

國立臺灣大學生命科學院生命科學系



碩士論文

Department of Life Science

College of Life Science

National Taiwan University

Master Thesis

*Islet* 基因調節序列的演化與脊椎動物成對的鰭  
起源有關

Evolution of *Islet* Gene Regulation is Linked to the Origin  
of Paired Fins in Vertebrates

許子易

Tzu-Yi Hsu

指導教授：李士傑 博士

Advisor: Shyh-Jye Lee, Ph.D.

中華民國 112 年 8 月

August 2023



國立臺灣大學碩士學位論文  
口試委員會審定書  
MASTER'S THESIS ACCEPTANCE CERTIFICATE  
NATIONAL TAIWAN UNIVERSITY

Islet 基因調節序列的演化與脊椎動物成對鰭的起源有關

Evolution of *Islet* Gene Regulation is Linked to the Origin of Paired Fins  
in Vertebrates

本論文係 許子昂 (姓名) R11B21001 (學號) 在國立臺灣大學  
生命科學 (系/所/學位學程) 完成之碩士學位論文，於民國 112 年  
7 月 3 日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Department / Institute of Life Science  
on 3 (date) Jul (month) 2023 (year) have examined a Master's thesis entitled above presented  
by Tzu-Yi Hsu (name) R11B21001 (student ID) candidate and hereby certify  
that it is worthy of acceptance.

口試委員 Oral examination committee:

<u>李士偉</u> (指導教授 Advisor)	<u>薛收璇</u>	<u>游育凱</u>
<u>郭慧珍</u>	<u>黃淑萍</u>	

系主任/所長 Director: 鄭怡生

## 致謝



假使時間重來，我會義無反顧的選擇同一間實驗室。四年來受李士傑老師的薰陶如沐春風，獲益良多。老師在大學時便誘導了我對發育的好奇心，並為我孕育了一個能自由探索的環境，讓我有空間探索興趣、發揮潛力。老師更教我做人比做科學更難的道理，這點我終身受用。謝謝蘇怡璇與游智凱老師在演化發育學領域樹立典範，不但在課堂裡引我入門，更總是額外花時間關心我的研究，並且常常在我遇到瓶頸時給予建議。謝謝郭典翰老師教我做學問的嚴謹態度，也帶我認識森羅萬象的無脊椎動物。謝謝黃禎祥老師在科學數據解讀上的指引，並且在我面臨人生選擇題時幫助我少踩了很多地雷。本論文得以完成是受老師們指教，在這裡致上最深刻的謝意。

倒轉著時光，我與實驗室夥伴的回憶歷歷在目，很高興也很珍惜能在你們的陪伴下度過這些日子。從認識最久的 Mark、Jean、Simon、Tina、Janet、Chris、Andy、Eve、Eugene、Paul、Scott，每一位都在我心中留下了深刻的刻痕，你們使我秉持著對於研究的熱忱和興趣，繼續朝著這條道路邁進，也祝福你們都能朝自己期待的人生道路前進。更謝謝我的家人和朋友，若沒有你們，這份研究不可能來臨。

## 中文摘要



脊椎動物成對附肢的起源和演化是演化發育生物學中一個長久存在的問題。最近的化石證據發現有頷類祖先具有一個允許成對側鰭發育的「鰭潛能區」，而此區域在演化過程被局限演化出現存有頷類動物單基底的側肢。我試圖通過研究成對後肢發育過程的保守轉錄因子來了解鰭潛能區的分子組成。我針對有頷類演化支動物取樣，在不同物種的腹鰭中皆發現了 *Islet* 同源基因的保守表達，並確定了一個名為 CR2 的保守增強子可調控此基因表達。我同時證明在斑馬魚胚中過度表達 *isll* 可在少數個體誘導出一個類似鰭的結構，此類鰭構造可表達數個鰭的標誌基因，這表明了 *isll* 足以誘導鰭的發育。因此我假設祖先的「鰭潛能區」可能是由 *Islet* 同源基因所組成，並受到其增強子 CR2 的調控，而增強子序列的改變可能成為單基底鰭演化的基礎。

關鍵字：成對附肢、演化、*Islet* 基因、有頷類、鰭潛能區

## Abstract



The origin and evolution of vertebrate paired appendages is a long-standing question in evolutionary developmental biology. Recent fossil discoveries indicate the presence of an ancestral “zone of fin competence” where paired lateral fins emerge, and the subsequent restriction of this zone led to a monobasal fin seen in extant gnathostomes. I sought to identify the molecular components of the zone of fin competence by examining the conserved factors in paired pelvic appendage development. I sampled the gnathostome clade, found a conserved expression of *Islet* orthologs in the pelvic fin, and identified a conserved enhancer CR2 driving such expression. I also show that overexpression of *isll* in zebrafish embryos induces outgrowth of a fin-like structure expressing several fin marker genes, suggesting the sufficiency of *isll* to initiate a fin program. I propose that the ancestral “zone of fin competence” is established by *Islet* genes, which are regulated by the enhancer CR2. Modification to the enhancer may underlie the evolution of monobasal fins.

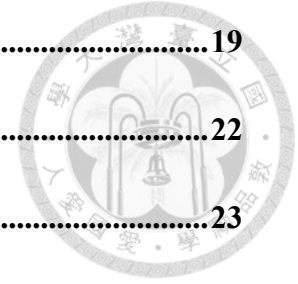
Key words: Paired appendages, Evolution, Islet gene, Gnathostomes, Zone of fin competence

# Table of contents

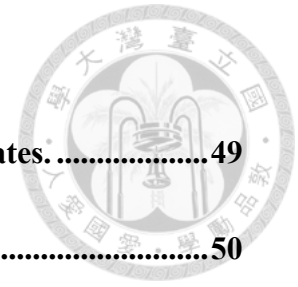


致謝.....	ii
中文摘要.....	iii
Abstract.....	iv
Introduction.....	1
Materials and Methods.....	6
Animals .....	6
cDNA sample preparation and quantitative polymerase chain reaction (qPCR) .....	7
Genomic DNA collection and genotyping .....	7
Whole-mount in situ hybridization (WISH).....	8
Phylogenetic Analysis.....	10
Enhancer identification .....	11
Microinjection .....	12
isl1 overexpression by heat shock.....	13
Transgenic Assays .....	13
CRISPR/Cas .....	14
Synteny map .....	15
Results .....	16
Islet orthologs are expressed in the developing pelvic appendages of gnathostomes. ....	16

Identification of enhancers regulating Islet gene expression .....	19
Generation of CR2 mutant fish .....	22
Functional validation of Isl1 activity.....	23
CR2 transcription factor binding motif.....	24
Discussion.....	28
Perspectives and future studies.....	32
Comparison between zebrafish and lamprey isl2a .....	32
On the chordate homolog of the cranial sensory ganglia .....	33
ISL1 and ISL2 complementarity .....	34
Lack of CR2 homology beyond gnathostomes .....	34
References .....	36
Tables.....	45
Table 1. List of primers used.....	45
Figures.....	48



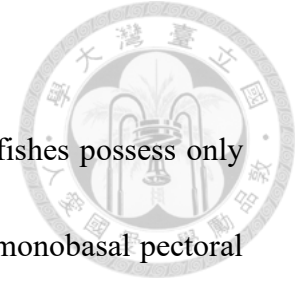
## List of figures



<b>Figure 1. Conserved synteny between amphioxus and vertebrates.....</b>	<b>49</b>
<b>Figure 2. Inferred Islet gene duplication tree.....</b>	<b>50</b>
<b>Figure 3. Phylogenetic relationship of ISLET orthologs.....</b>	<b>51</b>
<b>Figure 4. Whole-mount in situ hybridization (WISH) staining of Islet orthologs in vertebrates.....</b>	<b>52</b>
<b>Figure 5. Sequence comparison identifies conserved putative enhancers of Islet genes.....</b>	<b>54</b>
<b>Figure 6. CR2 drives reporter expression in the pelvic fin.....</b>	<b>56</b>
<b>Figure 7. CRISPR/Cas9 design to knockout CR2.....</b>	<b>57</b>
<b>Figure 8. Is11 overexpression induces the outgrowth of a fin with an unspecified identity.....</b>	<b>59</b>
<b>Figure 9. Probing the transcription factors of CR2.....</b>	<b>61</b>



## Introduction



While extant jawless vertebrate species such as lampreys and hagfishes possess only median fins, extant jawed vertebrates (gnathostomes) have paired monobasal pectoral and pelvic appendages supported by an endoskeleton. The paired appendages subsequently diversified into various forms, including the wing-shaped pectoral fins of batoid fishes (skates and rays), the elongated limbs of land animals, and the feathered wings of birds. Despite being morphologically distinct, these appendages are considered deeply homologous as they share a similar developmental process and core genetic regulatory network (GRN) (Shubin et al., 2009). Therefore, the question arises of how paired appendages first emerged in evolutionary history, and how deployment of the core GRN may contribute to this event.

