHARVGQPSGEHPGNRGHRQEQE
LPLAHVLLHLQPGGGRPAGERIERDRNGCDGADHQRQPDHLWRRREKHGQCVRLHDLQLTPGLHLESPGH

RRGPLRHHLLRPALPQHHDPTPRGAHHRMHMELLHGVRCALHHLLGERYSPHLPYQHVLHHAGPHGLALR
AHVPLGAASHETHRRLTGERPRVAGGQHEGRRDAHHPARSVCRVLGAVFSPPHSHDLLSEEPVLRLLHVS
LQHVPDSDHVQLGDRPAHLRVQESGDEEDLPGDLLRLGFGMELRLELRRLRREAQLL



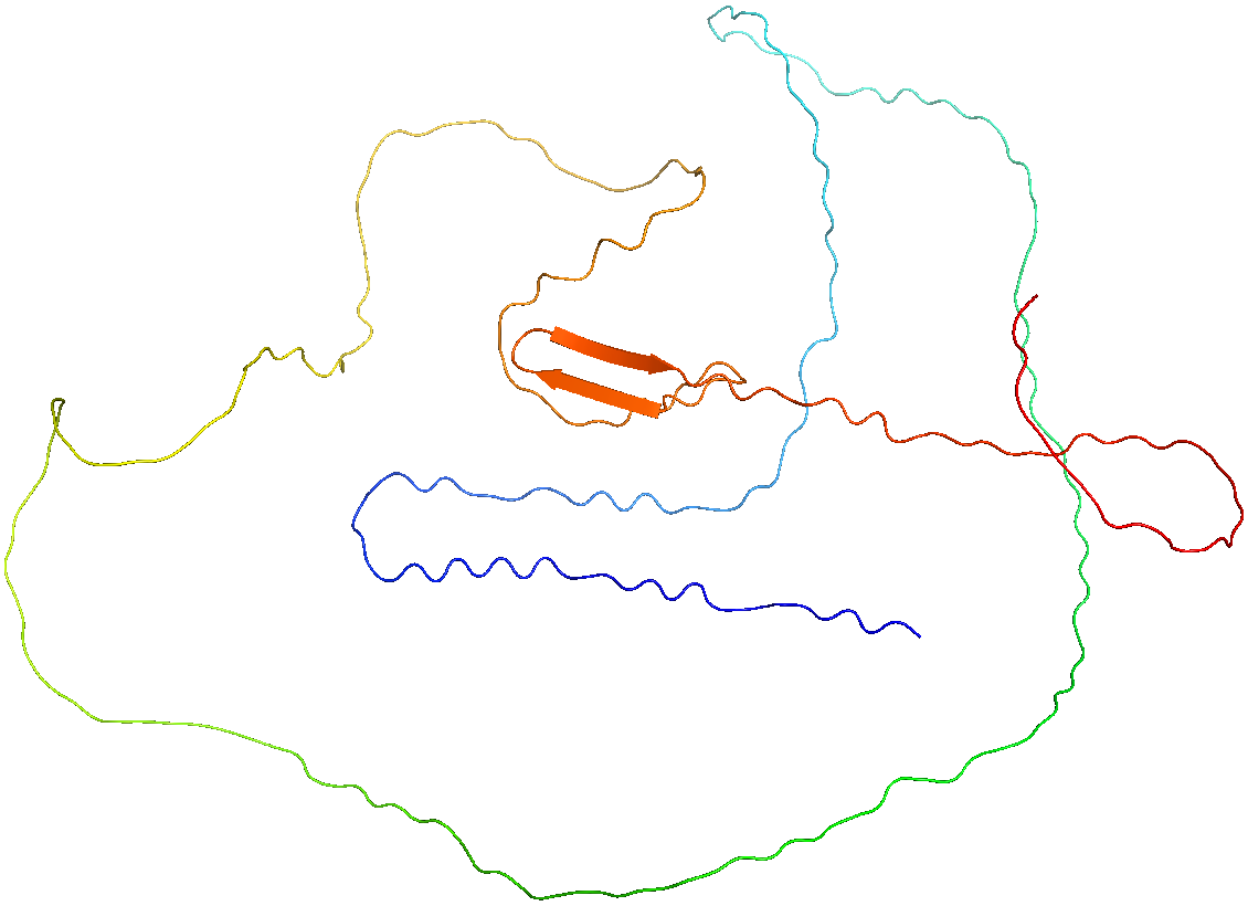
ssGE1-MUTATION#39: (mc4r-mutation-ssGE1, #39) -4 AGGT

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCTCCACCGCTTCATCACGCTA
GGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGCCCATGT
ACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTGTGATGG
CGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACTCCATGA
TCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTCTACGCC
CTGCGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCACGGCGT
CCGGTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACCATGCTG
GCCCTCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTTACCGGG
GAACGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGTTTGTCGT
GTGCTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCTTCATGTC

TCACTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGGAGTCAGG
AGATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTGCGTCGGC
TTCGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLISTASSRGWSAFWRTSWSRPSSRTRTST
RPCTSSSAAWRWPTCWAYRTRQKRLWRSPAATPSLETSKAWTMCSTPSAAHSWPPFGVSWPSPWTATSPS
STPCATTTSPNAARRSSSYAYGASARRPVCSSSSTRRALQSSSALSACSSPCWPSWPRFTCTCSSWRGFT
NASPPYRGTAPCGRRPTRAPRSPSCSECLSCAGRRFFSTSFSSLVRGTRIASASCLTSTCTFSCATRSTR
SSTRSGVRRGRPSGRSAAAGLRDGAAAGVASASTRGLTAI



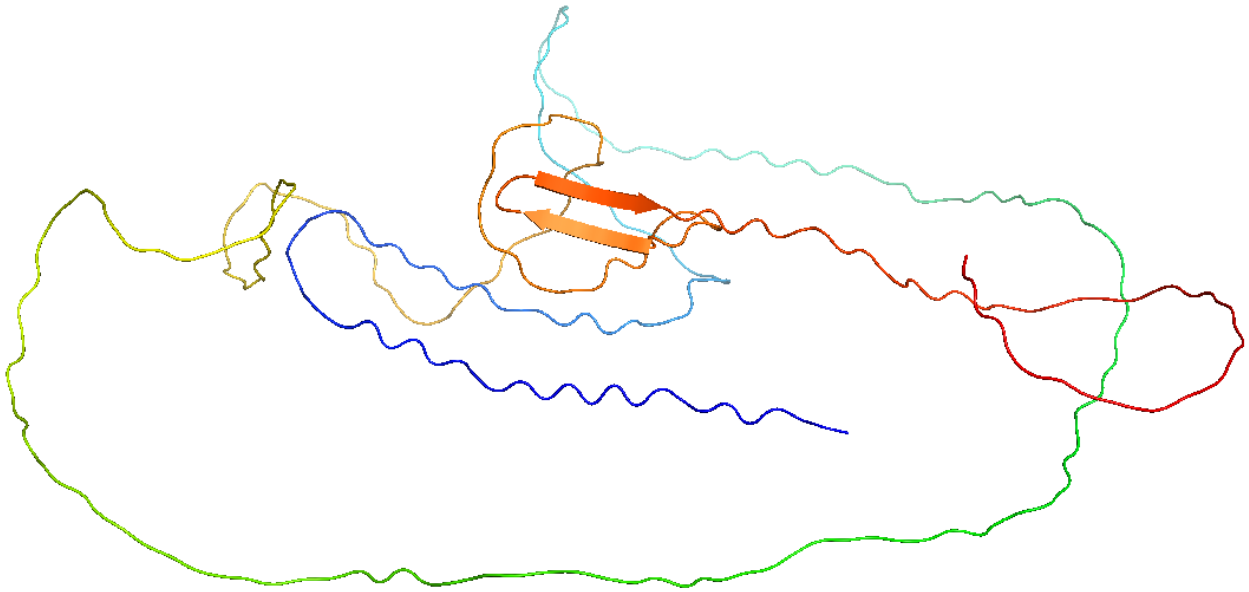
ssGE1-MUTATION#40, #45: (mc4r-mutation-ssGE1, #40, #45) -7 TCCACCG

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATC**AGG**TCTTCATCACGCTAGG
GTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGCCCATGTAC
TTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTGTGATGGCG
CTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACTCCATGATCT
GCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTCTACGCCCTG
CGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCACGGCGTCCG
GTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACCATGCTGGCCC

TCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTTACCGGGGAA
CGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGTTTGTCGTGTG
CTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCTTCATGTCTCA
CTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGGAGTCAGGAG
ATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTGCGTCGGCTT
CGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLIRSSSRGWSAFWRTSWSRPSSRTRTSTR
PCTSSSAAWRWPTCWAYRTRQKRLWRSPAATPSLETSKAWTMCSTPSAAHSWPPFGVSWPSPWTATSPSS
TPCATTTSPNAARRSSSYAYGASARRPVCSSSSTRRALQSSSALSACSSPCWPSWPRFTCTCSSWRGFTN
ASPPYRGTAPCGRRPTRAPRSPSCSECLSCAGRRFFSTSFSSLVRGTRIASASCLTSTCTFSCATRSTRS
STRSGVRRGRPSGRSAAAGLRDGAAAGVASASTRGLTAI



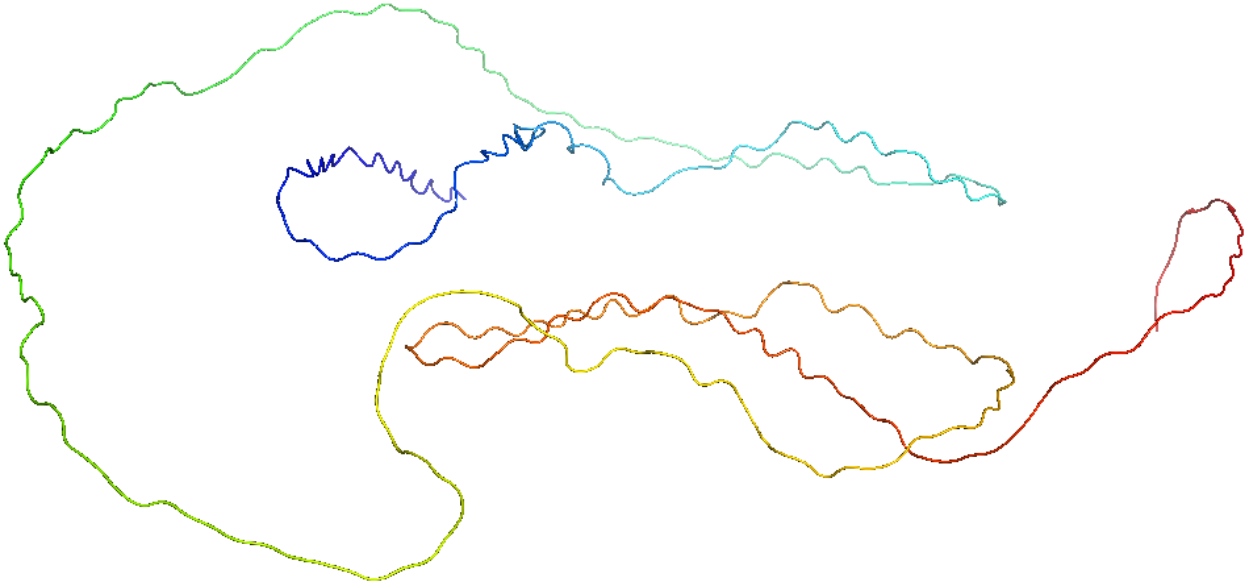
ssGE1-MUTATION#42: (mc4r-mutation-ssGE1, #42) +4 TAGC -2 CG

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCTCCATAGCC**AGG**TCTTCATC
ACGCTAGGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGC
CCATGTACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTG
TGATGGCGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACT
CCATGATCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTC
TACGCCCTGCGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCA
CGGCGTCCGGTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACC
ATGCTGGCCCTCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTT
ACCGGGGAACGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGT
TTGTCGTGTGCTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCT
TCATGTCTCACTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGG

AGTCAGGAGATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTG
CGTCGGCTTCGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLISIARSSSRGWSAFWRTSWSRPSSRTRT
STRPCTSSSAAWRWPTCWAYRTRQKRLWRSPAATPSLETSKAWTMCSTPSAAHSWPPFGVSWPSPWTATS
PSSTPCATTTSPNAARRSSSYAYGASARRPVCSSSSTRRALQSSSALSACSSPCWPSWPRFTCTCSSWRG
FTNASPPYRGTAPCGRRPTRAPRSPSCSECLSCAGRRFFSTSFSSLVRGTRIASASCLTSTCTFSCATRS
TRSSTRSGVRRGRPSGRSAAAGLRDGAAAGVASASTRGLTAI



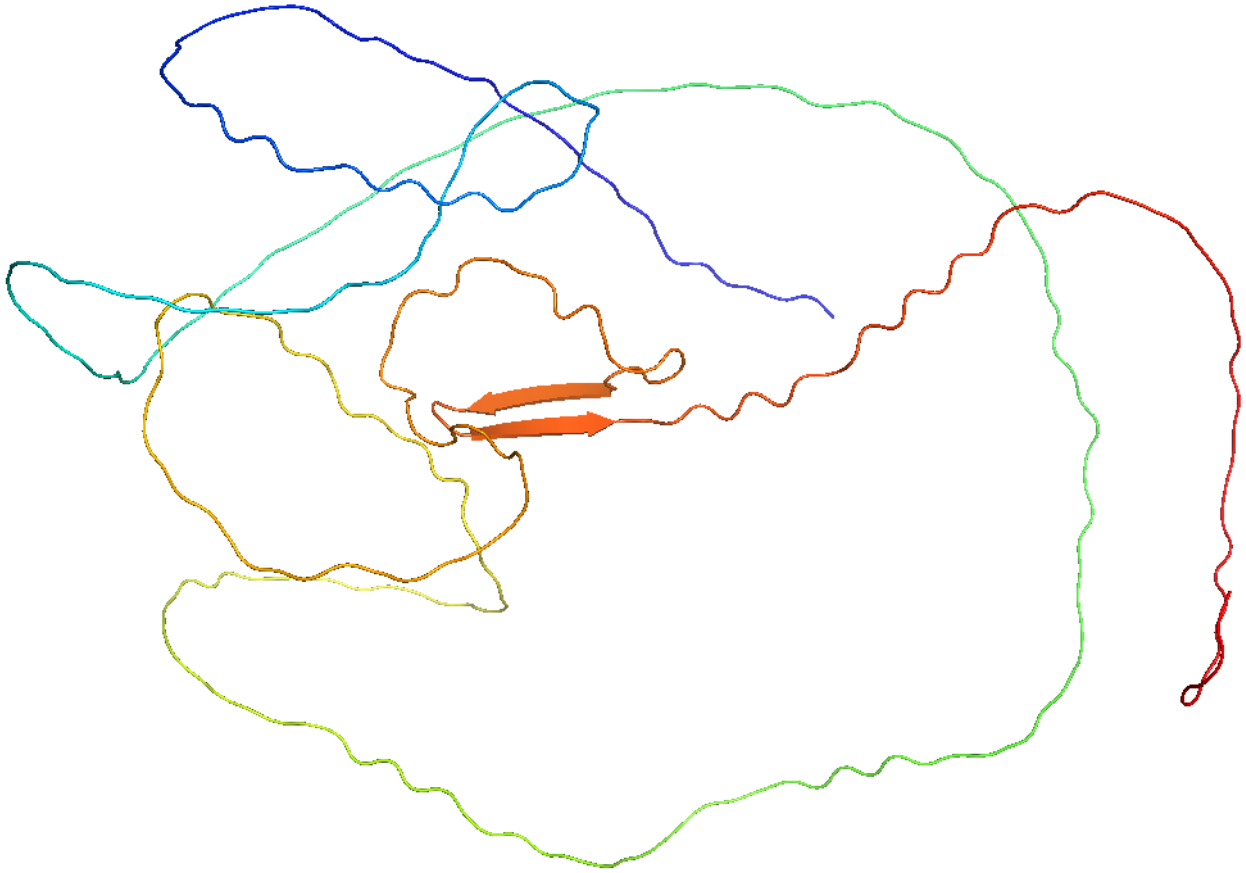
ssGE1-MUTATION #43: (mc4r-mutation-ssGE1, #43) +1 T -2 CT