Based on morphological observations of extant animals, several hypotheses were proposed regarding the emergence of paired appendages (Diogo, 2020). According to the lateral fin-fold theory, the paired pectoral and pelvic fins are the remnants of a continuous lateral fin fold along the trunk of an ancestor animal. This hypothesis posits that the fin fold between the pectoral and pelvic region was reduced. However, fossil evidence that supports this hypothesis has been limited until recently (Gai *et al.*, 2022), and there is little discussion regarding the molecular mechanism of fin fold reduction.

The second hypothesis, Gegenbaur's gill arch hypothesis, recognizes the developmental similarity between the gill arches and the paired pectoral fins and suggests that the paired pectoral appendages evolved by modifications to the last gill arch (Gegenbaur *et al.*, 1878). The hypothesis was extended to indicate that paired fins evolved by co-optioning genetic networks that pattern gill arches.



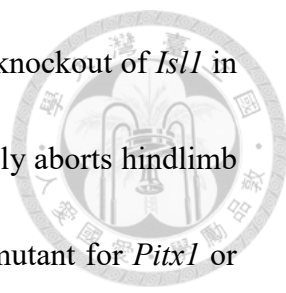
Fossil discoveries bridging the gap between extant species provide further insights into how morphologically distinct structures evolved, reshaping some hypotheses. Still a jawless species, the Silurian vertebrate Galeaspids *Tujiaaspis* seem to have had paired lateral fin-folds at the branchial to caudal level, indicating the presence of a zone of fin competence (Gai *et al.*, 2022). These fin-folds were thought to lack endoskeletal support and can only generate lift passively (Gai *et al.*, 2022). The Osteostracans, which branched off later along the lineage leading to gnathostomes, evolved paired pectoral appendages with skeletal support (Janvier *et al.*, 2004). Their monobasal fins are followed posteriorly by paired, elongated ventrolateral ridges. Therefore, it seems that an ancestral zone of fin competence was subdivided rostrocaudally, and the articulation of the pectoral appendage occurred before the pelvic appendage (Gai *et al.*, 2022). Later, the pelvic fin fold was restricted and differentiated to become the monobasal pelvic appendages seen in extant jawed vertebrates. These fossils exhibit compatibility with

the fin fold hypothesis and further demonstrate that restriction of the fin fold occurred in a pectoral-before-pelvic fashion.



A central focus of evolutionary developmental biology is to identify the genetic mechanisms underlying the evolution of animal design. Previous studies linked the subdivision of the zone of fin competence with the duplication of *Tbx4/5* (Minguillon *et al.*, 2009; Tanaka *et al.*, 2002). Whereas the amphioxus has a single *Tbx4/5* gene, in gnathostomes, it was duplicated into *Tbx4* and *Tbx5*, whose expression became restricted to the pelvic and pectoral fins, respectively. In addition, the cooption of an existing GRN from either the median fin fold or the gill arch has been linked with the differentiation and positioning of the fin endoskeleton (Diogo, 2020). However, the question remains what underlies the emergence of the first lateral fin fold, and how it became restricted in gnathostomes.

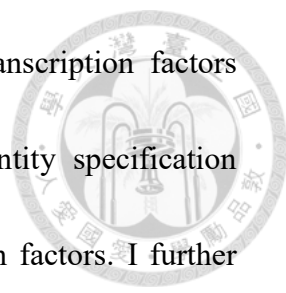
Here, I seek to probe the molecular underpinnings of paired fin emergence by identifying the conserved transcription factors in paired fin development. I focus on the lesser-studied pelvic appendage. Past studies identified *Pitx1*, *Tbx4*, and *Isl1* to be essential for pelvic appendage development in mice (Duboc *et al.*, 2021; Itou *et al.*, 2012; Kawakami *et al.*, 2011; Marcil *et al.*, 2003; Szeto *et al.*, 1999). Among them, the



expression of *Isl1* initiates the at earliest timepoint and conditional knockout of *Isl1* in the posterior mesendoderm (using the *T-cre; Isl1<sup>-flox</sup>* line) completely aborts hindlimb bud protrusion (Kawakami et al., 2011). On the other hand, mice mutant for *Pitx1* or *Tbx4* still develop hindlimbs with ossified endoskeleton (Duboc et al., 2021). Therefore, I consider *Isl1* a transcription factor necessary for hindlimb development, and aim to study the function of *Isl1* in the evolution and development of paired appendages.

I identified the conserved expression of *Isl1et* orthologs in the pelvic region of zebrafish, an actinopterygian, and the bamboo shark, a cartilaginous fish. In addition, By sequence analysis, I found a conserved enhancer region CR2 of downstream of the *Isl1et* locus that can drive the expression of enhanced green fluorescent protein (EGFP) in the pelvic fin. The CR2 enhancer is specific to gnathostomes, so the emergence of this enhancer may be linked with the initiation of pelvic fins.

I conditionally overexpressed *isl1* in zebrafish embryos and found ectopic fin outgrowth in a small population of zebrafish embryos, indicating that *isl1* can activate fin competency. This ectopic fin-like outgrowth expressed marker genes associated with fin polarization (*shh*) and differentiation (*and1*). Therefore, I propose that the CR2 enhancer drives *Isl1* and *Isl2* expression to define the zone of fin competence in



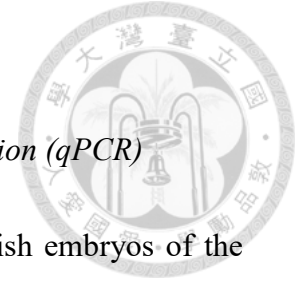
gnathostomes. Notably, the fin-like structure did not express transcription factors specific to the pelvic appendage. Thus, the pectoral/pelvic identity specification invokes additional stepwise co-option of downstream transcription factors. I further demonstrate that a gnathostome-specific Gata-binding site in the CR2 enhancer is required for the inhibition of CR2 activity in the trunk region. Therefore, I propose that evolution of the Gata binding site plays a role in constraining the expression of *Islet* genes, thereby delineating the zone of fin competence in the monobasal position of extant gnathostomes. These experiments demonstrate a conserved role of *Islet* orthologs during paired appendage development and provide insights into the molecular mechanisms underlying pelvic appendage origin and evolution.

## Materials and Methods



### *Animals*

Zebrafish (*Danio rerio*) were maintained at 28.5°C under a 14-hour light, 10-hour dark cycle. Embryos were collected and incubated in 0.3X Danieau's buffer (17.4 mM NaCl, 210 μM KCl, 120 μM MgSO<sub>4</sub>, 180 μM Ca(NO<sub>3</sub>)<sub>2</sub>, 1.5 mM HEPES in double-distilled water, pH 7.2). The developmental stages were defined according to Kimmel *et al.* (1995). Spotted gar (*Lepisosteus oculatus*) juvenile (12-14 cm) was purchased from a domestic aquarium (Suwenyen Aquarium, Kaohsiung) and maintained at 25°C in freshwater. White-spotted bamboo shark (*Chiloscyllium plagiosum*) eggs were purchased from a domestic aquarium (Ideal Jellyfish City (理想水母城), Pingtung). The eggs were incubated at 25°C in seawater prepared from marine salt mix (Coralife) to a density of 1026 kg/m<sup>3</sup>. The embryos were staged in reference to a comparable species, the brownbanded bamboo shark (*Chiloscyllium punctatum*) (Onimaru *et al.*, 2018). Fixed Japanese lamprey (*Lethenteron camtschaticum*) embryos were a kind gift of Daichi G. Suzuki, University of Tsukuba, Tsukuba. Embryos were fixed in PFA, stored in methanol, and shipped from Japan to Taiwan in dry ice. All experimental procedures were approved by the use of the laboratory animal committee at National Taiwan University, Taipei, Taiwan (IACUC Approval ID: 111 Animal Use document No. 00042) and carried out following the approved guidelines.



### *cDNA sample preparation and quantitative polymerase chain reaction (qPCR)*

Mixed-stage zebrafish cDNA samples were prepared from zebrafish embryos of the following stages: 1-cell, 4-cell, 8-cell, 4 hours post-fertilization (hpf), 30% epiboly, 6 hpf, 80% epiboly, 10 hpf, 12 hpf, and 1, 2, 3, 4, 5, 6 days post-fertilization (dpf) embryos.

Bamboo shark cDNA were prepared from one bamboo shark head at stage 28 and the extraembryonic tissues of several bamboo shark embryos at stage 27-28. RNA was extracted using the RNazol® RT kit (Molecular Research Center Inc., Cincinnati, OH).

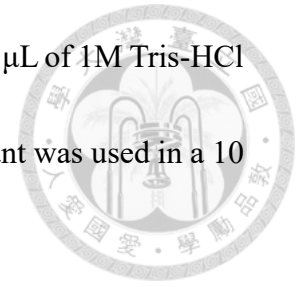
500  $\mu$ L RNazol® solution was used to extract RNA from 20 embryos for each zebrafish sample. RNA quality was confirmed by the visualization on 1% agarose gel in TAE buffer and the measurement of RNA absorbance. The isolated RNA was reverse-transcribed using M-MLV Reverse Transcriptase (Promega, Madison, MI) and oligo dT primers to obtain the first-strand cDNA, stored at  $-20^{\circ}\text{C}$  for further applications.

Japanese lamprey cDNA samples of stages 22-26 were a courtesy of Daichi G. Suzuki, University of Tsukuba, Tsukuba.

### *Genomic DNA collection and genotyping*

To prepare zebrafish and spotted gar genomic DNA, the fish was anesthetized before cutting a piece of the tail fin. The tissue was alkaline-lysed by adding 100  $\mu$ L of 50 mM

NaOH and heating to 98°C for 30 minutes, then neutralized with 10  $\mu$ L of 1M Tris-HCl pH8. The mixture was then centrifuged, and 0.5  $\mu$ L of the supernatant was used in a 10  $\mu$ L PCR reaction.

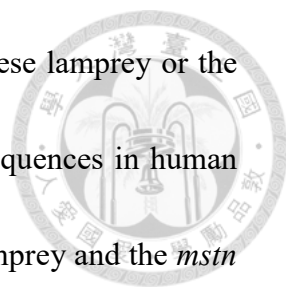


Live-genotyping of 4 dpf zebrafish larvae was performed according to Zhang et al. (2020) with modifications. Embryos were raised in Danieau's buffer instead of the E3 medium. The lysis buffer was not added following DNA collection to avoid inhibiting of downstream PCR by EDTA. In addition, 3 washes with Danieau's solution instead of the recovery solution to prevent sharp changes in osmolarity.

#### *Whole-mount in situ hybridization (WISH)*

WISH probes were constructed by cloning target gene cDNA into the pGEM®-T Easy Vector (Promega). The forward and reverse primers targeting different exons were used to amplify a sequence of 700-1000 bp. The plasmid was linearized after validation by sequencing and used to transcribe antisense DIG-labelled RNA probes using the MEGAscript® SP6 Transcription Kit (Invitrogen, Waltham, MA). Products were cleaned with the MicroSpin G-50 Column (Cytiva, Buckinghamshire, UK) to remove unincorporated nucleotides.





For the genes that are not annotated on the genomes of Japanese lamprey or the white spotted bamboo shark, we performed a blast search using sequences in human and zebrafish as query. Such genes include the *Islet* orthologs in lamprey and the *mstn* ortholog in bamboo shark. The blastn algorithm with default parameters was used to search for “somewhat similar sequences”. Primers were designed for all of the identified hits to clone them for WISH probe synthesis.

WISH on zebrafish, bamboo shark, and lamprey was performed according to Thisse and Thisse (2008) with slight modifications. Conditions of Proteinase K treatment were as follows: 10 µg/ml 15 min for 1 dpf embryos, 50 µg/ml 18 min for 2 dpf embryos, 70 µg/ml 15 min for 3-6 dff embryos and stage 22-26 lamprey embryos. For 22-26 dpf larvae and stage 27-28 bamboo shark embryos, 70 µg/ml proteinase K treatment for 20 min was sufficient to allow probe penetration. For bamboo shark embryos, two additional washes of 10 min with 100% hybridization medium was included after the probe hybridization step to allow higher stringency (O'Neill et al., 2007). After staining, the embryos were progressively dehydrated by gradually replacing to 100% methanol and stored at -20°C overnight to allow sharper staining contrast. The embryos were mounted in glycerol for observation and photography.

## Phylogenetic Analysis

Amino acid sequences were collected from the NCBI website (<https://www.ncbi.nlm.nih.gov/gene/>) and Ensembl website (<https://www.ensembl.org>)

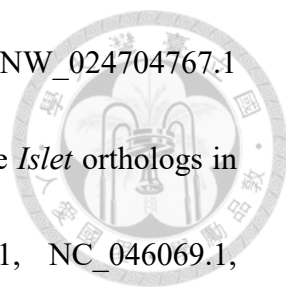


then aligned by MUSCLE (Edgar, 2004). The following sequences were used: *Homo sapiens* ISL-1 (NP\_002193.2) and ISL-2 (NP\_665804.1), *Mus musculus* ISL-1 (NP\_067434.3) and ISL-2 (NP\_081673.2), *Gallus gallus* ISL-1 (NP\_990745.1), *Xenopus tropicalis* ISL-1 (NP\_001016720.1) and ISL-2 (NP\_001159513.1), *Latimeria chalumnae* ISL-1 (XP\_005999085.1), ISL-1L (XP\_014345978.1), and ISL-2 (XP\_005995696.1), *Protopterus annectens* ISL-1 (XP\_043933573.1) and ISL-2 (XP\_043933401.1), *Erpetoichthys calabaricus* ISL-1 (XP\_028661461.1) and ISL-2 (XP\_028679339.1), *Lepisosteus oculatus* ISL-1 (XP\_015219661.1) and ISL-2 (XP\_006628961.1), *Danio rerio* ISL-1 (NP\_571037.1), ISL-1B (NP\_001002043.1), ISL-2A (NP\_571045.1), and ISL-2B (NP\_571039.1), *Takifugu rubripes* ISL-1 (XP\_029686106.1), ISL-2A (XP\_003967483.1), and ISL-3 (XP\_011607956.1), *Callorhinchus milii* ISL-1 (XP\_007890395.1) and ISL-2A (XP\_007906404.1), *Amblyraja radiata* ISL-1 (XP\_032886821.1), ISL-1L (XP\_032902406.1), and ISL-2 (XP\_032870769.1), *Chiloscyllium plagiosum* ISL-1 (XP\_043544206.1), ISL-1L (XP\_043577220.1), and ISL-2A (XP\_043536316.1), *Eptatretus burger* ENSEBUG00000011825, ENSEBUG00000004961, and ENSEBUG00000016346,

*Petromyzon marinus* ISL-1 (XP\_032829736.1), ISL-1L (XP\_032824418.1), and ISL-2L (XP\_032806737.1), *Ciona intestinalis* (NP\_001027767.1), *Styela clava* ISL-1L (XP\_039271423.1), *Branchiostoma belcheri* (XP\_019620436.1), *Branchiostoma floridae* (XP\_035666393.1). The construction of phylogenetic trees was performed by the maximum-likelihood method using the JTT+G model in MEGA (Tamura et al., 2021).