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCCATCCG**AGG**TCTTCATCAC
GCTAGGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGCCC
ATGTACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTGTG
ATGGCGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACTCC
ATGATCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTCTA
CGCCCTGCGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCACG
GCGTCCGGTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACCAT
GCTGGCCCTCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTTAC
CGGGGAACGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGTTT
GTCGTGTGCTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCTTC
ATGTCTCACTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGGAG
TCAGGAGATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTGCG
TCGGCTTCGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLIHPRSSSRGWSAFWRTSWSRPSSRTRTS
TRPCTSSSAAWRWPTCWAYRTRQKRLWRSPAATPSLETSKAWTMCSTPSAAHSWPPFGVSWPSPWTATSP

SSTPCATTTSPNAARRSSSYAYGASARRPVCSSSSTRRALQSSSALSACSSPCWPSWPRFTCTCSSWRGF
TNASPPYRGTAPCGRRPTRAPRSPSCSECLSCAGRRFFSTSFSSLVRGTRIASASCLTSTCTFSCATRST
RSSTRSGVRRGRPSGRSAAAGLRDGAAAGVASASTRGLTAI

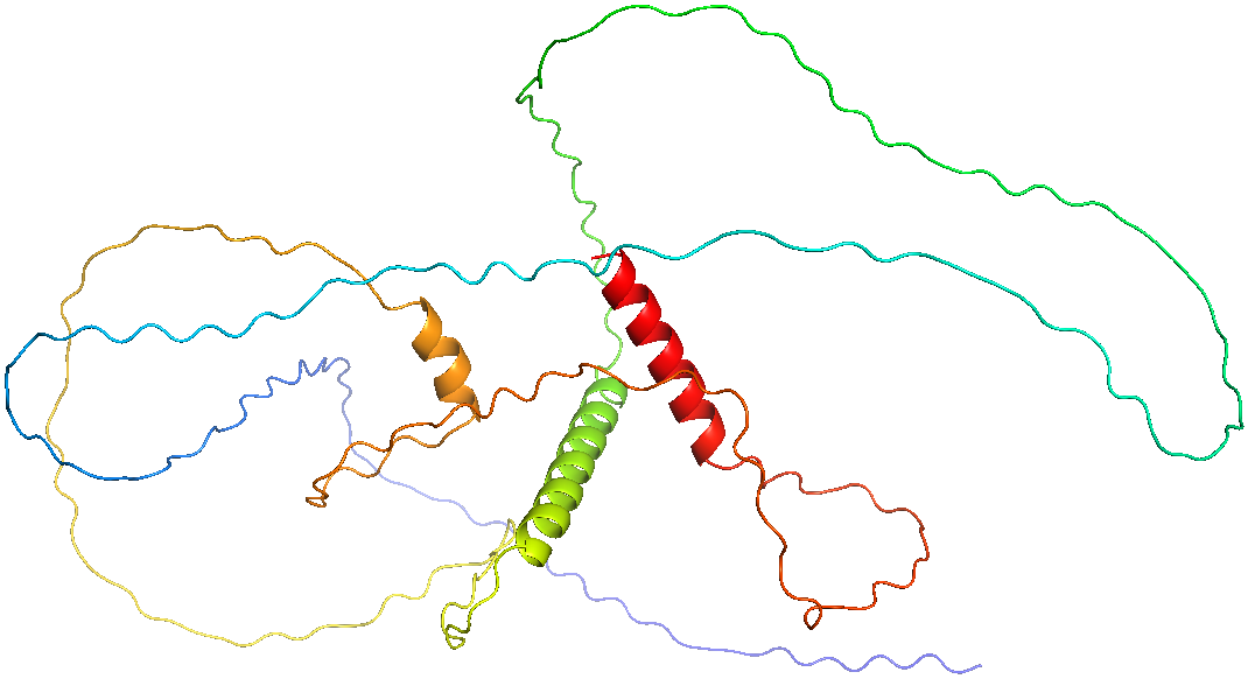


ssGE1-MUTATION#46: (mc4r-mutation-ssGE1, #46) +1 G

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCTCCGACCG**AGG**TCTTCATC
ACGCTAGGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGC
CCATGTACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTG
TGATGGCGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACT
CCATGATCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTC
TACGCCCTGCGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCA
CGGCGTCCGGTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACC
ATGCTGGCCCTCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTT
ACCGGGGAACGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGT
TTGTCGTGTGCTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCT
TCATGTCTCACTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGG
AGTCAGGAGATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTG
CGTCGGCTTCGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLISDRGLHHARVGQPSGEHPGNRGHRQEQ
ELPLAHVLLHLQPGGGRPAGERIERDRNGCDGADHQRQPDHLWRRREKHGQCVRLHDLQLTPGLHLESPG
HRRGPLRHHLLRPALPQHHDPTPRGAHHRMHMELLHGVRCALHHLLGERYSPHLPYQHVLHHAGPHGLAL
RAHVPLGAASHETHRRLTGERPRVAGGQHEGRRDAHHPARSVCRVLGAVFSPPHSHDLLSEEPVLRLLHV
SLQHVPDSDHVQLGDRPAHLRVQESGDEEDLPGDLLRLGFGMELRLELRRLRREAQLL



#2 Mutations from ssGE2 (Sequence, AA and predicted protein structure)

Mstn CDS for AA:

WT:

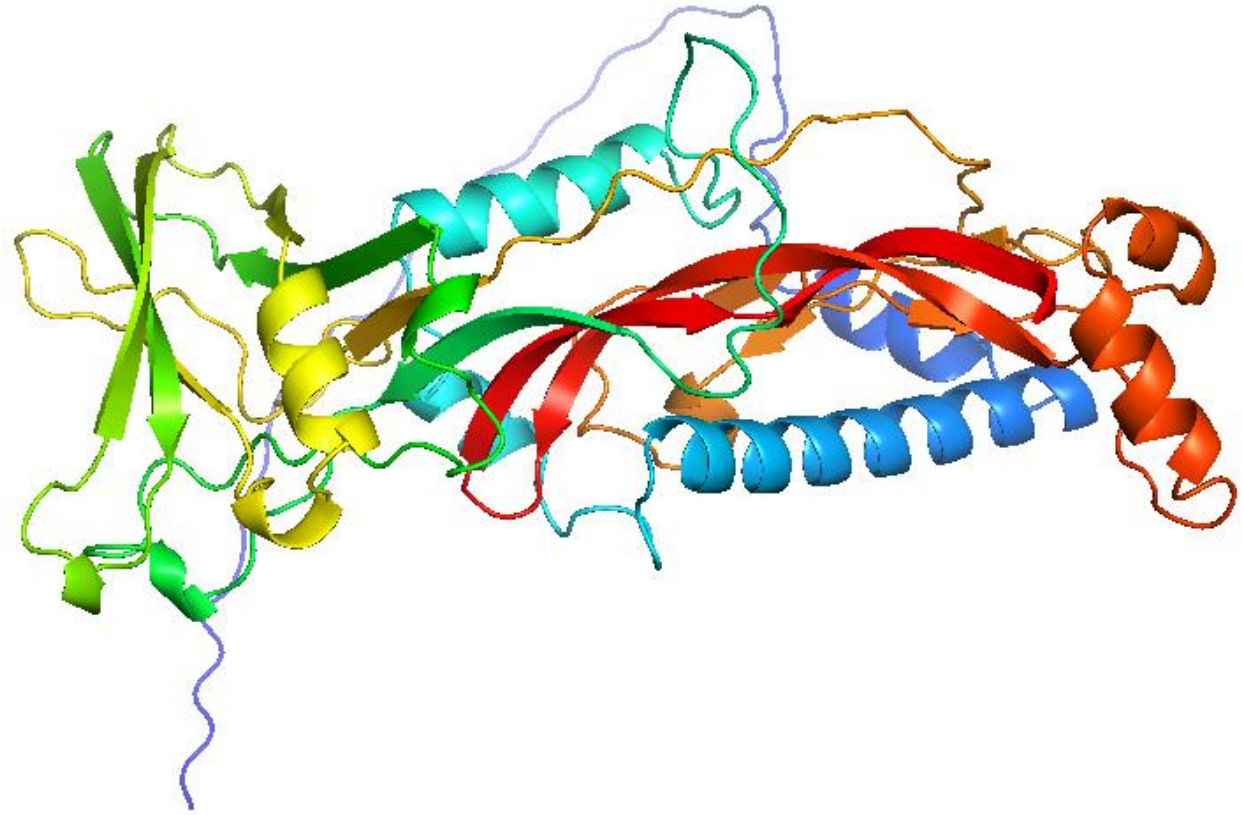
ATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTTCG**TGG**TGGCGTTCGGTCCGATGGCGCGCACTGACACCGGA
GCACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGG
CCAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACT
GCGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCA
GCTGCTCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGAC
GACGAGGAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGAGCCCAA**CCC**CGACGTTCAAGTCGACCA
AAAACCGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGT
GCACTTGCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAG
AAGGCACGTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGA
AGCAGGTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAA
AGCAACGATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCT
GAAGTTCCAAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGC
TACCCCCTTACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACT
GCTCGGGCGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTG

GCACTGCCGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGAT
CATCTACGGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

AA: >XP_017324606.1

MHLAQVLISLGFVVAFGPMARTDTGAPEQQQQQQQPTAVTEEREAQCSAASACAFRQHSKQLRLQAIKSQILSKLRLK
QAPNVSRDVVKQLLPKAPPVQQLLDLYDVLGDDGKPGTALQDEEEDDEEHATTETVMSMAAEPNPDVQVDQKPKCC
FFSFSPKIQASRIVRAQLWVHLRPADEATTVFLQISRLMPIKDGRRHVRIRSLKIDVDAGVSSWQSIDVKQVLAVWLRQP
ETNWGIEINAFDSKSNDLAITSAEPGEEGLLPFLEVKISEVPKRTRRESGLDCDENSSESRCCRYPLTVDFEDFGWDWIIAP
KRYKANYCSGECDYVHLQKYPHTHLVNKANPRGTAGPCCTPTKMSPINMLYFNGKEQIIYGKIPSMVVDRCGCS

MHLAQVLISLGLAFGPMARTDTGAPEQQQQQQQPTAVTEEREAQCSAASACAFRQHSKQLRLQAIKSQILSKLRLKQA
PNVSRDVVKQLLPKAPPVQQLLDLYDVLGDDGKPGTALQDEEEDDEEHATTETVMSMAAEPNPDVQVDQKPKCCFF
SFSPKIQASRIVRAQLWVHLRPADEATTVFLQISRLMPIKDGRRHVRIRSLKIDVDAGVSSWQSIDVKQVLAVWLRQPET
NWGIEINAFDSKSNDLAITSAEPGEEGLLPFLEVKISEVPKRTRRESGLDCDENSSESRCCRYPLTVDFEDFGWDWIIAPK
RYKANYCSGECDYVHLQKYPHTHLVNKANPRGTAGPCCTPTKMSPINMLYFNGKEQIIYGKIPSMVVDRCGCS



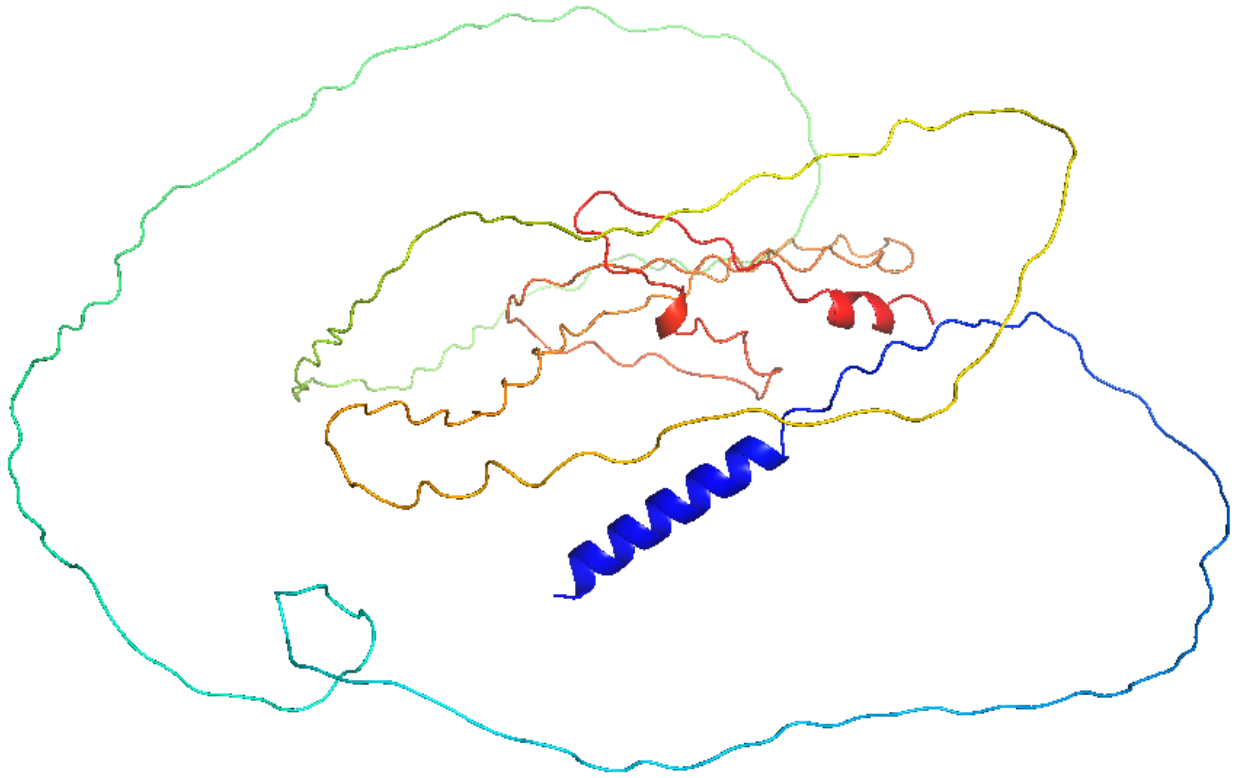
ssGE2-MUTATION#47, #52: (mstn1-mutation-ssGE2, #47, #52)+1 T -5TTCGT

ATGCATTTAGCGCAGGTTCTGATTTCGCTGGTGC**GG**TGGCGTTCGGTCCGATGGCGCGCACTGACACCGGAGCAC
CGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGCCAG
CGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTGCGC
CTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAGCTG

CTCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGACGACG
AGGAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGAGCCCAA**CCC**CGACGTTCAAGTCGACCAAAAAC
CGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGTGCACTT
GCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAGAAGGC
ACGTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGAAGCAG
GTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAAAGCAA
CGATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCTGAAGT
TCCAAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGCTACCC
CCTTACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACTGCTCG
GGCGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTGGCACT
GCCGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGATCATCT
ACGGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLVRWRSVRWRALTPEHRSSSSSSSNLPPRRSARRSVQRPARALSASTASSSVCKPSSPRFA
NCASNKLPTAAMWSSSCCRKRHRCSSCSTCTTCSGTTASRAQRSRTRRRTTRSTPPPRPSAWPPSPTPTF
KSTKNRSAVFSPSARRSKRAASGRSSGCTCARRMRRQRCSCRYRDSCPSKTGEGTYEYVRRSTWTQESVR
GRASTSRCLRCGGSRKPTGGLRSTRSTPKATISRSLLRSLEKRDCSRSWRKFLKFQSEPGENQDTVMRIR
PSPAAAATPLRWTLKTSAGTGLLPQNATRPTTARASATTCTCRSTRIHTWTRPTHVALPAPAARPPRCLP
STCSTSTEKSRSSTARSPPWWIAVAAR