#### *Enhancer identification*

To identify conserved regulatory elements, a sequence 10-kb upstream and downstream *Isl1* or *Isl2* genomic regions were downloaded from the NCBI website (<https://www.ncbi.nlm.nih.gov/gene/>). For alignment of the *Isl1* locus, the following sequences were used: NC\_000005.10 (*Homo sapiens*), NC\_000079.7 (*Mus musculus*), NC\_052572.1 (*Gallus gallus*), NC\_030677.2 (*Xenopus tropicalis*), NC\_041400.1 (*Erpetoichthys calabaricus*), NC\_023180.1 (*Lepisosteus oculatus*), NC\_007116.7 (*Danio rerio islla*), NC\_007121.7 (*Danio rerio isl1b*), NW\_024704742.1 (*Callorhinchus milii*), NC\_045956.1 (*Amblyraja radiata*). For *Isl2*, the following sequences were used: NC\_000015.10 (*Homo sapiens*), NC\_000075.7 (*Mus musculus*), NC\_052541.1 (*Gallus gallus*), NC\_030679.2 (*Xenopus tropicalis*), NC\_041410.1 (*Erpetoichthys calabaricus*), NC\_023181.1 (*Lepisosteus oculatus*), NC\_007136.7




(*Danio rerio is11a*), NC\_007118.7 (*Danio rerio is11b*), NW\_024704767.1 (*Callorhinchus milii*), NC\_045999.1 (*Amblyraja radiata*). The three *Islet* orthologs in lamprey (*Petromyzon marinus*) were included: NC\_046120.1, NC\_046069.1, NC\_046108.1. The above sequences were aligned using the mVista SLAGAN program (Frazer et al., 2004) with the default parameters, referencing the human sequence. Sequence identity greater than 50% in a sliding window of 50 bp are considered conserved, and the enhancers were annotated.

#### *Microinjection*

Glass capillaries with dimensions of 1.14 mm outer diameter and 0.5 mm inner diameter (World Precision Instrument Inc., Sarasota, FL) were pulled using a horizontal puller (P-97, Sutter Instrument, Navato, CA). Embryos were immobilized in 1-mm ditches of a pre-made agar plate. Injection mixtures at 2.3 nL were injected into zebrafish embryos at the 1-cell stage using a Nanoliter injector (World Precision-Instrument Inc., Sarasota, FL). To generate transgenic zebrafish, co-injections of 25 pg tol2 transposon mRNA and 100 pg transgenic plasmids were made. For overexpression, up to 75 pg of plasmid was injected. For CRISPR/Cas, 100 pg of Cas9 protein and 100pg gRNA was injected. Plasmids were injected into the embryo proper to enhance gene incorporation.

### *isl1* overexpression by heat shock

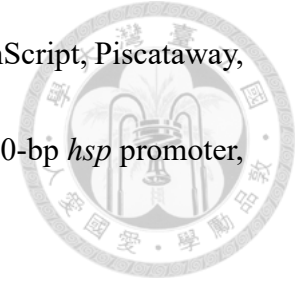


To generate the *hsp:isl1-2A-EGFP* construct, *isl1* was amplified from mixed-stage zebrafish cDNA, and the 2A-EGFP sequence was amplified from the Tol2kit (Kwan et al., 2007). The T2Khsp backbone was generated by digesting pT2KhspGGFF (Asakawa et al., 2008) using the NcoI and SmaI restriction enzymes (New England Biolabs, Gaithersburg, MD). The three DNA fragments (*isl1*, 2A-EGFP, and T2Khsp) were connected using the Gibson assembly method with the GenBuilder™ Plus Cloning Kit (GenScript, Piscataway, NJ). This resulting plasmid was injected in the embryo proper at the one cell stage, and the petri dish containing the embryos were placed in a 37°C air incubator for an hour for heat shock at the appropriate stages. Embryos that developed a protrusion were fixed with 4% PFA at 2 dpf and used for WISH staining.

### *Transgenic Assays*

To assay enhancer expression patterns, the *Isl1* and *Isl2* enhancers from human, gar, and zebrafish were inserted into the pT2KhspGGFF plasmid that drives Gal4-EGFP fusion protein expression (Asakawa et al., 2008). Primers were designed to clone all enhancer elements (CR1-3) and consistently be 3-3.5 kb in all three species. The pT2KhspGGFF plasmid was digested with ApaI, and the enhancers were inserted based

on Gibson assembly using the GenBuilder™ Plus Cloning Kit (GenScript, Piscataway, NJ). The resulting construct contains an enhancer, followed by a 600-bp *hsp* promoter, and the Gal4-EGFP sequence.

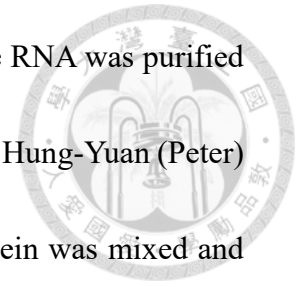


Transgenic zebrafish were generated by microinjecting the plasmids and *tol2* transposon mRNA as described. We synthesized mature *tol2* mRNA by *in vitro* transcription with the mMessage mMachine SP6 kit (Ambion, Austin, TX, USA). Injected embryos (F0) were raised to adult, and F0 adult whose offspring exhibited fluorescence signal after heat shock were kept as founders. For the examination of enhancer activity, F1 embryos or larvae are used to perform WISH.

#### *CRISPR/Cas*

gRNAs were generated from annealed DNA oligos with *in vitro* transcription. The template DNA oligo were order from Integrated DNA Technologies IDT. A 60 bp target-specific oligo and a 80 bp constant oligo with 23 bp overlap was annealed by incubating at 95 °C for 5 min, cooling at -2°C/sec to 85°C, then cooling at -0.1°C/sec to 25°C. The single-strand ends were filled with T4 polymerase (New England Biolabs, Graiciuno, Vilnius) at 12°C for 20 min. The product was used for *in vitro* transcription using the Ambion MEGAscript™ T7 Kit (Invitrogen, Waltham, MA). DNA template was

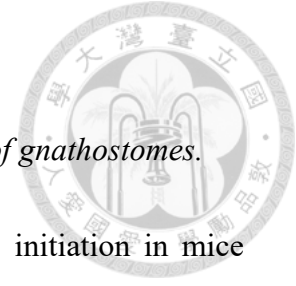
digested in resulting mixture using DNase (included in kit), and the RNA was purified following the manual instructions. Cas9 protein was kind gift of Dr. Hung-Yuan (Peter) Chi, National Taiwan University, Taipei. The gRNA and Cas9 protein was mixed and incubated at 37°C for 15 minutes prior to injections.



### *Synteny map*

Genomes of amphioxus (*Branchiostoma floridae*, GCF\_000003815.2), spotted gar (*Lepisosteus oculatus*, GCF\_000242695.1), and chicken (*Gallus gallus*, GCF\_000002315.6) was downloaded from NCBI. Synteny maps were generated by CoGe SynMap(Lyons & Freeling, 2008). The maximum distance between two matches are adjusted to include the whole genome, and the minimum number of aligned pairs is specified to 1. C-score was set to 0.99 to show the top mutual blast hit only. I generated the SVG Version of Syntenic Dotplot for the final visualization.

## Results

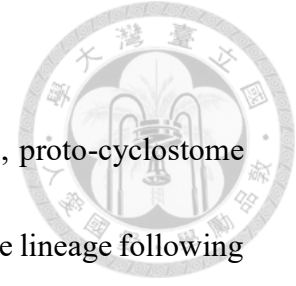


*Islet* orthologs are expressed in the developing pelvic appendages of gnathostomes.

*Isl1* is expressed in developing hindlimbs and is required for its initiation in mice (Kawakami *et al.*, 2011; Yang *et al.*, 2006). I thus asked whether *Islet* orthologs have a conserved role in pelvic appendage development in gnathostomes.

I first generated Oxford dot plots, as previously described, to infer the evolutionary history of *Islet* orthologs (**Figure 1**). Comparisons between genomes of amphioxus and vertebrates revealed densely clustered blocks of dots that reflect the conserved synteny. The single amphioxus *Islet* gene is located on chromosome 2 (BFL2), the two *Islet* homologs in spotted gar on LOC2 and LOC3, and the two *Islet* homologs in chicken on GGA10 and GGAZ. The distribution of these orthologs coincide with the conserved synteny blocks, supporting that the vertebrate *Isl1* and *Isl2* arose by whole genome duplication events. Indeed, according to the inferred chromosomal organization in (Simakov *et al.*, 2020), the gar chromosomes LOC2 and LOC3, and the chicken chromosomes GGA10 and GGAZ, belongs to the same chordate linkage group and were separated in the first round of genome duplication (1R) in the vertebrate ancestor. Therefore, this supported a duplication in the ancestor of vertebrates, and independent losses of two duplicates after the second genome duplication (2R) (**Figure 2**).





A recent study reconstructed the genomes of proto-vertebrate, proto-cyclostome and proto-gnathostome, and indicated a triplication in the cyclostome lineage following 1R (Nakatani *et al.*, 2021). We searched the orthologs of *Isl1* and *Isl2* in the dataset and found three orthologs in cyclostomes, all of which are located on the proto-cyclostome chromosome 13. Thus, the model from Nakatani *et al.* (2021) suggests that the three *Islet* orthologs in cyclostomes arisen by the triplication event followed by gene losses. However, the relationship of the three cyclostome *Islet* and the two gnathostome *Islet* cannot be defined due to the lack of a two-to-three orthology correspondence and an apparent quasi-random gene retention pattern (Nakatani *et al.*, 2021).

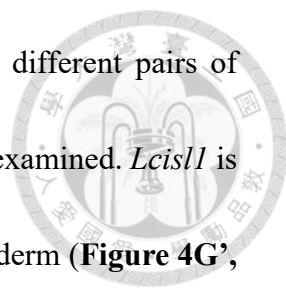
Given this information, I examined the phylogenetic relationships of the *Islet* orthologs on the neighbor-joining tree (**Figure 3**). Consistent with the model of Nakatani *et al.* (2021), the number of *Islet* orthologs in gnathostomes and cyclostomes exhibited a two-to-three ratio. *Isl1* and *Isl2* orthologs in gnathostomes form two distinct clusters, while the three *Islet* orthologs in cyclostomes cluster together. The ISL-1-like genes in lungfish, skate, and dog shark are clustered together with high support, suggesting that they are orthologous and may be derived from the 2R duplication and were not lost as in other gnathostome species. However, their relationship with the

cyclostome *Islet* genes is not well supported.



I examined *Islet* ortholog expression by WISH in selected species representing non-tetrapod gnathostomes: zebrafish (*Danio rerio*) of the actinopterygians (ray-finned fishes) and white-spotted bamboo shark (*Chiloscyllium plagiosum*) of the cartilaginous fishes. Additionally, I included the Japanese lamprey (*Lethenteron camtschaticum*) as an outgroup species for comparative analysis. *DrIsl1a*, *DrIsl2a*, and *DrIsl2b* are expressed in the central nervous system (including the tegmentum ganglia and spinal cord), prechordal plate, cardiopharyngeal mesoderm, and lateral plate mesoderm of 16 hpf and 24 hpf zebrafish embryos (**Figure 4A-D**). The expression of *DrIsl1b* is relatively weak at the stages examined and displays no specific staining. *DrIsl2a* and *DrIsl2b*, but not *DrIsl1a* and *DrIsl1b*, is expressed in the developing pelvic fin of metamorphosing zebrafish larvae (22-26 dpf). *Cpis11* and *Cpis12* are notably expressed in the pelvic region of stage 27-28 bamboo shark embryos when the pelvic fin emerged (**Figure 4E-F**). *Cpis11* and *Cpis12* are also expressed in the nervous system, consistent with the known role of *Islet* orthologs in metazoan nervous system development (Wong *et al.*, 2020).

I cloned two of the three annotated *Islet* homologs (*Lcis11* and *Lcis12*) in Japanese



lamprey. The amplification of *LcIsl1* was unsuccessful with two different pairs of primers, which might be due to its low expression level at the stage examined. *Lcisl1* is expressed in the gill arch, heart, and the posterior lateral plate mesoderm (**Figure 4G**, **G'**). The expression in the posterior lateral plate mesoderm is reminiscent of the *Isl2a* expression in 16 hpf zebrafish embryos. The expression of *LcIsl2* is restricted in the central nervous system, with the most robust expression in the cranial sensory ganglia (**Figure 4H**, **H'**). The expression in the nervous system is consistent with the known role of *Islet* orthologs in metazoan nervous system development (Wong et al., 2020).

Together, these results show a conserved expression of the *Islet* orthologs in the pelvic fin across gnathostomes. In addition, these orthologs exhibit varying degrees of specialization in different species. Thus, different enhancers may exist in the *Isl1/2* region of these species to regulate the overall expression pattern.

#### *Identification of enhancers regulating Islet gene expression*

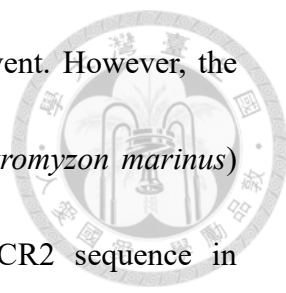
I sought to identify enhancers driving the expression of *Islet* genes in the pelvic fin. Three conserved regions (CR1-3) downstream of the *Isl1* locus are known to drive the reporter expression in the mouse heart (Kang et al., 2009). Given that the pelvic fin, akin to the heart, originates from the lateral plate mesoderm (Prummel et al., 2020), I

hypothesized that these conserved regions might also possess the potential to drive gene expression in the pelvic fin.



To examine the presence of those conserved regions in the *Isl1* and *Isl2* locus, I broadened the taxonomic input across gnathostomes by including the genomes of the non-teleost actinopterygians reedfish (*Erpetoichthys calabaricus*), spotted gar (*Lepisosteus oculatus*), cartilaginous fish white-spotted bamboo shark and thorny skate (*Amblyraja radiata*) in the comparison. Sequence alignment using the VISTA program (Frazer et al., 2004) revealed that CR1 and CR2 could be traced back to the ancestor of gnathostomes, as they are readily found in cartilaginous fishes (**Figure 5A**). CR1 is lost in actinopterygians, and CR1 and CR2 are absent in zebrafish. The loss of the zebrafish CR1 and CR2 downstream of *Isl1a/b* may account for the lack of *Isl1a/b* expression in the pelvic fin.

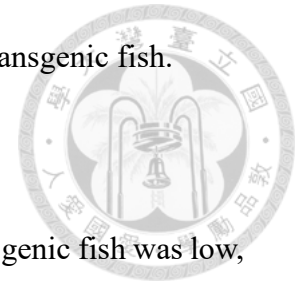
I further identified a region of sequence downstream of the *Isl2* locus that is conserved in gnathostomes (**Figure 5B**). Due to its similarity with CR2, I named this region the CR2-like sequence (**Figure 5D**). Both the CR2 and CR2-like sequences exhibit putative binding sites for transcription factors Smads and Gata (**Figure 5C, D**). The high similarity of CR2 and CR2-like sequence supports its emergence before the



duplication of *Isl1* and *Isl2* – that is, before the 1R duplication event. However, the homologs of CR2 could not be identified in the sea lamprey (*Petromyzon marinus*) using VISTA alignment. Furthermore, BLAST search of the CR2 sequence in cyclostomes (*Eptatretus burgeri*, *Lethenteron camtschaticum*, and *Petromyzon marinus*) or other basal chordates (*Branchiostoma floridae* and *Ciona intestinalis*) with a query using CR2 sequences from any species examined could not identify a hit (data not shown). It remains conceivable that rearrangement of the transcription factor binding sites masked the search. However, I did not find cases of Smad and Gata binding sites clustered closely (in a 50 bp range) near the upstream or downstream region of *Islet* orthologs in these species. Therefore, I conclude that while CR2 emerged in the ancestor of all vertebrates, the sequence conservation was lost in the cyclostome lineage.

To assess the activity of the CR2 enhancer, I generated transgenic reporter zebrafish lines by cloning a 3-kb region spanning CR1 (if present) and CR2 from human *Isl1*, spotted gar *Isl1*, and zebrafish *Isl2a* (**Figure 6A**). As a negative control, I also generated a transgenic line using the corresponding region from zebrafish *Isl1a*, which lacks CR1 and CR2 sequences. I linked the enhancer with an *hsp* minimal promoter and a Gal4-EGFP reporter. With this construct design, the enhancer could drive the EGFP reporter expression in the absence of heat shock, while after heat shock, the entire fish

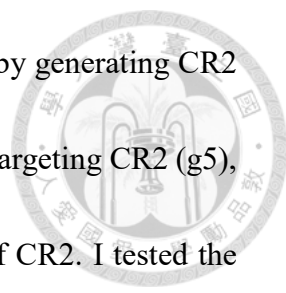
would express EGFP fluorescence that helped in the screening of transgenic fish.