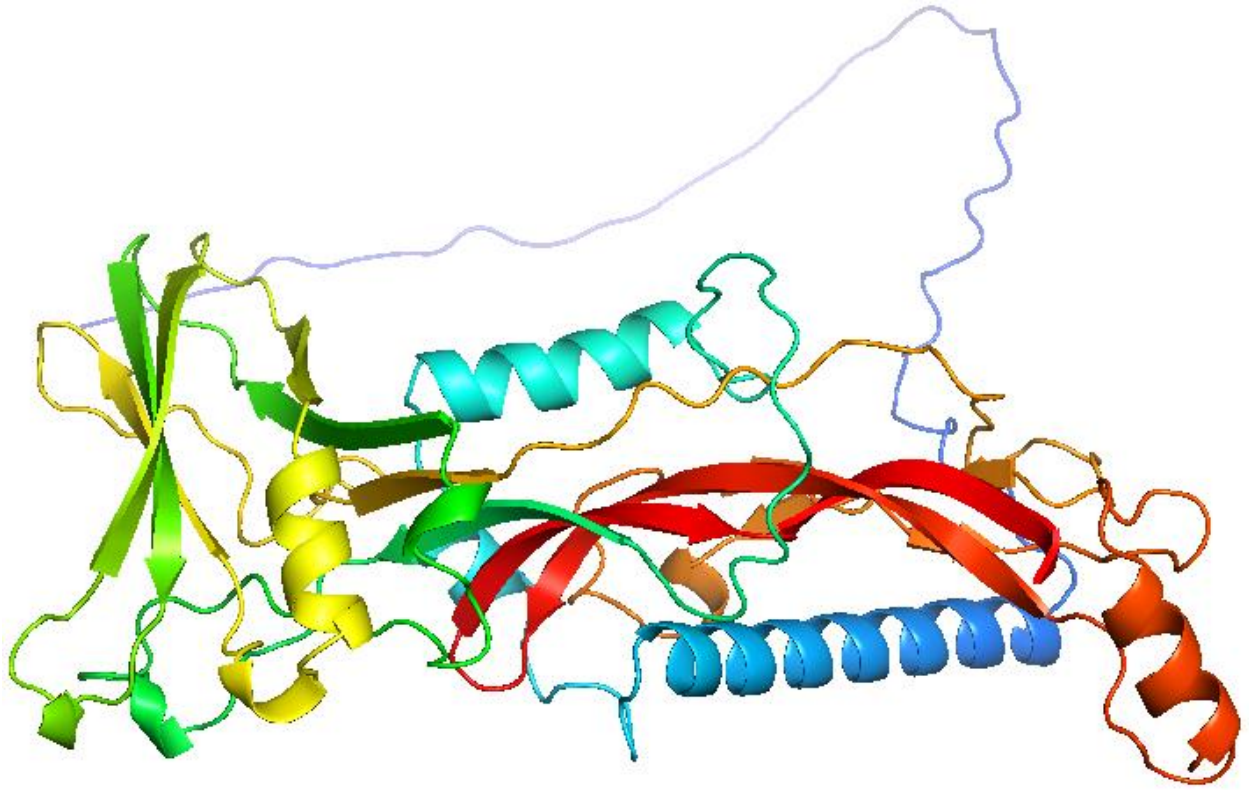
ssGE2-MUTATION#48, #50: (mstn1-mutation-ssGE2 #48, #50) -3 GGC

ATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTTCG**TGG**TGTTCGGTCCGATGGCGCGCACTGACACCGGAGCA
CCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGCCA

GCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTGC
GCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAGC
TGCTCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGACGA
CGAGGAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGAGCCCAA**CCC**CGACGTTCAAGTCGACCAAAA
ACCGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGTGCAC
TTGCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAGAAG
GCACGTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGAAGC
AGGTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAAAGC
AACGATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCTGAA
GTTCCAAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGCTAC
CCCCTTACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACTGCT
CGGGCGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTGGCA
CTGCCGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGATCAT
CTACGGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLGFVVFGPMARTDTGAPEQQQQQQQPTAVTEEREAQCSAASACAFRQHSKQLRLQAIKSQI
LSKLRLKQAPNVSRDVVKQLLPKAPPVQQLLDLYDVLGDDGKPGTALQDEEEDDEEHATTETVMSMAAEP
NPDVQVDQKPKCCFFSFSPKIQASRIVRAQLWVHLRPADEATTVFLQISRLMPIKDGRRHVRIRSLKIDV
DAGVSSWQSIDVKQVLAVWLRQPETNWGIEINAFDSKSNDLAITSAEPGEEGLLPFLEVKISEVPKRTRR
ESGLDCDENSSESRCCRYPLTVDFEDFGWDWIIAPKRYKANYCSGECDYVHLQKYPHTHLVNKANPRGTA
GPCCTPTKMSPINMLYFNGKEQIIYGKIPSMVVDRCGCS

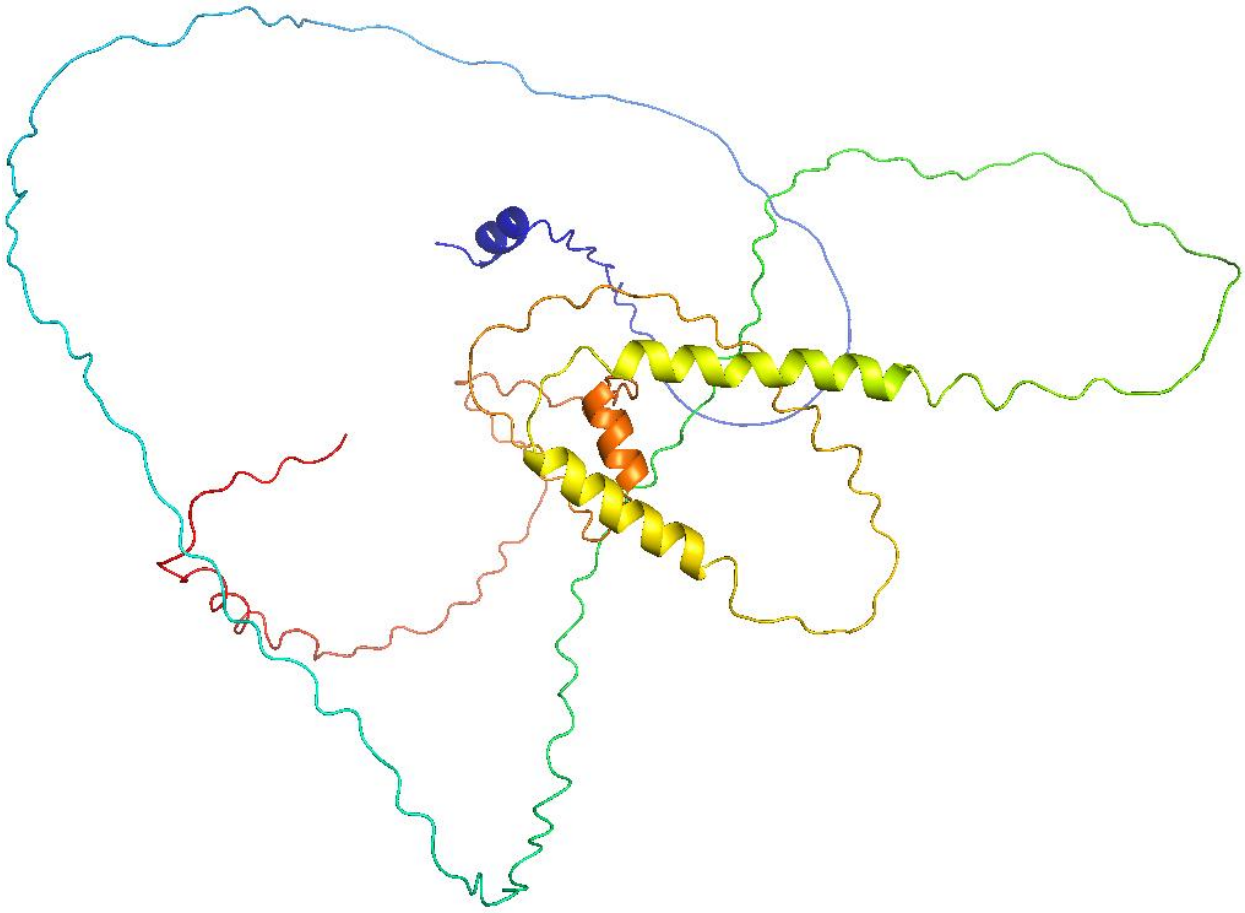


ssGE2-MUTATION#49: (mstn1-mutation-ssGE2 #49) -5 GTGGC

ATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTTCG**TG**GTTCGGTCCGATGGCGCGCACTGACACCGGAGCACC
GGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGCCAGC
GCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTGCGCC
TCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAGCTGC
TCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGACGACGA
GGAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGAGCCCAA**CCC**CGACGTTCAAGTCGACCAAAAACC
GAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGTGCACTTG
CGCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAGAAGGCA
CGTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGAAGCAGG
TGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAAAGCAAC
GATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCTGAAGTT
CCAAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGCTACCCC
CTTACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACTGCTCGG
GCGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTGGCACTG
CCGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGATCATCTA
CGGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLGFVVRSDGAHHRSTGAAAAAAATYRRDGGARGAVFSGQRVRFPPAQQAAPSASHQVPDSE
QTAPQTSSQREPRCGQAAAAESATGAAAARPVRRARGRRQAGHSAPGRGGGRRGARHHRDRHEHGRRAQP
RRSSRPKTEVLFFLLQPEDPSEPHRKGAALGALAPGGGDNGVLADIATHAHQRREKARTNTFAEDRRGRR
SQFVAEHRREAGACGVAEAAGNQLGDDQRVRLQKQRSRDHFCGAWRRGTAPVLGGENFSSKANQERIRTR
LEFVRVPLLPLPPYGGLRLRLGLDYCPKTLQGQLLLGRVRLRALAEVPAYTLGEQGQPTWHCRPLLHAHQ
DVSHQHALLQRKRADHLRQDPLHGSGSLWLLV



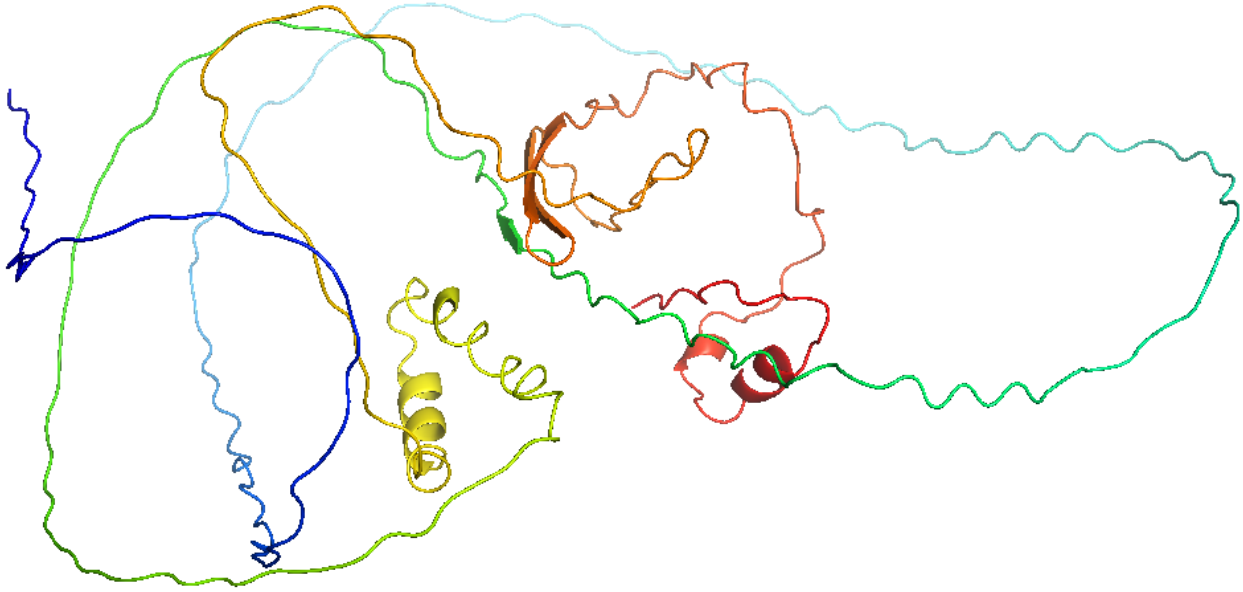
ssGE2-MUTATION#51: (mstn1-mutation-ssGE2 #51) +1 A

ATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTTCG**TGAG**TGGCGTTCGGTCCGATGGCGCGCACTGACACCGG
AGCACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCG
GCCAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAA
CTGCGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAG
CAGCTGCTCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGG
ACGACGAGGAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGAGCCCAA**CCC**CGACGTTCAAGTCGACC
AAAAACCGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGT
GCACTTGCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAG
AAGGCACGTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGA
AGCAGGTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAA
AGCAACGATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCT
GAAGTTCCAAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGC
TACCCCCTTACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACT
GCTCGGGCGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTG

GCACTGCCGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGAT
CATCTACGGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLGFVSGVRSDGAHHRSTGAAAAAAATYRRDGGARGAVFSGQRVRFPPAQQAAPSASHQVPD
SEQTAPQTSSQREPRCGQAAAAESATGAAAARPVRRARGRRQAGHSAPGRGGGRRGARHHRDRHEHGRRA
QPRRSSRPKTEVLFFLLQPEDPSEPHRKGAALGALAPGGGDNGVLADIATHAHQRREKARTNTFAEDRRG
RRSQFVAEHRREAGACGVAEAAGNQLGDDQRVRLQKQRSRDHFCGAWRRGTAPVLGGENFSSKANQERIR
TRLEFVRVPLLPLPPYGGLRLRLGLDYCPKTLQGQLLLGRVRLRALAEVPAYTLGEQGQPTWHCRPLLHA
HQDVSHQHALLQRKRADHLRQDPLHGSGSLWLLV



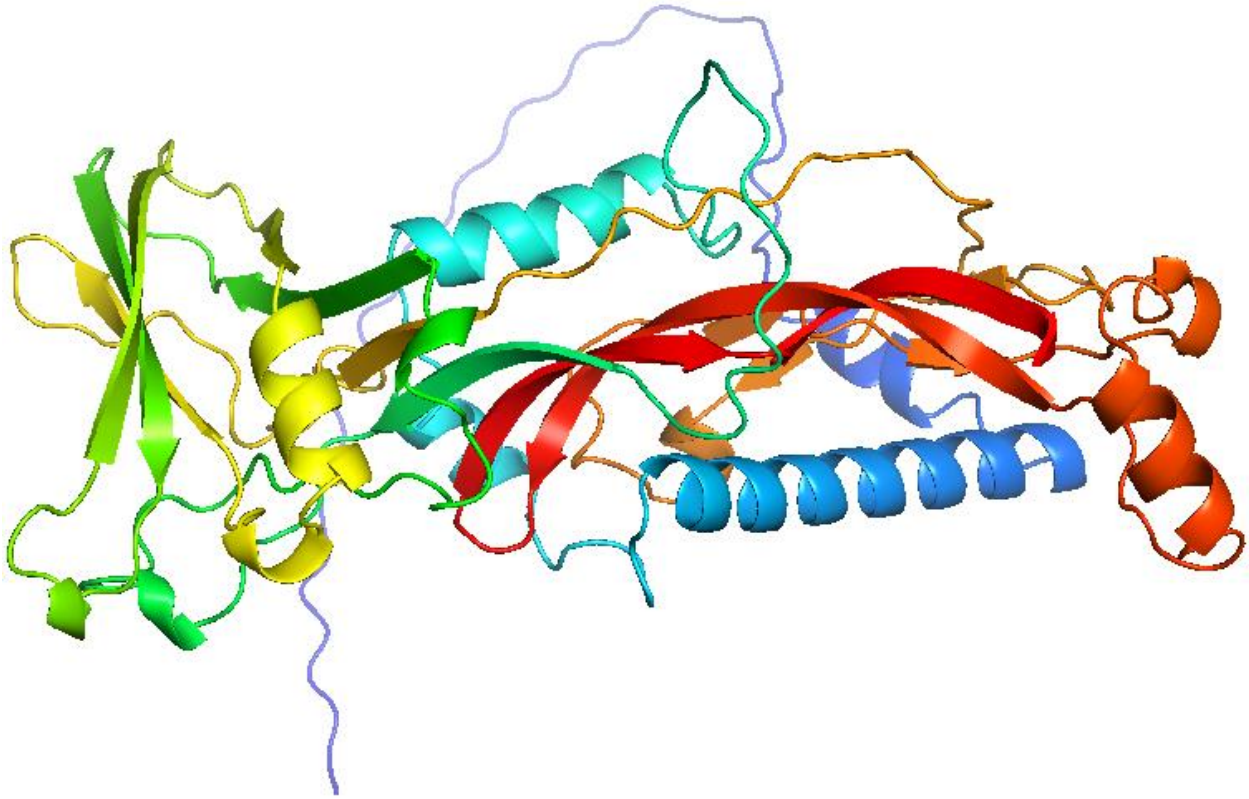
ssGE2-MUTATION#53, #54: (mstn1-mutation-ssGE2 #53, #54) -6 TCGTGG

ATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTTGGCGTTCGGTCCGATGGCGCGCACTGACACCGGAGCACCG
GAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGCCAGCG
CGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTGCGCCT
CAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAGCTGCT
CGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGACGACGAG
GAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGAGCCCAA**CCC**CGACGTTCAAGTCGACCAAAAACCG
AAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGTGCACTTGC
GCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAGAAGGCAC
GTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGAAGCAGGT
GCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAAAGCAACG
ATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCTGAAGTTCC
AAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGCTACCCCCT
TACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACTGCTCGGG
CGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTGGCACTGC

CGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGATCATCTAC
GGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLGLAFGPMARTDTGAPEQQQQQQQPTAVTEEREAQCSAASACAFRQHSKQLRLQAIKSQIL
SKLRLKQAPNVSRDVVKQLLPKAPPVQQLLDLYDVLGDDGKPGTALQDEEEDDEEHATTETVMSMAAEPN
PDVQVDQKPKCCFFSFSPKIQASRIVRAQLWVHLRPADEATTVFLQISRLMPIKDGRRHVRIRSLKIDVD
AGVSSWQSIDVKQVLAVWLRQPETNWGIEINAFDSKSNDLAITSAEPGEEGLLPFLEVKISEVPKRTRRE
SGLDCDENSSESRCCRYPLTVDFEDFGWDWIIAPKRYKANYCSGECDYVHLQKYPHTHLVNKANPRGTAG
PCCTPTKMSPINMLYFNGKEQIIYGKIPSMVVDRCGCS



#3 Mutations from msMGE (Sequence, AA and predicted protein structure)

1. *lh* mutation:

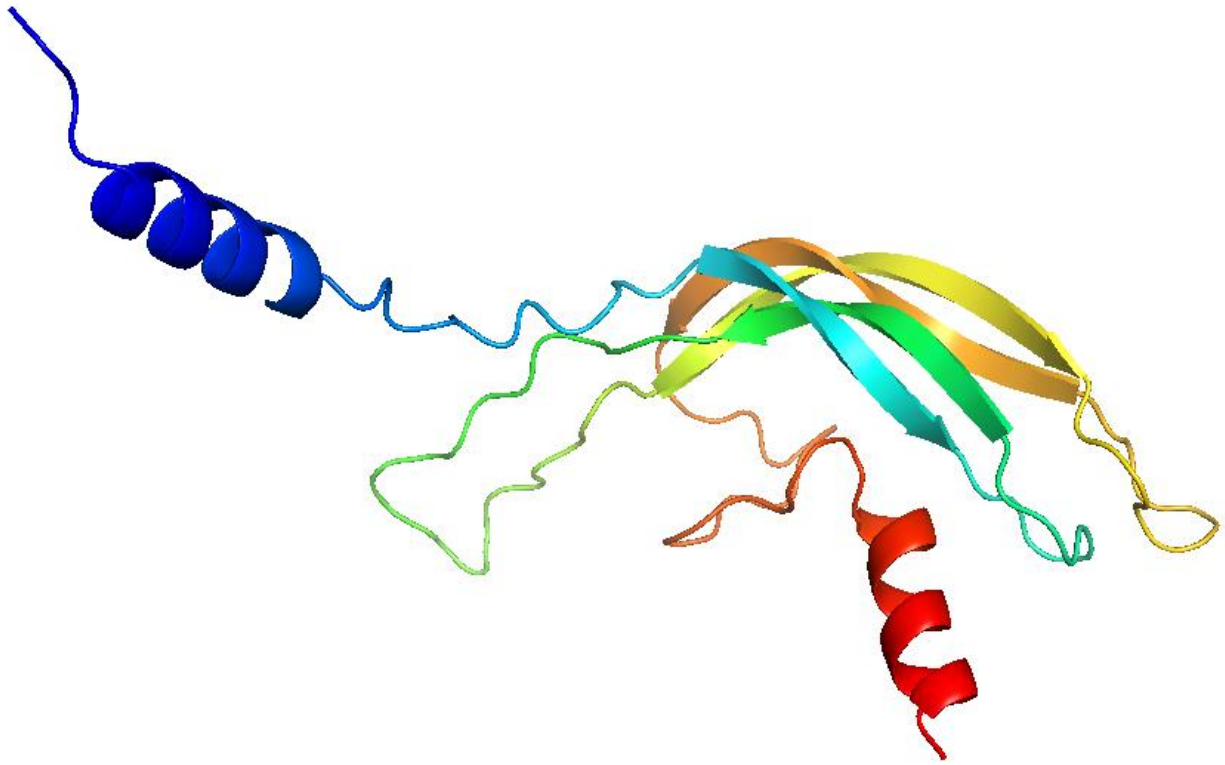
***lh* CDS sequence for AA:**

WT CDS:

ATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGC
CACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGC
CATCTGCAGC**GGG**CACTGCTTCACCAAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGT
ACCTACAGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATC
CTGTCGCACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAATCCGG
ATTTCTGTATGACACAGAAAGAGTACATCCTGGATTAC**TGA**

AA:

MSVPASSFLLLCFLMNSFSPAQSYILPHCEPVNETVSVEKDGCPKCLVFQTAICSGHCFTKEPVYKSPFSSIYQHVCTYRD
VRYETVRLPDCRPGVDPHVTYPVALSCECSLCTMDTSDCTIESLNPDFCMTQKEYILDY

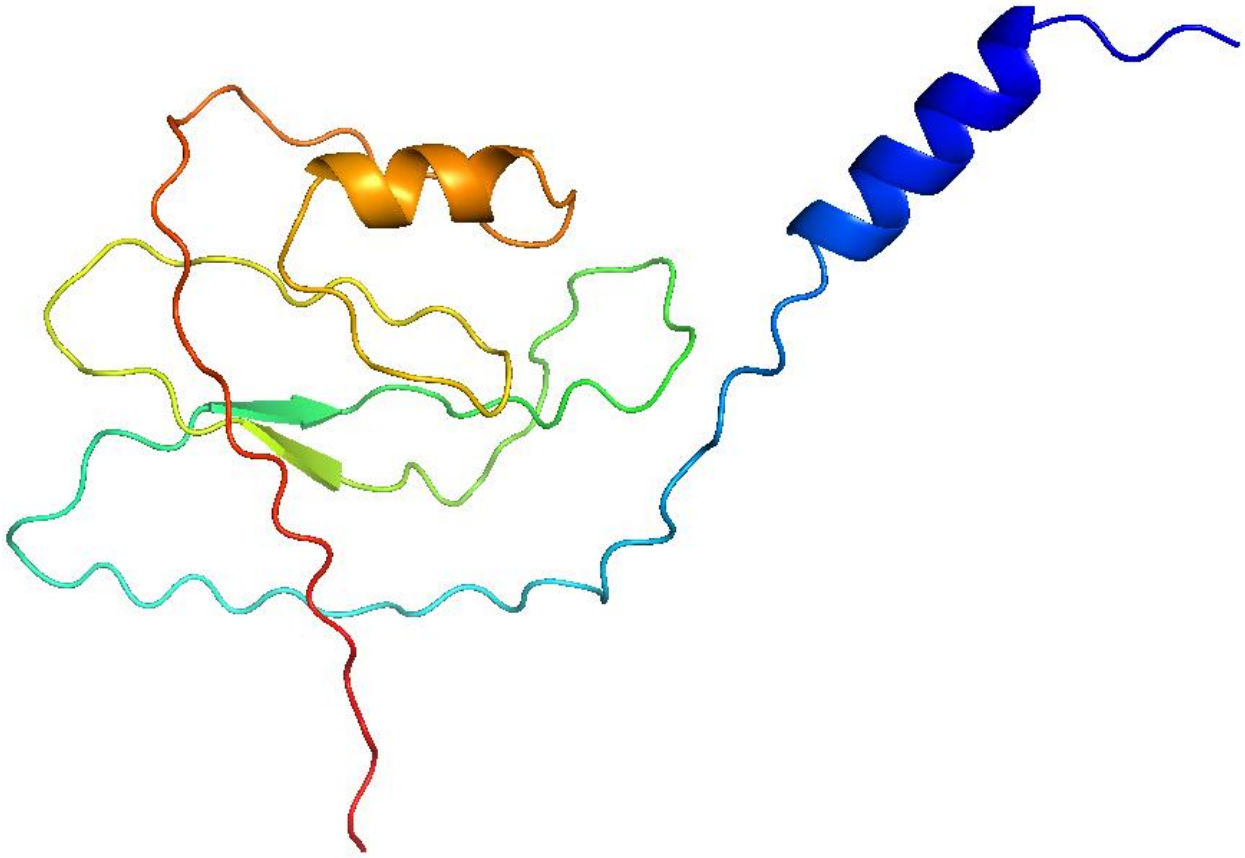


msMGE-MUTATION#1': (lh-mutation-msMGE, #1')

ATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGC
CACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGC
CATCTGCAACGCTACGACTGCTTCACCAAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTG
TACCTACAGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATAT
CCTGTCGCACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAATCCG
GATTTCTGTATGACACAGAAAGAGTACATCCTGGATTAC**TGA**

Predicted AA:

MSVPASSFLLLCFLMNSFSPAQSYILPHCEPVNETVSVEKDGCPKCLVFQTAICNATTASPRNLCTRARS
LPSISMCVPTGTFAMKPYACLTVGPVWILTSHILSHAASAACAPWTPRTVPSRAIRISVHRKSTSWIT

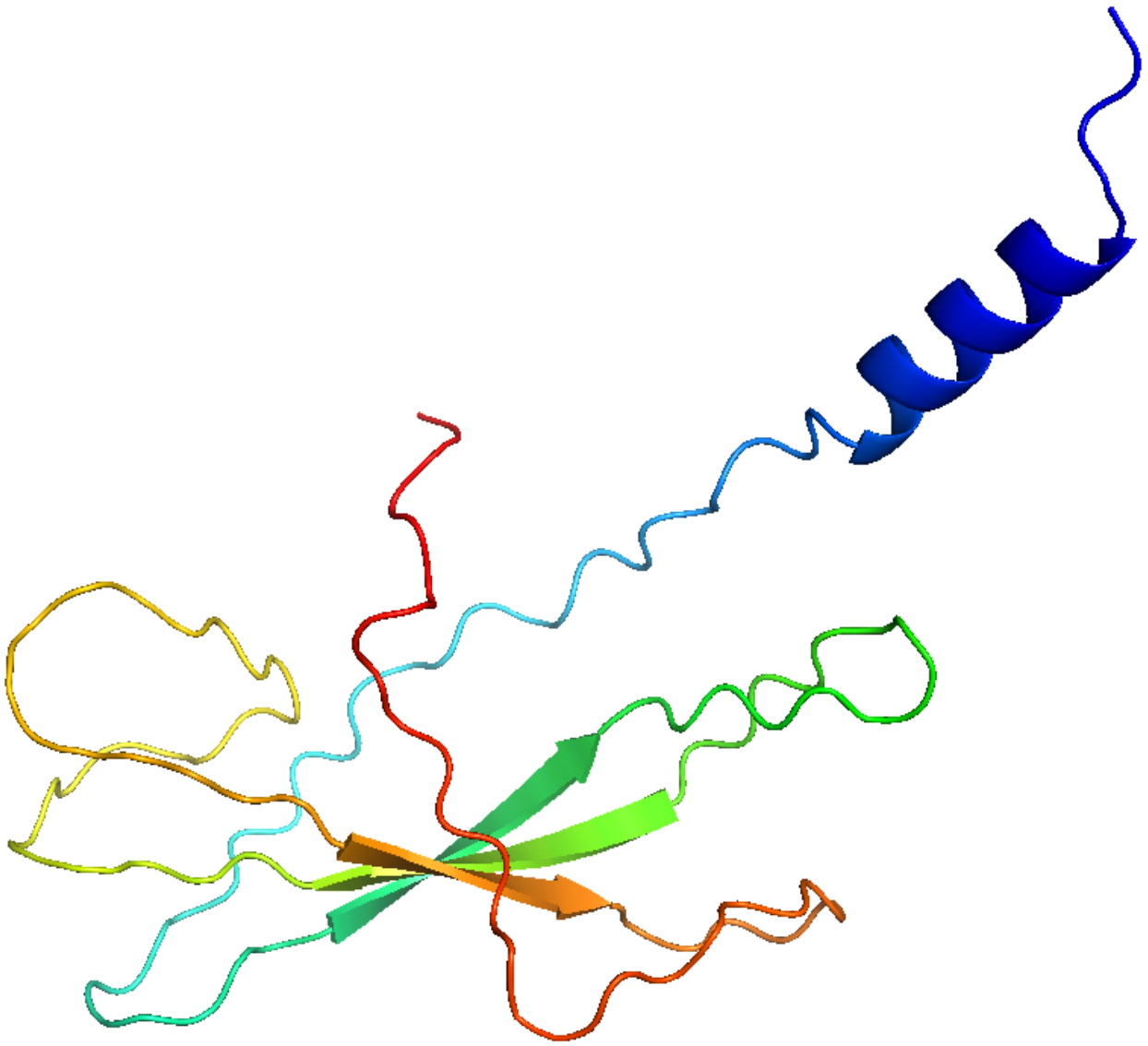


msMGE-MUTATION#2' and #5': (lh-mutation-msMGE, #2' and #5') +1 A, -3**GGG**

ATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGC
CACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGC
CATCTGCAGCACACTGCTTCACCAAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGTACC
TACAGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATCCTG
TCGCACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAATCCGGATTT
CTGTATGACACAGAAAGAGTACATCCTGGATTAC**TGA**

Predicted AA:

MSVPASSFLLLCFLMNSFSPAQSYILPHCEPVNETVSVEKDGCPKCLVFQTAICSTLLHQGTCVQEPVLF
HLSACVYLQGRSLNRTPALSARCGSSRHISCRTKLRVQPVHHGHLGLYHREPESGFLYDTERVHPGLL

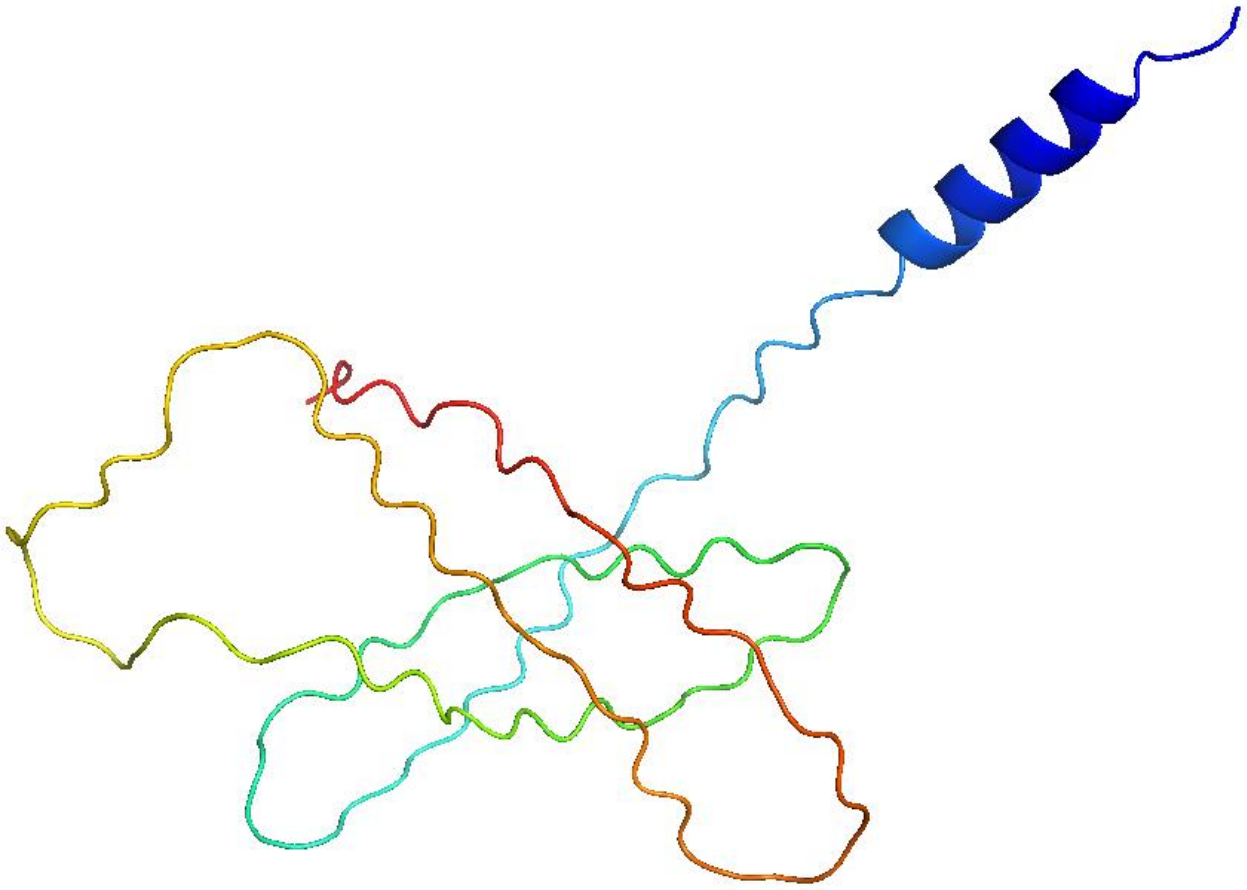


msMGE-MUTATION#8', #15, #19': (lh-mutation-msMGE, #8', #15', #19') +1 T, -5AGCGGG

ATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGC
CACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGC
CATCTGCTCACTGCTTCACCAAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGTACCTAC
AGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATCCTGTCG
CACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAATCCGGATTTCTG
TATGACACAGAAAGAGTACATCCTGGATTAC**TGA**

Predicted AA:

MSVPASSFLLLCFLMNSFSPAQSYILPHCEPVNETVSVEKDGCPKCLVFQTAICSLLHQGTCVQEPVLFH
LSACVYLQGRSLNRTPALSARCGSSRHISCRTKLRVQPVHHGHLGLYHREPESGFLYDTERVHPGLL

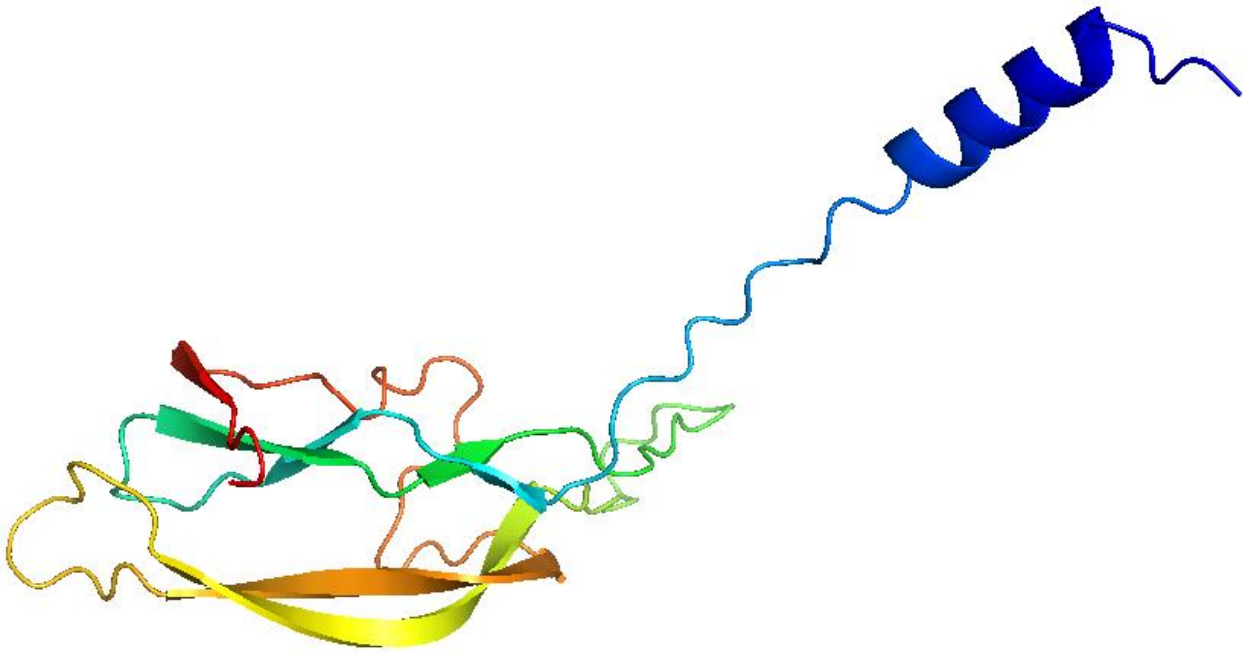


msMGE-MUTATION#14': (lh-mutation-msMGE, #14') -2/+2 CT(to GA), -3AGC

ATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGC
CACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGC
CATGAGC**GGG**CACTGCTTCACCAAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGTACC
TACAGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATCCTG
TCGCACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAATCCGGATTT
CTGTATGACACAGAAAGAGTACATCCTGGATTAC**TGA**

Predicted AA:

MSVPASSFLLLCFLMNSFSPAQSYILPHCEPVNETVSVEKDGCPKCLVFQTAMSGHCFTKEPVYKSPFSS
IYQHVCTYRDVRYETVRLPDCRPGVDPHVTYPVALSCECSLCTMDTSDCTIESLNPDFCMTQKEYILDY

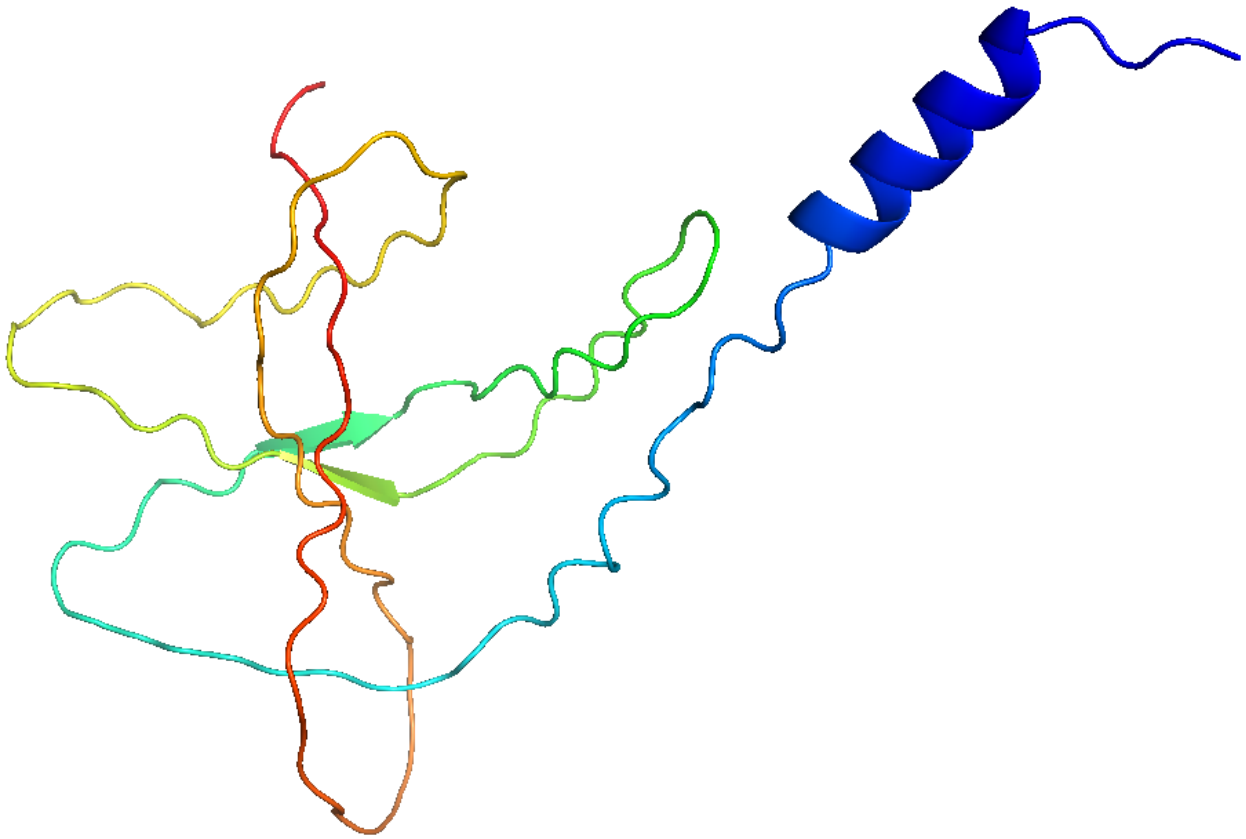


msMGE-MUTATION#18': **(lh-mutation-msMGE, #18')** +1 T, +2 AC, -2CA

ATGTCAGTGCCAGCTTCCTCTTTTCTTCTCCTGTGTTTCTTGATGAACTCCTTCTCCCCCGCTCAAAGCTACATTCTGC
CACACTGCGAACCTGTTAATGAGACTGTTTCTGTGGAGAAAGATGGCTGCCCGAAATGCCTTGTGTTTCAAACCGC
CATCTGCAACGC**TGGG**CTGCTTCACCAAGGAACCTGTGTACAAGAGCCCGTTCTCTTCCATCTATCAGCATGTGTGT
ACCTACAGGGACGTTCGCTATGAAACCGTACGCCTGCCTGACTGTCGGCCCGGTGTGGATCCTCACGTCACATATC
CTGTCGCACTAAGCTGCGAGTGCAGCCTGTGCACCATGGACACCTCGGACTGTACCATCGAGAGCCTGAATCCGG
ATTTCTGTATGACACAGAAAGAGTACATCCTGGATTAC**TGA**

Predicted AA:

MSVPASSFLLLCFLMNSFSPAQSYILPHCEPVNETVSVEKDGCPKCLVFQTAICNAGLLHQGTCVQEPVL
FHLSACVYLQGRSLNRTPALSARCGSSRHISCRTKLRVQPVHHGHLGLYHREPESGFLYDTERVHPGLL



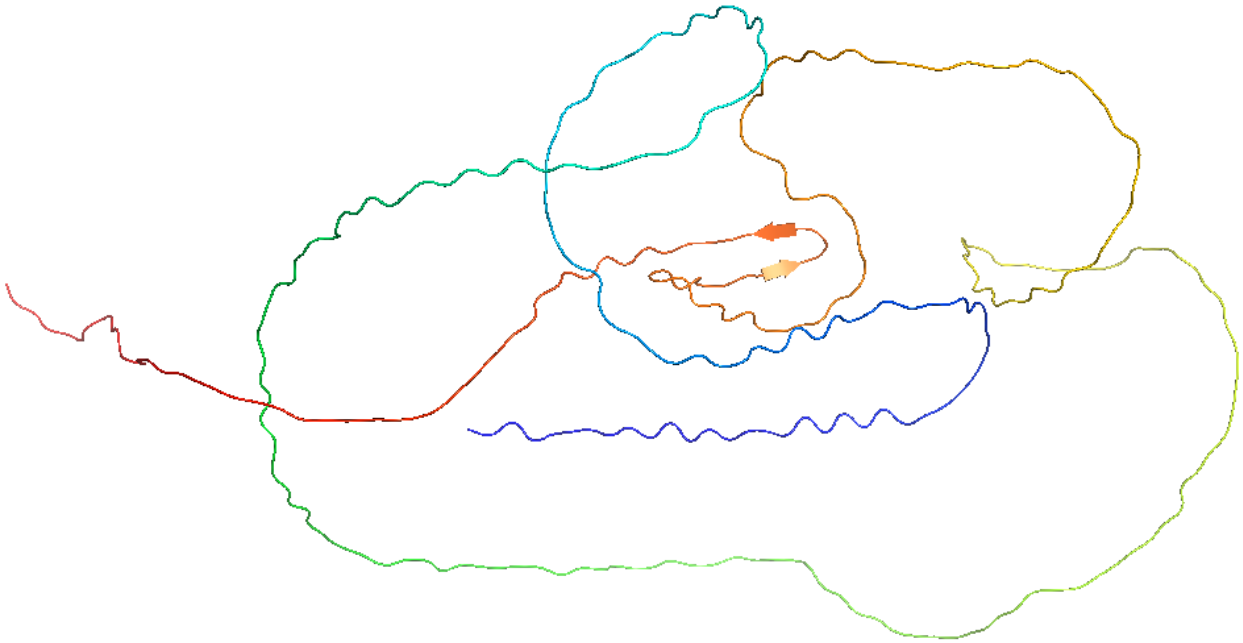
2. *mc4r* mutation:

msMGE-MUTATION#3': (mc4r-mutation-msMGE, #3')

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCTCCACCCG**A**TCTTCATCACG
CTAGGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGCCCA
TGTACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTGTGA
TGGCGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACTCCA
TGATCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTCTAC
GCCCTGCGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCACGG
CGTCCGGTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACCATGC
TGGCCCTCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTTACCG
GGGAACGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGTTTGT
CGTGTGCTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCTTCAT
GTCTCACTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGGAGTC
AGGAGATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTGCGTC
GGCTTCGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLISTRSSSRGWSAFWRTSWSRPSSRTRTS
TRPCTSSSAAWRWPTCWAYRTRQKRLWRSPAATPSLETSKAWTMCSTPSAAHSWPPFGVSWPSPWTATSP
SSTPCATTTSPNAARRSSSYAYGASARRPVCSSSSTRRALQSSSALSACSSPCWPSWPRFTCTCSSWRGF
TNASPPYRGTAPCGRRPTRAPRSPSCSECLSCAGRRFFSTSFSSLVRGTRIASASCLTSTCTFSCATRST
RSSTRSGVRRGRPSGRSAAAGLRDGAAAGVASASTRGLTAI