As the fluorescent intensity of CR2-driven EGFP in those transgenic fish was low, I performed WISH against EGFP on F1 transgenic embryos to better visualize the signal (**Figure 6B**). The gar *Isl1* CR2 enhancer can drive the EGFP expression in the cardiopharyngeal mesoderm, prechordal plate, and posterior lateral plate mesoderm derivatives, recapitulating the endogenous mesodermal *DrIsl2a* expression. Similarly, the human CR2 *Isl1* enhancer displays activities in the lateral plate mesoderm and prechordal plate, although the expression is much less stringent. On the other hand, the zebrafish *DrIsl2a* enhancer exhibited robust and stringent activity in the posterior lateral plate mesoderm. As anticipated, no discernible activity of *DrIsl1a* was observed despite the successful transgenesis, as evidenced by the whole-body fluorescence after heat shock. Most importantly, regardless of which species the enhancer is cloned from, sequences spanning CR2 consistently drive reporter expression in the pelvic fin. It is thus possible that the evolution of the CR2 enhancer co-opts *Isl1/2* to the pelvic fin region.

#### *Generation of CR2 mutant fish*

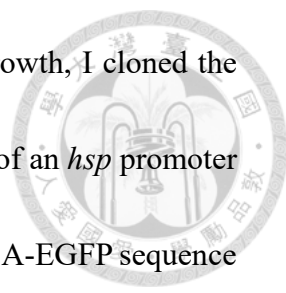
To examine the necessity of CR2 enhancer to drive mesodermal *isl2a* expression and



pelvic fin outgrowth, I planned to perform a loss of function assay by generating CR2 mutant zebrafish (**Figure 7A**). I synthesized guide RNAs (gRNAs) targeting CR2 (g5), the upstream (g1 and g2), or the downstream (g3 and g4) region of CR2. I tested the efficiency of individual gRNAs by genotyping 24 hpf embryos injected with the gRNA/Cas9 protein mixture. By capillary electrophoresis, the PCR products from embryos injected with g1 or g4 showed no noticeable difference compared to un-injected siblings. Signal distribution of the PCR products from embryos injected with g2, g3, or g5 was spread compared to un-injected siblings, indicating the presence of insertions or deletions (indels). Therefore, these gRNAs were used to generate CR2 mutant zebrafish.

I injected zebrafish embryos with a mixture of g2 and g3 with Cas9 protein. This combination could cause a large deletion to remove the entire CR2 region. I performed live embryo genotyping on those larvae at 4 dpf, and some injected animals had a disrupted CR2 out with multiple bands (**Figure 7B**). I also raised zebrafish injected with only g5, seeking to collect mutant fish with a disrupted Gata binding site (**Figure 7C**). These presumptive *isl2a* CR2 mutants will be assayed for pelvic fin development.

*Functional validation of Isl1 activity*

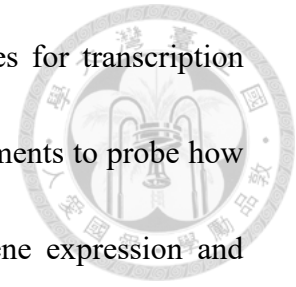


To test whether *Islet* genes have a role in inducing pelvic fin outgrowth, I cloned the *isl1* coding sequence from zebrafish and placed it under the control of an *hsp* promoter for heat shock-induced overexpression (**Figure 8A**). I also linked a 2A-EGFP sequence downstream of *isl1*, allowing visualization of *isl1*-overexpressing cells. I injected the construct at the 1-cell stage and performed heat shock at 6, 9, or 12 hpf to induce overexpression (**Figure 8B**). Over half of the embryos heat-shocked were deformed and formed a bent trunk. About 5-10% of the embryos developed an ectopic protrusion with a fin-like structure (**Figure 8C**). I consistently observed the EGFP signal associated with the ectopic protrusion (**Figure 8D**). The protrusions expressed *actinodin1* (*and1*), a fin fold marker, suggesting the activation of a fin developmental program (**Figure 8E**). *shha* is also expressed in the protrusion but is not polarized as in regular fin buds (**Figure 8E**). I further asked whether the ectopic protrusion has the identity of a pelvic fin. WISH Staining of pelvic appendage-specific transcription factors *tbx4* and *pitx1* revealed no apparent signal in the ectopic protrusion (**Figure 8E**). Taken together, these results suggest that *isl1* has a role in inducing a fin-like structure expressing several fin marker genes, but the polarity and pectoral/pelvic identity of the structure are not well defined.

*CR2 transcription factor binding motif*

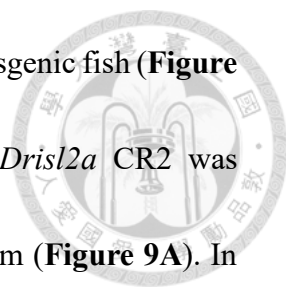


The conserved CR2 and CR2-like sequences exhibit binding sites for transcription factors Smads and Gata (**Figure 5C, D**). I designed several experiments to probe how each of the transcription factor binding motifs contributes to gene expression and phenotypic robustness.



Gata6 is an inhibitor that restricts *Isl1* expression to the posterior hindlimb bud in mice, and loss of Gata6 increased the number of cells expressing *Isl1* (Tahara et al., 2018). The sequence analysis reveals conservation of Gata binding sites in the CR2 and CR2-like enhancers. I hypothesized that the evolution of the Gata binding site restricts the expression of *isl1*, hence the zone of fin competence, to a single point in the posterior part of the developing embryo. If validated, this hypothesis could provide an explanation for the evolution from the paired elongated ventrolateral ridges of Osteostracans to the monobasal fins of modern jawed fishes.

To test how the Gata binding site affects CR2 activity, I employed conventional site-directed mutagenesis by PCR in the *Drisl2a* enhancer reporter construct (**Figure 6B**) to introduce an TA-to-AT replacement specifically within the Gata binding site (CR2 Gata<sup>mut</sup>), and verified by sanger sequencing (**Figure 9A**). Subsequently, the constructs were injected into 1-cell stage zebrafish zygotes and the transient EGFP



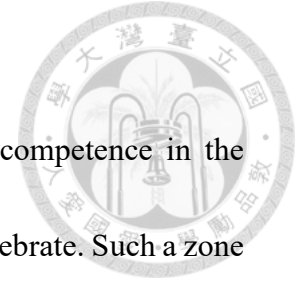
expression was documented at 16 hpf. Consistent with the stable transgenic fish (**Figure 6B**), the transient expression of EGFP driven by the wildtype *Drisl2a* CR2 was observed at the prechordal plate and posterior lateral plate mesoderm (**Figure 9A**). In contrast, the CR2 *Gata<sup>mut</sup>*-driven EGFP expression in the lateral plate mesoderm is expanded, spanning either along the entire trunk (6 out of 8 embryos) or the head region (2 out of 8 embryos) (**Figure 9A**). The activity in the prechordal plate and posterior lateral plate mesoderm is less stringent but is still observable. Using this reporter construct, I will raise stable transgenic fish to assay whether the reporter expression pattern is truly different from wildtype-CR2 reporter fish at the later stages.

CR2 also contains conserved Smad binding sites, which are targeted by the Smad transcription factors of the TGF $\beta$  family. I thus hypothesized that TGF $\beta$  signaling acts upstream of *Isl1/2* to induce fin protrusion. The TGF $\beta$  member *Gdf11* acts upstream of *Isl1* during mice axial development (Jurberg et al., 2013). As *Gdf11* and *Mstn* have overlapping functions (McPherron et al., 2009), I sought and found both *gdf11* and *mstnb* expression in the protruding pelvic fin bud of zebrafish larvae (**Figure 9B, C**). And while I could not amplify the *gdf11* ortholog in the white-spotted bamboo shark, its *mstn* ortholog is also expressed in the pelvic region (**Figure 9D, D'**). This implies that the activation of TGF $\beta$  signaling is conserved in gnathostome pelvic fin

development. Further functional tests will be performed to assay the requirement of TGF $\beta$  signaling during paired fin protrusion and development.



## Discussion



Fossil discoveries have indicated the presence of a zone of fin competence in the ventrolateral region from the gill to the tail level in the ancestral vertebrate. Such a zone of fin competence was preserved and evident as the continuous fin fold of galeaspid (Gai et al., 2022). The zone was divided in osteostracans and further restricted in extant gnathostomes (Janvier et al., 2004). Thus, the evolution of skeletal support, the subdivision of fin competence along the body axis, and the subsequent restriction and differentiation of the pelvic fin fold highlight the gradual morphological changes that occurred during the evolution of fins in vertebrates.

Here, I demonstrate that *Islet* orthologs may contribute, at least in part, to establishing the zone of fin competence. Examining representative species of the gnathostome clade, I found a conserved *Islet* gene expression in the paired pelvic appendages associated with the early stage of appendage outgrowth. Using a combination of genome sequence analysis and transgenesis, I identified the enhancer CR2 that drives reporter expression in the posterior lateral mesoderm of zebrafish embryos and the protruding pelvic fin of zebrafish larvae. CR2 has conserved binding sites for transcription factors such as Smads and Gata, which collectively determine the activity of the enhancer. We inferred that CR2 emerged in the vertebrate ancestor.

Despite lacking a sequence homolog of CR2, lamprey maintained the expression of *Islet* orthologs in the posterior lateral plate mesoderm. Therefore, the data presented here presents a previously undescribed deep conservation of *Islet* gene expression pattern in spite of the lack of regulatory sequence identity or morphological homology, suggesting a homolog of the zone of fin competence in ancestor vertebrates.

Using a conditional gain-of-function experimental setting, I demonstrate that *Isl1* can induce ectopic outgrowth of a fin-like structure expressing the fin fold marker *and1* and polarization marker *shh* in zebrafish. The lack of pelvic appendage-specific markers *tbx4* and *pitx1* indicates that *isl1* define the zone of fin competency but not fin identity in zebrafish. Therefore, the co-option of downstream transcription factors is required for specifying the pectoral/pelvic identity of the appendages in the evolution of pelvic fins. Given the function of *Islet* in fin development and the activity of its regulatory element CR2 in the pelvic region, maintaining an *Islet* gene expression domain controlled by CR2 was likely an essential step in acquiring paired appendages.

I further hypothesized that the evolution of the Gata binding site restricts the *Islet* expression domain, hence the zone of fin competence, to a single point in the posterior part of the developing embryo. This could explain the evolutionary transition from the

ventrolateral fin fold seen in osteostracans to the monobasal fin in extant gnathostomes.

To test this, I asked whether the loss of the Gata binding site allows the expansion of CR2 activity to the trunk region. First, I designed a gRNA targeting the nucleotide sequence adjacent to the Gata binding site to obtain zebrafish embryos containing a mutated Gata binding site (**Figure 6**). Second, I synthesized a reporter construct driven by a CR2 sequence harboring a disrupted Gata binding site. Using this reporter construct, I will raise stable transgenic fish to assay whether the reporter expression pattern is noticeably different from wildtype-CR2 reporter fish. My experimental approach using the two strategies will enable us to assess whether the absence of CR2 activity in the lateral plate mesoderm of the trunk of zebrafish embryos is due to the inhibitory action of Gata.

While I propose the emergence of the Gata binding site to restrict the zone of fin competence, it is crucial to gather further evidence to establish this evolutionary event at the cis-regulatory level. It is necessary to ascertain that the expression of Gata in the trunk region is conserved during the evolution of gnathostomes from their ancestor. One way to test this is to examine Gata expression in outgroup species such as lamprey or the more basal chordate amphioxus. However, interpretation of the results may be challenging as vertebrate evolution accompanied the expansion of Gata orthologs.

Subfunctionalization, neofunctionalization, and developmental system drift of the duplicated genes could make ortholog mapping more intricate.



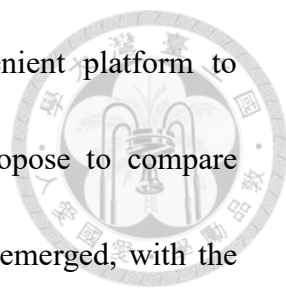
## Perspectives and future studies



### *Comparison between zebrafish and lamprey *isl2a**

My data indicate that the expression of *Islet* gene in the lateral plate mesoderm does not consistently coincide with fin protrusion. This observation is exemplified by the expression of *Lcis11* in lamprey and *Isl2a* in 16 hpf zebrafish embryos, where no concurrent fin protrusion was observed. Nevertheless, the overexpression of *isl1* in zebrafish embryos induced ectopic protrusions. These findings contradict in suggesting whether *Islet* gene expression alone is sufficient to initiate fin development. I propose several explanations to resolve the contradiction. First, *Islet* may induce fin outgrowth only when its dosage exceeds a certain threshold, which is achieved in our overexpression experiment. Utilization of single-copy insertions could provide finer control over the dosage of *Islet*, enabling a more comprehensive investigation into its dosage-dependent effects on fin development (Hwang et al., 2013; Mosimann et al., 2013). Second, *Islet* may be regulated beyond its RNA level. Past studies demonstrated that ISL protein is localized in the cytoplasm of zebrafish embryos and is in the nucleus of the protruding zebrafish pelvic fin (Moriyama et al., 2019). In addition, ISL forms complexes with other proteins, such as Lhx3 or Ldb1, which may alter its binding site recognition (Caputo et al., 2015; Thaler et al., 2002).





The CR2 transgenic reporter fish could serve as a convenient platform to understand the role of *Islet* in the lateral plate mesoderm. I propose to compare fluorescent cells in 16 hpf zebrafish, when the pelvic fin not yet emerged, with the fluorescent cells in the protruding pelvic fin of metamorphosing zebrafish (at around 21 days). A fluorescent activated cell sorting (FACS) followed by RNA-seq could reveal the transcriptomic changes in the same tissue at different stages, potentially resolving the molecular mechanism underlying the regulation of fin competency establishment and protrusion induction.

*On the chordate homolog of the cranial sensory ganglia*

A recent study that examined *hmx* gene expression in chordates suggested the homology of bipolar tail neurons (BTN) of the ascidian with the cranial sensory ganglia (CSG) of vertebrates (Papadogiannis et al., 2022). Data showed the expression of *Lcisl2* in the cranial sensory ganglia of lamprey. Combined with published data that showed *islet* expression in the BTN (Stolfi et al., 2015), it also supports the homology of BTN and the CSG. Examining how ISL1 and ISL2 cooperates with HMX transcription factors in specifying neural fate would be interesting. As *Islet* orthologs play a profoundly conserved role in metazoan nervous system development, we could trace the origin of different neural subtypes further back in evolution.



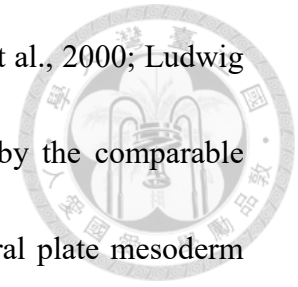
### *ISL1 and ISL2 complementarity*

Given the conservation of CR2 and CR2-like sequences in gnathostomes, I presume that the ancestral gnathostome expresses both *Isl1* and *Isl2* in the pelvic fin. However, our investigation in zebrafish revealed the absence of *DrIsl1a* and *DrIsl1b* and the presence of *DrIsl2a* and *DrIsl2b* expression in the pelvic fin. This observation raises the possibility that *Isl1* and *Isl2* are complimentary. Indeed, *Isl1a* and *Isl2a* still exhibit a significant sequence identity of 75%, with the homeodomain, which is crucial for DNA binding, sharing 98% identity. Therefore, it is not surprising that despite zebrafish pelvic fin development does not coincide with the expression of *DrIsl1a*, the overexpression of *DrIsl1a* still leads to the induction of ectopic fin outgrowth. Nonetheless, further biochemical assays would be necessary to discern the precise functional distinctions of ISL1 and ISL2, but such investigation extends beyond the scope of our current study.

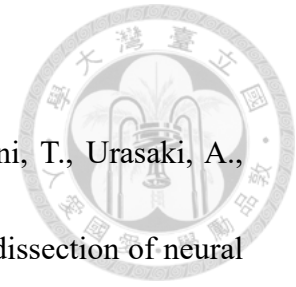
### *Lack of CR2 homology beyond gnathostomes*

While I could not identify CR2 homologs in lamprey or other basal chordates based on sequence identity, I do not rule out the possibility that these species possess regulatory elements functionally similar to CR2. Such developmental systems drift concerning

enhancer code utilization is not uncommon in evolution (Ludwig et al., 2000; Ludwig et al., 1998; True & Haag, 2001). This possibility is supported by the comparable expression patterns of *Isl1/2* orthologs in lamprey's posterior lateral plate mesoderm and 16 hpf zebrafish embryos. It could be rewarding to survey the *Islet* loci to map the regulatory sequences and determine to what degree such regulatory sequences are conserved throughout evolution.