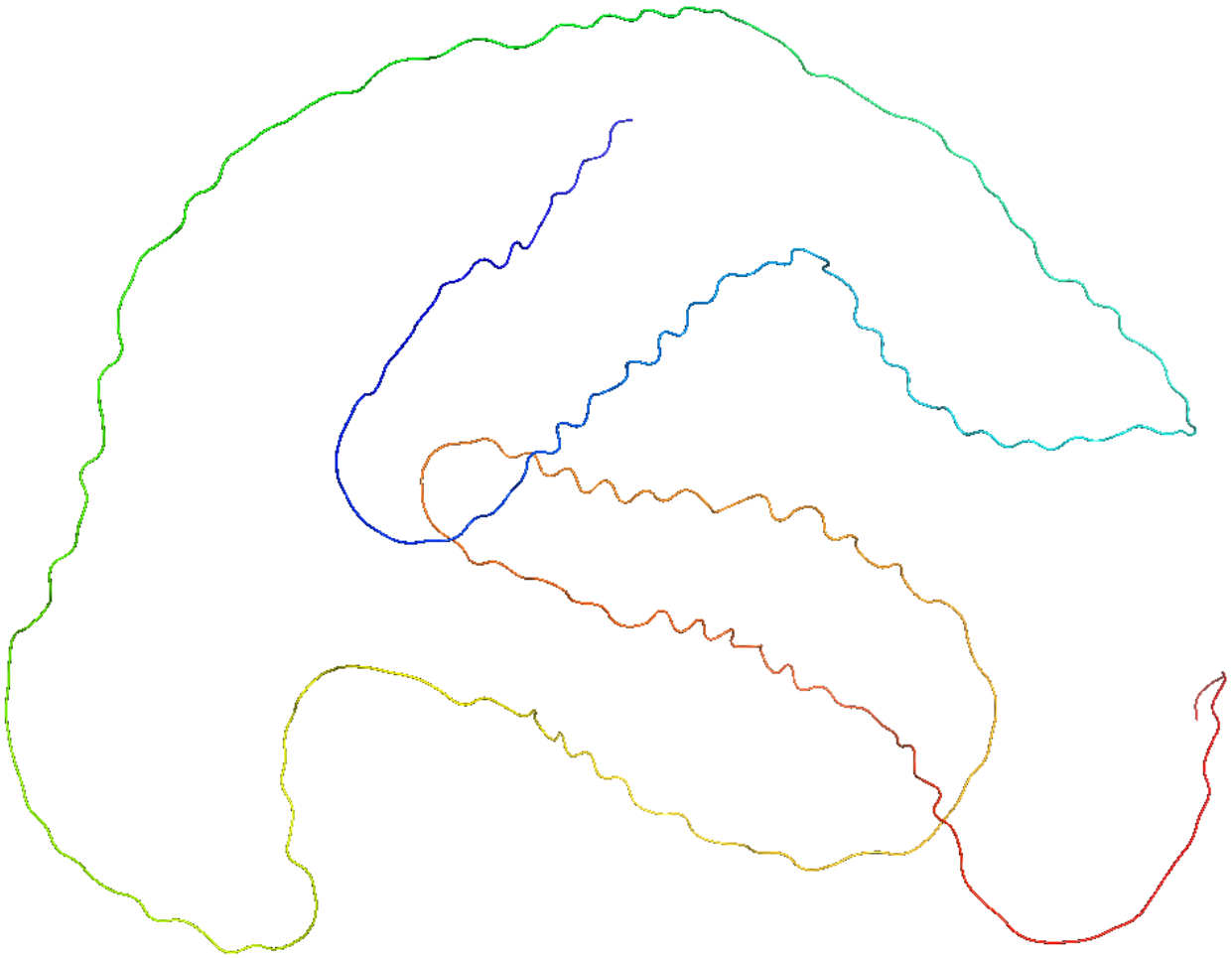


msMGE-MUTATION#5': **(mc4r-mutation-msMGE, #5')** +2 TA, +3 ACA

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCTCCTAACACACG**AGG**TCTT
CATCACGCTAGGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCAC
TCGCCCATGTACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACG
GCTGTGATGGCGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTC
GACTCCATGATCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCAT
CTTCTACGCCCTGCGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCT
GCACGGCGTCCGGTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTC
ACCATGCTGGCCCTCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGC
CTTACCGGGGAACGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAG
TGTTTGTCGTGTGCTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCT
GCTTCATGTCTCACTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTC
AGGAGTCAGGAGATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGA
GTTGCGTCGGCTTCGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLISHTRSSSRGWSAFWRTSWSRPSSRTRT
STRPCTSSSAAWRWPTCWAYRTRQKRLWRSPAATPSLETSKAWTMCSTPSAAHSWPPFGVSWPSPWTATS
PSSTPCATTTSPNAARRSSSYAYGASARRPVCSSSSTRRALQSSSALSACSSPCWPSWPRFTCTCSSWRG
FTNASPPYRGTAPCGRRPTRAPRSPSCSECLSCAGRRFFSTSFSSLVRGTRIASASCLTSTCTFSCATRS
TRSSTRSGVRRGRPSGRSAAAGLRDGAAAGVASASTRGLTAI

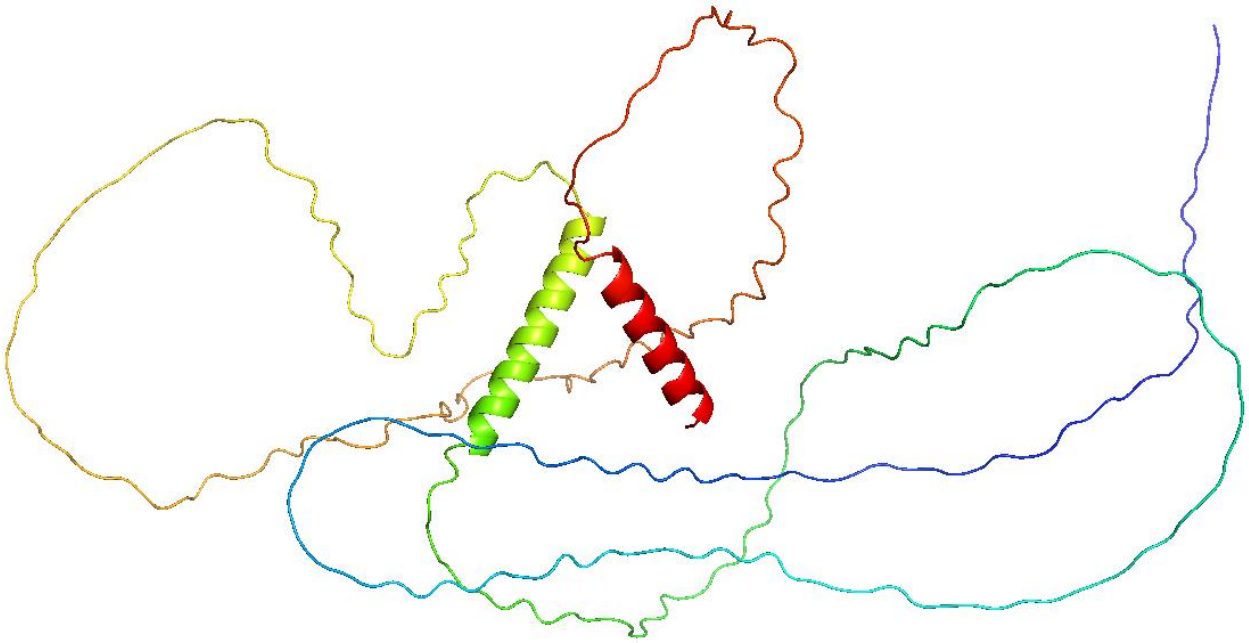


msMGE-MUTATION#8': **(mc4r-mutation-msMGE, #8')** -5 CACCG

ATGAACGTGTCGGAGCACCACGGGATGCAGCATGCACACCGGAACCACAGCCTGGGCGTGCAGATTGGAAACAA
AGCCGGCTCGGGGGAAAGGAACTCGGAGTCGGGCTGCTACGAGCAGCTGTTGATCTC**AGG**TCTTCATCACGCTA
GGGTTGGTCAGCCTTCTGGAGAACATCCTGGTAATCGCGGCCATCGTCAAGAACAAGAACTTCCACTCGCCCATGT
ACTTCTTCATCTGCAGCCTGGCGGTGGCCGACCTGCTGGTGAGCGTATCGAACGCGACAGAAACGGCTGTGATGG
CGCTGATCACCAGCGGCAACCTGACCATCTCTGGAGACGTCGTGAAAAGCATGGACAATGTGTTCGACTCCATGA
TCTGCAGCTCACTCCTGGCCTCCATTTGGAGTCTCCTGGCCATCGCCGTGGACCGCTACGTCACCATCTTCTACGCC
CTGCGCTACCACAACATCATGACCCAACGCCGCGCGGCGCTCATCATCGTATGCATATGGAGCTTCTGCACGGCGT
CCGGTGTGCTCTTCATCATCTACTCGGAGAGCGCTACAGTCCTCATCTGCCTTATCAGCATGTTCTTCACCATGCTG
GCCCTCATGGCCTCGCTTTACGTGCACATGTTCCTCTTGGCGCGGCTTCACATGAAACGCATCGCCGCCTTACCGGG
GAACGGCCCCGTGTGGCAGGCGGCCAACATGAAGGGCGCCGTGACGCTCACCATCCTGCTCGGAGTGTTTGTCGT
GTGCTGGGCGCCGTTTTTTCTCCACCTCATTCTCATGATCTCTTGTCCGAGGAACCCGTATTGCGTCTGCTTCATGTC
TCACTTCAACATGTACCTGATTCTGATCATGTGCAACTCGGTGATCGACCCGCTCATCTACGCGTTCAGGAGTCAGG
AGATGAGGAAGACCTTCCGGGAGATCTGCTGCGGCTGGGCTTCGGGATGGAGCTGCGGCTGGAGTTGCGTCGGC
TTCGACGAGAGGCTTAACAGCTAT**TGA**

Predicted AA:

MNVSEHHGMQHAHRNHSLGVQIGNKAGSGERNSESGCYEQLLISGLHHARVGQPSGEHPGNRGHRQEQEL
PLAHVLLHLQPGGGRPAGERIERDRNGCDGADHQRQPDHLWRRREKHGQCVRLHDLQLTPGLHLESPGHR
RGPLRHHLLRPALPQHHDPTPRGAHHRMHMELLHGVRCALHHLLGERYSPHLPYQHVLHHAGPHGLALRA
HVPLGAASHETHRRLTGERPRVAGGQHEGRRDAHHPARSVCRVLGAVFSPPHSHDLLSEEPVLRLLHVSL
QHVPDSDHVQLGDRPAHLRVQESGDEEDLPGDLLRLGFGMELRLELRRLRREAQLL



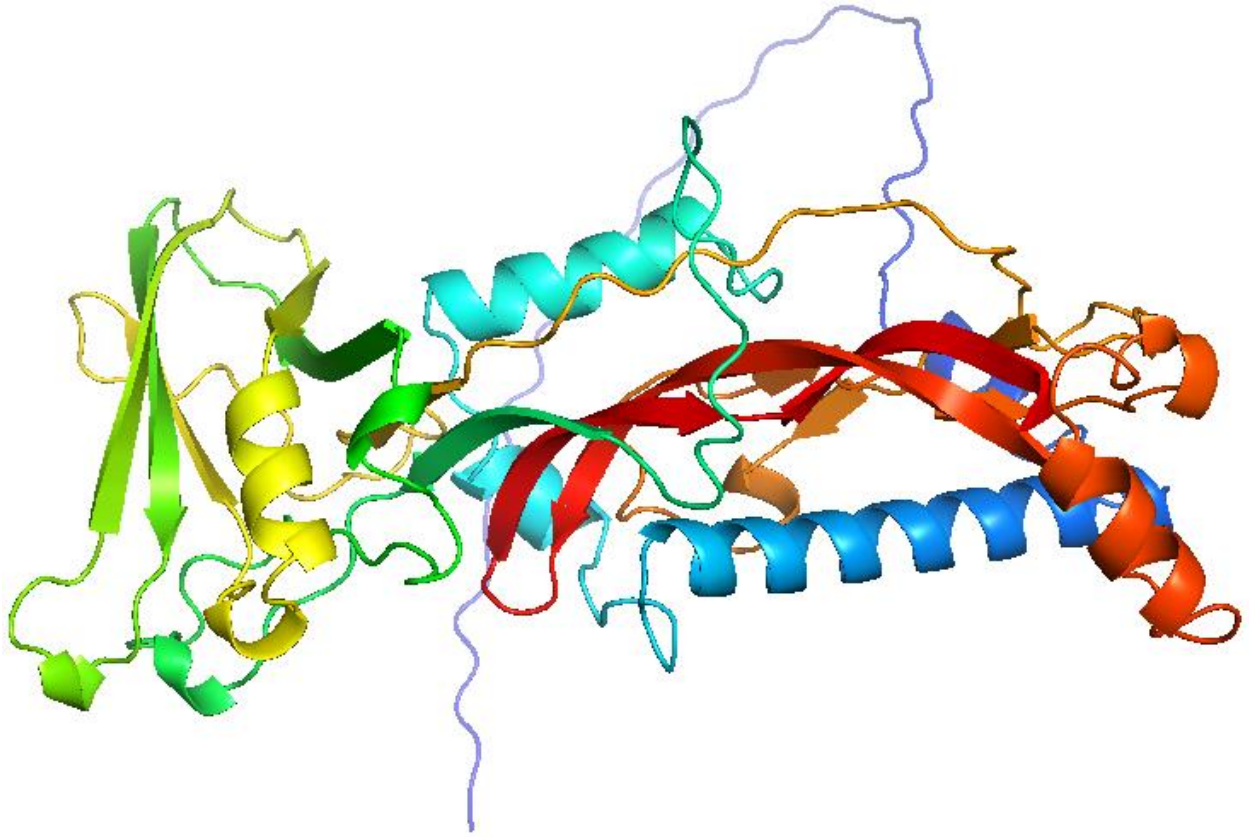
3. *mstn* mutation:

msMGE-MUTATION#4': (mstn1-mutation-msMGE, #4')

ATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTTGGCGTTCGGTCCGATGGCGCGCACTGACACCGGAGCACCG
GAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGCCAGCG
CGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTGCGCCT
CAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAGCTGCT
CGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGACGACGAG
GAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGAGCCCAA**CCC**CGACGTTCAAGTCGACCAAAAACCG
AAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGTGCACTTGC
GCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAGAAGGCAC
GTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGAAGCAGGT
GCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAAAGCAACG
ATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCTGAAGTTCC
AAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGCTACCCCCT
TACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACTGCTCGGG
CGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTGGCACTGC
CGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGATCATCTAC
GGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLGLAFGPMARTDTGAPEQQQQQQQPTAVTEEREAQCSAASACAFRQHSKQLRLQAIKSQIL
SKLRLKQAPNVSRDVVKQLLPKAPPVQQLLDLYDVLGDDGKPGTALQDEEEDDEEHATTETVMSMAAEPN
PDVQVDQKPKCCFFSFSPKIQASRIVRAQLWVHLRPADEATTVFLQISRLMPIKDGRRHVRIRSLKIDVD
AGVSSWQSIDVKQVLAVWLRQPETNWGIEINAFDSKSNDLAITSAEPGEEGLLPFLEVKISEVPKRTRRE
SGLDCDENSSESRCCRYPLTVDFEDFGWDWIIAPKRYKANYCSGECDYVHLQKYPHTHLVNKANPRGTAG
PCCTPTKMSPINMLYFNGKEQIIYGKIPSMVVDRCGCS



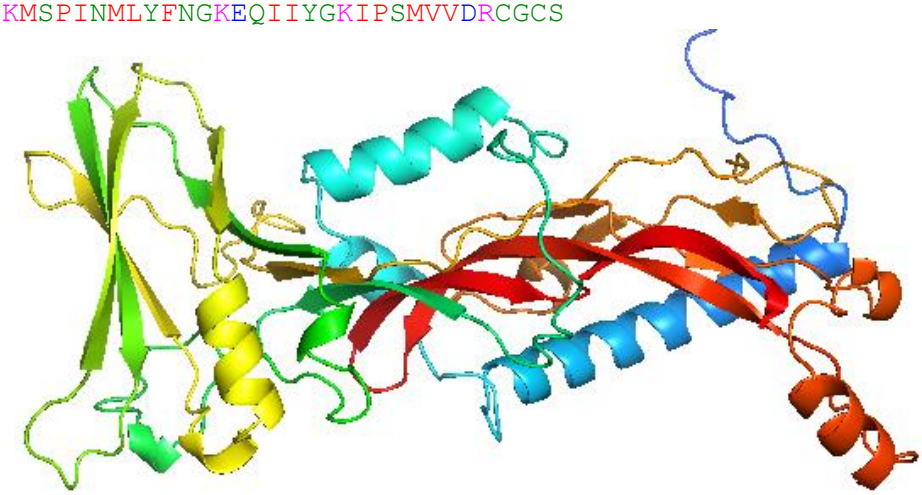
msMGE-MUTATION13#: (mstn1-mutation-msMGE, #13') -1/+1 G(to T)

ATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTTCT**TGG**TGGCGTTCGGTCCGATGGCGCGCACTGACACCGGA
GCACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGG
CCAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACT
GCGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCA
GCTGCTCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGAC
GACGAGGAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGAGCCCAA**CCC**CGACGTTCAAGTCGACCA
AAAACCGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGT
GCACTTGCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAG
AAGGCACGTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGA
AGCAGGTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAA

AGCAACGATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCT
GAAGTTCCAAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGC
TACCCCCTTACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACT
GCTCGGGCGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTG
GCACTGCCGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGAT
CATCTACGGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLGFLVAFGPMARTDTGAPEQQQQQQQPTAVTEEREAQCSAASACAFRQHSKQLRLQAIKSQ
ILSKLRLKQAPNVSRDVVKQLLPKAPPVQQLLDLYDVLGDDGKPGTALQDEEEDDEEHATTETVMSMAAE
PNPDVQVDQKPKCCFFSFSPKIQASRIVRAQLWVHLRPADEATTVFLQISRLMPIKDGRRHVRIRSLKID
VDAGVSSWQSIDVKQVLAVWLRQPETNWGIEINAFDSKSNDLAITSAEPGEEGLLPFLEVKISEVPKRTR
RESGLDCDENSSESRCCRYPLTVDFEDFGWDWIIAPKRYKANYCSGECDYVHLQKYPHTHLVNKANPRGT
AGPCCTPTKMSPINMLYFNGKEQIIYGKIPSMVVDRCGCS

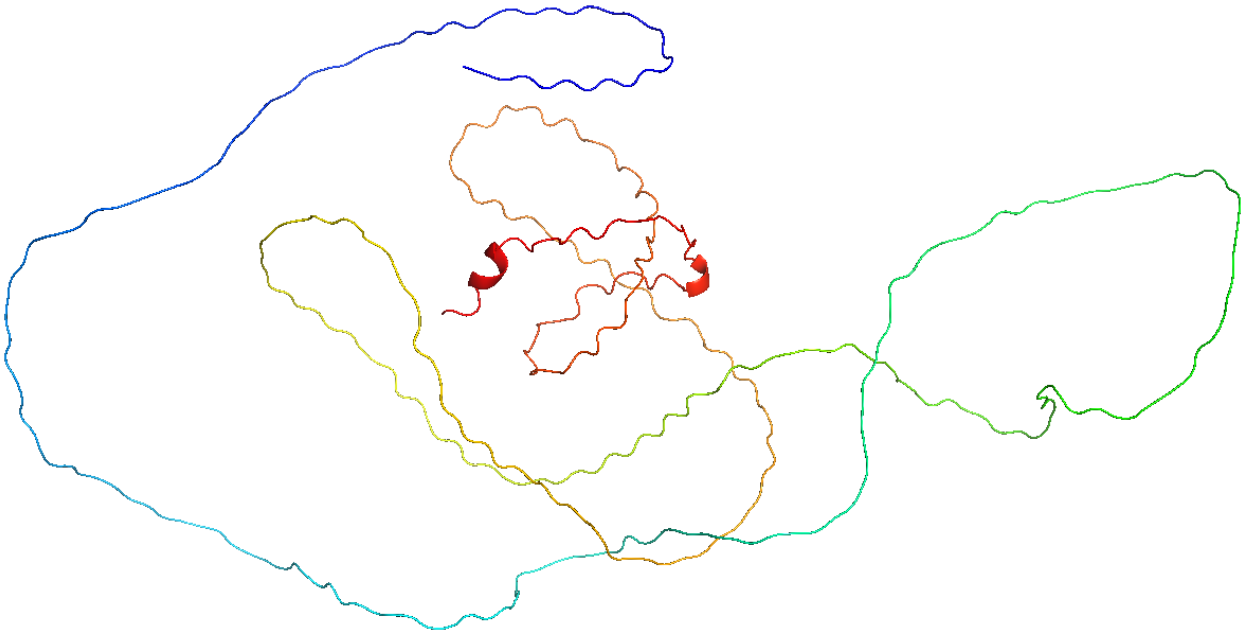


msMGE-MUTATION#7': (mstn1-mutation-msMGE, #7') +1 G, +3 TAC, -5GGGCT

ATGCATTTAGCGCAGGTTCTGATTTCGCTTCG**TGG**TGGTACGCGTTCGGTCCGATGGCGCGCACTGACACCGGAG
CACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGC
CAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTG
CGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAG
CTGCTCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGACG
ACGAGGAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGAGCCCAA**CCC**CGACGTTCAAGTCGACCAAA
AACCGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGTGCA
CTTGCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAGAA
GGCACGTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGAAG
CAGGTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAAAG
CAACGATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCTGA
AGTTCCAAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGCTA
CCCCCTTACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACTGC
TCGGGCGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTGGC
ACTGCCGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGATCA
TCTACGGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLRGGTRSVRWRALTPEHRSSSSSSSNLPPRRSARRSVQRPARALSASTASSSVCKPSSPRF
ANCASNKLPTAAMWSSSCCRKRHRCSSCSTCTTCSGTTASRAQRSRTRRRTTRSTPPPRPSAWPPSPTPT
FKSTKNRSAVFSPSARRSKRAASGRSSGCTCARRMRRQRCSCRYRDSCPSKTGEGTYEYVRRSTWTQESV
RGRASTSRCLRCGGSRKPTGGLRSTRSTPKATISRSLLRSLEKRDCSRSWRKFLKFQSEPGENQDTVMRI
RPSPAAAATPLRWTLKTSAGTGLLPQNATRPTTARASATTCTCRSTRIHTWTRPTHVALPAPAARPPRCL
PSTCSTSTEKSRSSTARSPPWWIAVAAR

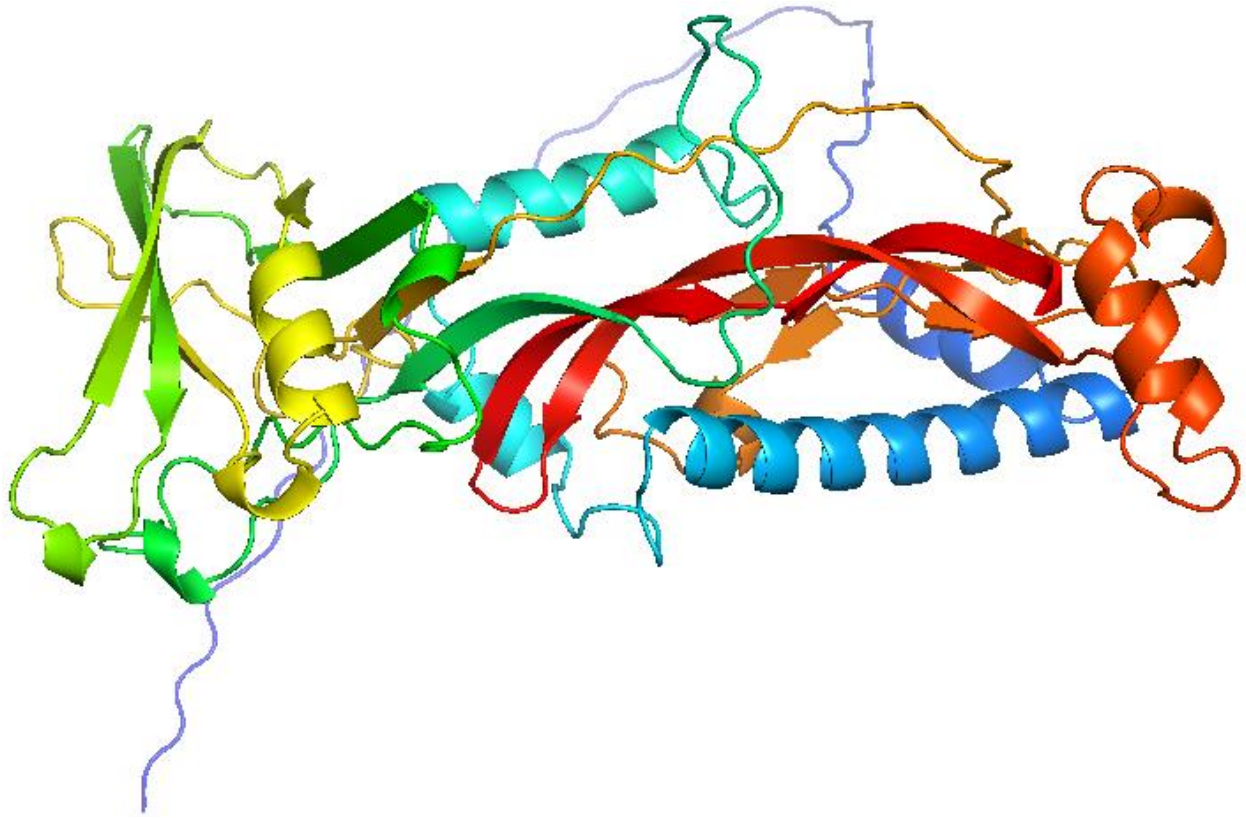


msMGE-MUTATION#9', #11' and #18': (mstn1-mutation-msMGE, #9', #11' and #18') +3 TAC

ATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTTCG**TGG**TGTACGCGTTCGGTCCGATGGCGCGCACTGACACCG
GAGCACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGC
GGCCAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAA
ACTGCGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCA
GCAGCTGCTCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAG
GACGACGAGGAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGAGCCCAA**CCC**CGACGTTCAAGTCGA
CCAAAAACCGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGG
GTGCACTTGCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGG
AGAAGGCACGTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGT
GAAGCAGGTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCA
AAAGCAACGATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTT
CTGAAGTTCCAAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCC
GCTACCCCCTTACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTA
CTGCTCGGGCGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACG
TGGCACTGCCGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAG
ATCATCTACGGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLGFVVYAFGPMARTDTGAPEQQQQQQQPTAVTEEREAQCSAASACAFRQHSKQLRLQAIKS
QILSKLRLKQAPNVSRDVVKQLLPKAPPVQQLLDLYDVLGDDGKPGTALQDEEEDDEEHATTETVMSMAA
EPNPDVQVDQKPKCCFFSFSPKIQASRIVRAQLWVHLRPADEATTVFLQISRLMPIKDGRRHVRIRSLKI
DVDAGVSSWQSIDVKQVLAVWLRQPETNWGIEINAFDSKSNDLAITSAEPGEEGLLPFLEVKISEVPKRT
RRESGLDCDENSSESRCCRYPLTVDFEDFGWDWIIAPKRYKANYCSGECDYVHLQKYPHTHLVNKANPRG
TAGPCCTPTKMSPINMLYFNGKEQIIYGKIPSMVVDRCGCS



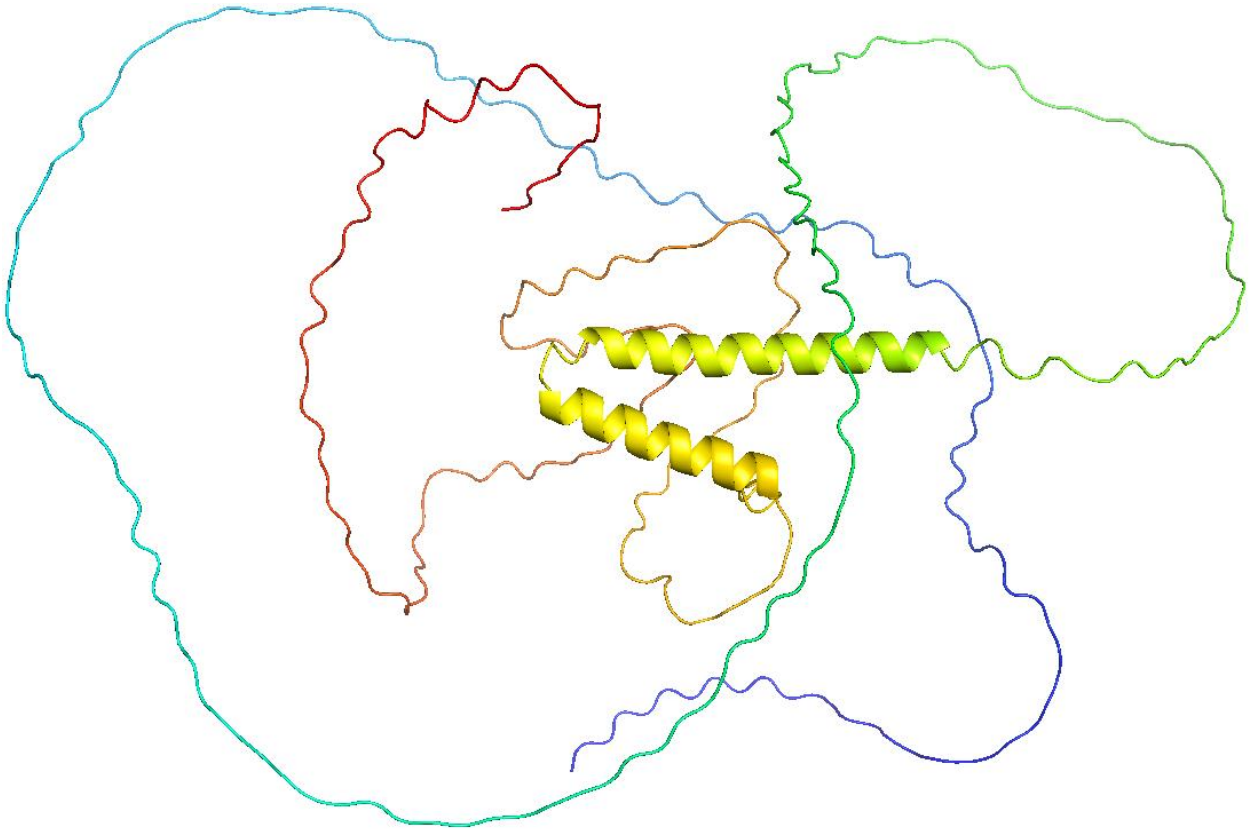
msMGE-MUTATION#10', #12', #15' and #17': (mstn1-mutation-msMGE, #10', #12' and #15', #17')
+1A, -3TCG

ATGCATTTAGCGCAGGTTCTGATTTCGCTGAGGCT**TGG**TGGCGTTCGGTCCGATGGCGCGCACTGACACCGGAGC
ACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGCC
AGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTGC
GCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAGC
TGCTCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGACGA
CGAGGAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGAGCCCAA**CCC**CGACGTTCAAGTCGACCAAAA
ACCGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGTGCAC
TTGCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAGAAG
GCACGTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGAAGC
AGGTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAAAGC
AACGATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCTGAA
GTTCCAAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGCTAC
CCCCTTACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACTGCT
CGGGCGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTGGCA

CTGCCGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGATCAT
CTACGGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLRLGGVRSDGAHHRSTGAAAAAAATYRRDGGARGAVFSGQRVRFPPAQQAAPSASHQVPDS
EQTAPQTSSQREPRCGQAAAAESATGAAAARPVRRARGRRQAGHSAPGRGGGRRGARHHRDRHEHGRRAQ
PRRSSRPKTEVLFFLLQPEDPSEPHRKGAALGALAPGGGDNGVLADIATHAHQRREKARTNTFAEDRRGR
RSQFVAEHRREAGACGVAEAAGNQLGDDQRVRLQKQRSRDHFCGAWRRGTAPVLGGENFSSKANQERIRT
RLEFVRVPLLPLPPYGGLRLRLGLDYCPKTLQGQLLLGRVRLRALAEVPAYTLGEQGQPTWHCRPLLHAH
QDVSHQHALLQRKRADHLRQDPLHGSGSLWLLV



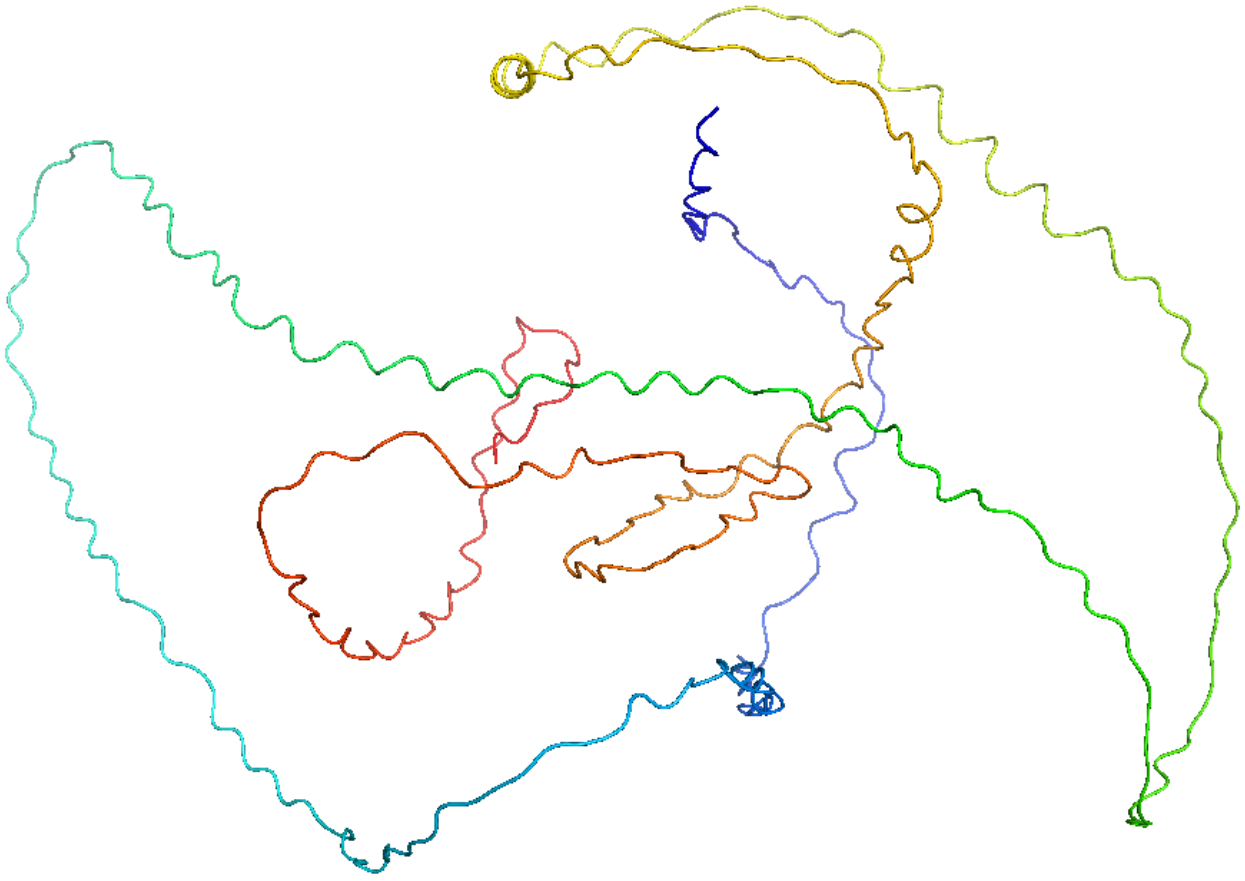
msMGE-MUTATION#16': (mstn1-mutation-msMGE, #16') +1C, -3/+3 GGG (to ATG)

ATGCATTTAGCGCAGGTTCTGATTTCGCTATGCTTCG**TGC**G**T**GGCGTTCGGTCCGATGGCGCGCACTGACACCGGA
GCACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGG
CCAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACT
GCGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCA
GCTGCTCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGAC
GACGAGGAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGAGCCCAA**CCC**CGACGTTCAAGTCGACCA
AAAACCGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGT
GCACTTGCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAG
AAGGCACGTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGA
AGCAGGTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAA

AGCAACGATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCT
GAAGTTCCAAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGC
TACCCCCTTACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACT
GCTCGGGCGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTG
GCACTGCCGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGAT
CATCTACGGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLCFVRGVRSDGAHHRSTGAAAAAAATYRRDGGARGAVFSGQRVRFPPAQQAAPSASHQVPD
SEQTAPQTSSQREPRCGQAAAAESATGAAAARPVRRARGRRQAGHSAPGRGGGRRGARHHRDRHEHGRRA
QPRRSSRPKTEVLFFLLQPEDPSEPHRKGAALGALAPGGGDNGVLADIATHAHQRREKARTNTFAEDRRG
RRSQFVAEHRREAGACGVAEAAGNQLGDDQRVRLQKQRSRDHFCGAWRRGTAPVLGGENFSSKANQERIR
TRLEFVRVPLLPLPPYGGLRLRLGLDYCPKTLQGQLLLGRVRLRALAEVPAYTLGEQGQPTWHCRPLLHA
HQDVSHQHALLQRKRADHLRQDPLHGSGSLWLLV



msMGE-MUTATION#5': (Both mstn1- and mstn2-mutation-msMGE, #5')

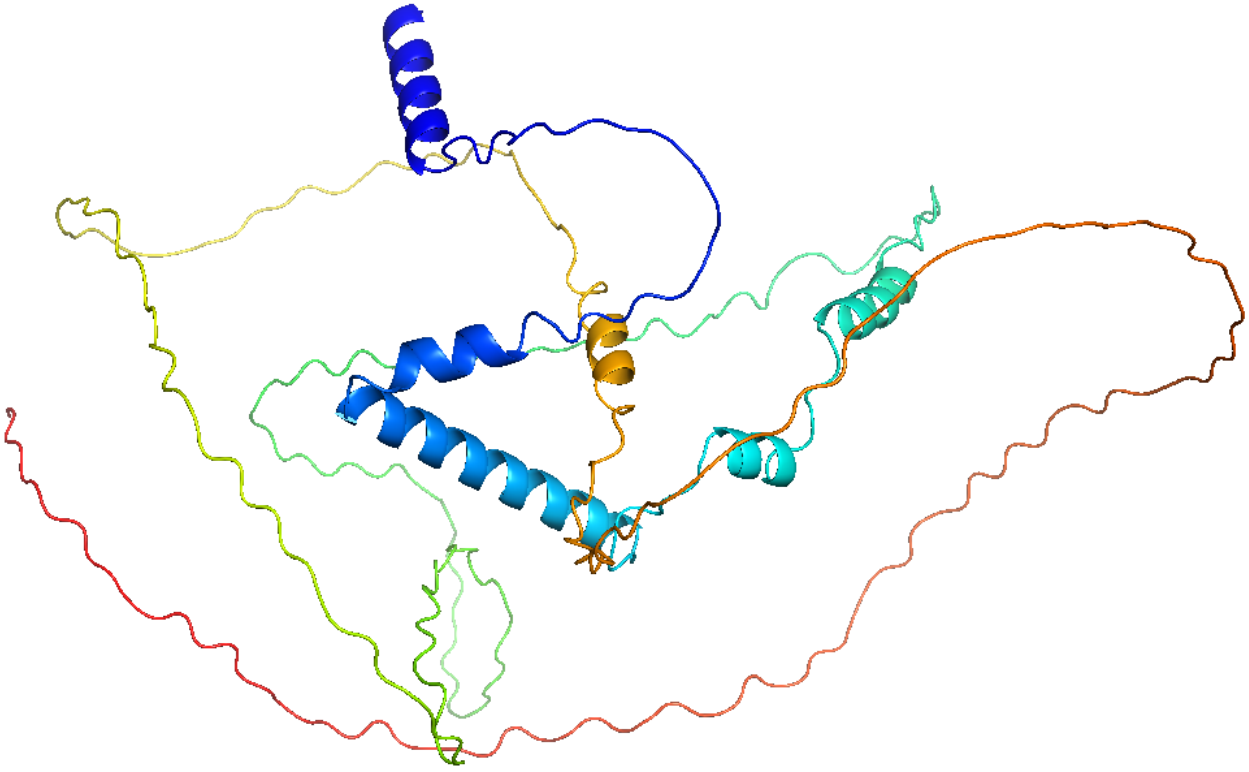
-1/+1 G(to T) for mstn1; +1T, -8 GCCCAA**CC**for mstn2

ATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTTCT**TGG**TGGCGTTCGGTCCGATGGCGCGCACTGACACCGGA
GCACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGG
CCAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACT

GCGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCA
GCTGCTCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGAC
GACGAGGAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGACCGTACGTTCAAGTCGACCAAAAACCG
AAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGTGCACTTGC
GCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAGAAGGCAC
GTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGAAGCAGGT
GCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAAAGCAACG
ATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCTGAAGTTCC
AAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGCTACCCCCT
TACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACTGCTCGGG
CGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTGGCACTGC
CGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGATCATCTAC
GGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLGFLVAFGPMARTDTGAPEQQQQQQQPTAVTEEREAQCSAASACAFRQHSKQLRLQAIKSQ
ILSKLRLKQAPNVSRDVVKQLLPKAPPVQQLLDLYDVLGDDGKPGTALQDEEEDDEEHATTETVMSMAAD
RTFKSTKNRSAVFSPSARRSKRAASGRSSGCTCARRMRRQRCSCRYRDSCPSKTGEGTYEYVRRSTWTQE
SVRGRASTSRCLRCGGSRKPTGGLRSTRSTPKATISRSLLRSLEKRDCSRSWRKFLKFQSEPGENQDTVM
RIRPSPAAAATPLRWTLKTSAGTGLLPQNATRPTTARASATTCTCRSTRIHTWTRPTHVALPAPAARPPR
CLPSTCSTSTEKSRSSTARSPPWWIAVAAR



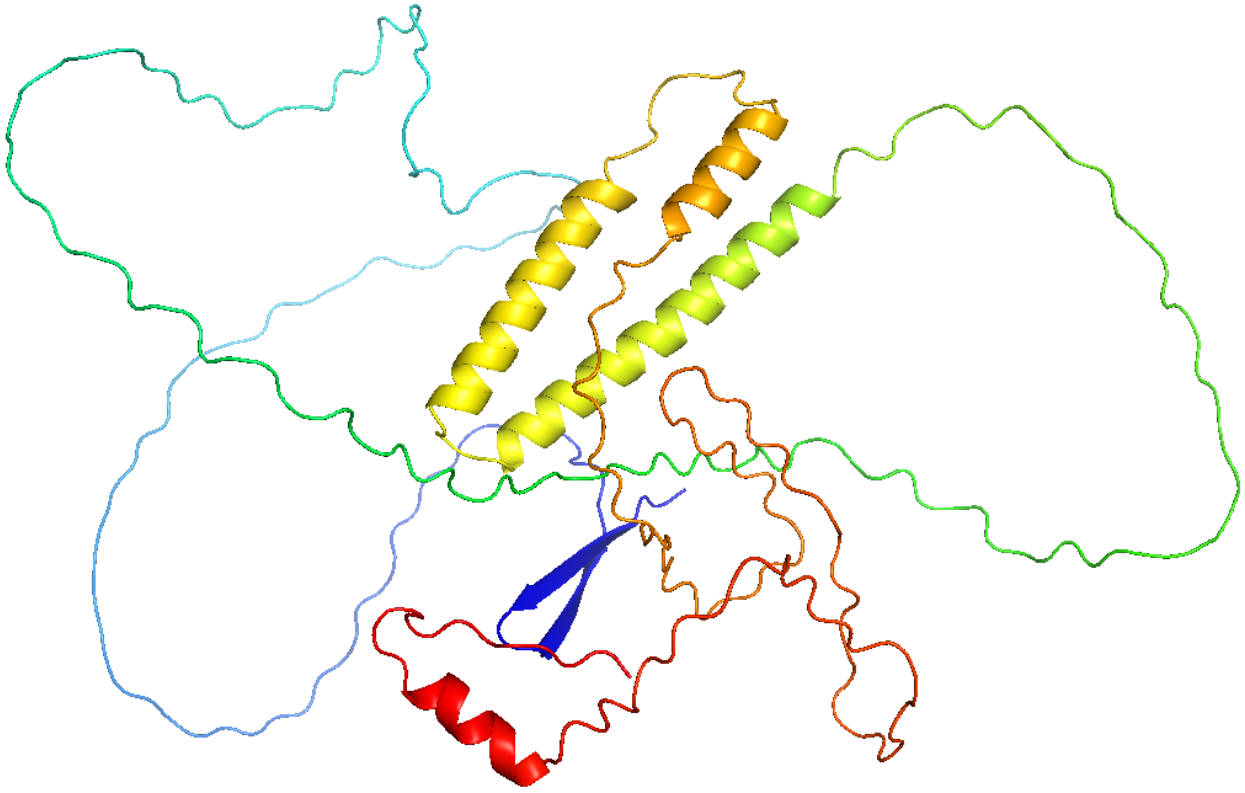
msMGE-MUTATION#8': (Both mstn1- and mstn2-mutation-msMGE, #8')

+1 G, +3 TAC, -5GGGCT for mstn1; +1T, -8 GCCCAA**CC**for mstn2

ATGCATTTAGCGCAGGTTCTGATTTCGCTTCG**TGG**TGGTACGCGTTCGGTCCGATGGCGCGCACTGACACCGGAG
CACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGGC
CAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACTG
CGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCAG
CTGCTCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGACG
ACGAGGAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGA**C**CGTACGTTCAAGTCGACCAAAAACCGA
AGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGTGCACTTGCG
CCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAGAAGGCACG
TACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGAAGCAGGTG
CTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAAAGCAACGA
TCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCTGAAGTTCCA
AAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGCTACCCCCTT
ACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACTGCTCGGGC
GAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTGGCACTGCC
GGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGATCATCTACG
GCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLRGGTRSVRWRALTPEHRSSSSSSSNLPPRRSARRSVQRPARALSASTASSSVCKPSSPRF
ANCASNKLPTAAMWSSSCCRKRHRCSSCSTCTTCSGTTASRAQRSRTRRRTTRSTPPPRPSAWPPTVRSS
RPKTEVLFFLLQPEDPSEPHRKGAALGALAPGGGDNGVLADIATHAHQRREKARTNTFAEDRRGRRSQFV
AEHRREAGACGVAEAAGNQLGDDQRVRLQKQRSRDHFCGAWRRGTAPVLGGENFSSKANQERIRTRLEFV
RVPLLPLPPYGGLRLRLGLDYCPKTLQGQLLLGRVRLRALAEVPAYTLGEQGQPTWHCRPLLHAHQDVSH
QHALLQRKRADHLRQDPLHGSGSLWLLV



msMGE-MUTATION#6': (mstn2-mutation-msMGE, #6')

ATGCATTTAGCGCAGGTTCTGATTTCGCTGGGCTTCG**TGG**TGGCGTTCGGTCCGATGGCGCGCACTGACACCGGA
GCACCGGAGCAGCAGCAGCAGCAGCAGCAACCTACCGCCGTGACGGAGGAGCGCGAGGCGCAGTGTTCAGCGG
CCAGCGCGTGCGCTTTCCGCCAGCACAGCAAGCAGCTCCGTCTGCAAGCCATCAAGTCCCAGATTCTGAGCAAACT
GCGCCTCAAACAAGCTCCCAACGTGAGCCGCGATGTGGTCAAGCAGCTGCTGCCGAAAGCGCCACCGGTGCAGCA
GCTGCTCGACCTGTACGACGTGCTCGGGGACGACGGCAAGCCGGGCACAGCGCTCCAGGACGAGGAGGAGGAC
GACGAGGAGCACGCCACCACCGAGACCGTCATGAGCATGGCCGCCGAGCCCAA**CCC**CGACTAGTCGACCAAAAA
CCGAAGTGCTGTTTTTTCTCCTTCAGCCCGAAGATCCAAGCGAGCCGCATCGTAAGGGCGCAGCTCTGGGTGCACT
TGCGCCCGGCGGATGAGGCGACAACGGTGTTCTTGCAGATATCGCGACTCATGCCCATCAAAGACGGGAGAAGG
CACGTACGAATACGTTCGCTGAAGATCGACGTGGACGCAGGAGTCAGTTCGTGGCAGAGCATCGACGTGAAGCA
GGTGCTTGCGGTGTGGCTGAGGCAGCCGGAAACCAACTGGGGGATTGAGATCAACGCGTTCGACTCCAAAAGCA
ACGATCTCGCGATCACTTCTGCGGAGCCTGGAGAAGAGGGACTGCTCCCGTTCTTGGAGGTGAAAATTTCTGAAG
TTCCAAAGCGAACCAGGAGAGAATCAGGACTAGACTGTGATGAGAATTCGTCCGAGTCCCGCTGCTGCCGCTACC
CCCTTACGGTGGACTTTGAAGACTTCGGCTGGGACTGGATTATTGCCCCAAAACGCTACAAGGCCAACTACTGCTC
GGGCGAGTGCGACTACGTGCACTTGCAGAAGTACCCGCATACACACTTGGTGAACAAGGCCAACCCACGTGGCAC
TGCCGGCCCCTGCTGCACGCCCACCAAGATGTCTCCCATCAACATGCTCTACTTCAACGGAAAAGAGCAGATCATC
TACGGCAAGATCCCCTCCATGGTAGTGGATCGCTGTGGCTGCTCG**TGA**

Predicted AA:

MHLAQVLISLGFVVAFGPMARTDTGAPEQQQQQQQPTAVTEEREAQCSAASACAFRQHSKQLRLQAIKSQ
ILSKLRLKQAPNVSRDVVKQLLPKAPPVQQLLDLYDVLGDDGKPGTALQDEEEDDEEHATTETVMSMAAE
PNPDSTKNRSAVFSPSARRSKRAASGRSSGCTCARRMRRQRCSCRYRDSCPSKTGEGTYEYVRRSTWTQE
SVRGRASTSRCLRCGGSRKPTGGLRSTRSTPKATISRSLLRSLEKRDCSRSWRKFLKFQSEPGENQDTVM
RIRPSPAAAATPLRWTLKTSAGTGLLPQNATRPTTARASATTCTCRSTRIHTWTRPTHVALPAPAARPPR
CLPSTCSTSTEKSRSSTARSPPWWIAVAAR