## References



Asakawa, K., Suster, M. L., Mizusawa, K., Nagayoshi, S., Kotani, T., Urasaki, A.,

Kishimoto, Y., Hibi, M., & Kawakami, K. (2008). Genetic dissection of neural circuits by Tol2 transposon-mediated Gal4 gene and enhancer trapping in zebrafish. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(4), 1255-1260. <https://doi.org/10.1073/pnas.0704963105>

Caputo, L., Witzel, H. R., Kolovos, P., Cheedipudi, S., Looso, M., Mylona, A., van

IJcken, W. F. J., Laugwitz, K. L., Evans, S. M., Braun, T., Soler, E., Grosveld, F., & Dobрева, G. (2015). The Isl1/Ldb1 Complex Orchestrates Genome-wide Chromatin Organization to Instruct Differentiation of Multipotent Cardiac Progenitors. *Cell Stem Cell*, *17*(3), 287-299.

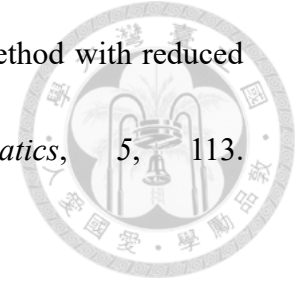
<https://doi.org/10.1016/j.stem.2015.08.007>

Diogo, R. (2020). Cranial or postcranial-Dual origin of the pectoral appendage of

vertebrates combining the fin-fold and gill-arch theories? *Dev Dyn*, *249*(10), 1182-1200. <https://doi.org/10.1002/dvdy.192>

Duboc, V., Sulaiman, F. A., Feneck, E., Kucharska, A., Bell, D., Holder-Espinasse, M.,

& Logan, M. P. O. (2021). Tbx4 function during hindlimb development reveals a mechanism that explains the origins of proximal limb defects. *Development*, *148*(19). <https://doi.org/10.1242/dev.199580>



Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, 5, 113.

<https://doi.org/10.1186/1471-2105-5-113>

Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M., & Dubchak, I. (2004). VISTA: computational tools for comparative genomics. *Nucleic Acids Res*, 32(Web

Server issue), W273-279. <https://doi.org/10.1093/nar/gkh458>

Gai, Z., Li, Q., Ferron, H. G., Keating, J. N., Wang, J., Donoghue, P. C. J., & Zhu, M.

(2022). Galeaspid anatomy and the origin of vertebrate paired appendages.

*Nature*, 609(7929), 959-963. <https://doi.org/10.1038/s41586-022-04897-6>

Gegenbaur, C., Bell, F. J., & Lankester, E. R. (1878). *Elements of comparative anatomy*.

Macmillan and Co.

Hwang, W. Y., Fu, Y. F., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson,

R. T., Yeh, J. R. J., & Joung, J. K. (2013). Efficient genome editing in zebrafish

using a CRISPR-Cas system. *Nature Biotechnology*, 31(3), 227-229.

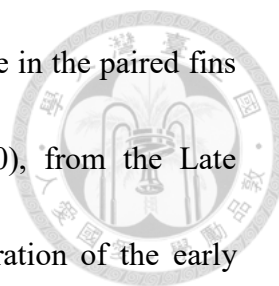
<https://doi.org/10.1038/nbt.2501>

Itou, J., Kawakami, H., Quach, T., Osterwalder, M., Evans, S. M., Zeller, R., &

Kawakami, Y. (2012). Islet1 regulates establishment of the posterior hindlimb

field upstream of the Hand2-Shh morphoregulatory gene network in mouse

embryos. *Development*, 139(9), 1620-1629. <https://doi.org/10.1242/dev.073056>



Janvier, P., Arsenault, M., & Desbiens, S. (2004). Calcified cartilage in the paired fins of the osteostracan *Escuminaspis laticeps* (Traquair 1880), from the Late Devonian of Miguasha (Quebec, Canada), with a consideration of the early evolution of the pectoral fin endoskeleton in vertebrates. *Journal of Vertebrate Paleontology*, 24(4), 773-779. [https://doi.org/10.1671/0272-4634\(2004\)024\[0773:Ccitpf\]2.0.Co;2](https://doi.org/10.1671/0272-4634(2004)024[0773:Ccitpf]2.0.Co;2)

Jurberg, A. D., Aires, R., Varela-Lasheras, I., Novoa, A., & Mallo, M. (2013). Switching axial progenitors from producing trunk to tail tissues in vertebrate embryos. *Development*, 142(5), 451-462. <https://doi.org/10.1016/j.devcel.2013.05.009>

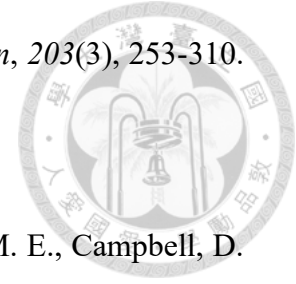
Kang, J., Nathan, E., Xu, S. M., Tzahor, E., & Black, B. L. (2009). *Isl1* is a direct transcriptional target of Forkhead transcription factors in second heart field-derived mesoderm. *Developmental Biology*, 334(2), 513-522. <https://doi.org/10.1016/j.ydbio.2009.06.041>

Kawakami, Y., Marti, M., Kawakami, H., Itou, J., Quach, T., Johnson, A., Sahara, S., O'Leary, D. D. M., Nakagawa, Y., Lewandoski, M., Pfaff, S., Evans, S. M., & Belmonte, J. C. I. (2011). *Islet1*-mediated activation of the beta-catenin pathway is necessary for hindlimb initiation in mice. *Development*, 138(20), 4465-4473. <https://doi.org/10.1242/dev.065359>

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., & Schilling, T. F. (1995).

Stages of embryonic development of the zebrafish. *Dev Dyn*, 203(3), 253-310.

<https://doi.org/10.1002/aja.1002030302>



Kwan, K. M., Fujimoto, E., Grabher, C., Mangum, B. D., Hardy, M. E., Campbell, D.

S., Parant, J. M., Yost, H. J., Kanki, J. P., & Chien, C. B. (2007). The Tol2kit: A multisite Gateway-based construction kit for Tol2 transposon transgenesis constructs. *Developmental Dynamics*, 236(11), 3088-3099.

<https://doi.org/10.1002/dvdy.21343>

Ludwig, M. Z., Bergman, C., Patel, N. H., & Kreitman, M. (2000). Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature*, 403(6769), 564-

567. <https://doi.org/10.1038/35000615>

Ludwig, M. Z., Patel, N. H., & Kreitman, M. (1998). Functional analysis of eve stripe 2 enhancer evolution in *Drosophila*: rules governing conservation and change.

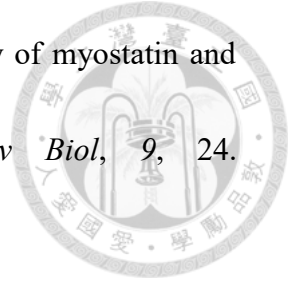
*Development*, 125(5), 949-958. <https://doi.org/10.1242/dev.125.5.949>

Lyons, E., & Freeling, M. (2008). How to usefully compare homologous plant genes and chromosomes as DNA sequences. *Plant Journal*, 53(4), 661-673.

<https://doi.org/10.1111/j.1365-313X.2007.03326.x>

Marcil, A., Dumontier, E., Chamberland, M., Camper, S. A., & Drouin, J. (2003). Pitx1 and Pitx2 are required for development of hindlimb buds. *Development*, 130(1),

45-55. <https://doi.org/10.1242/dev.00192>



McPherron, A. C., Huynh, T. V., & Lee, S. J. (2009). Redundancy of myostatin and growth/differentiation factor 11 function. *BMC Dev Biol*, 9, 24.

<https://doi.org/10.1186/1471-213X-9-24>

Minguillon, C., Gibson-Brown, J. J., & Logan, M. P. (2009). Tbx4/5 gene duplication and the origin of vertebrate paired appendages. *Proceedings of the National Academy of Sciences of the United States of America*, 106(51), 21726-21730.

<https://doi.org/10.1073/pnas.0910153106>

Moriyama, Y., Pratiwi, H. M., Ueda, S., & Tanaka, M. (2019). Localization of beta-Catenin and Islet in the Pelvic Fin Field in Zebrafish. *Zoological Science*, 36(5),

365-371. <https://doi.org/10.2108/zs180185>

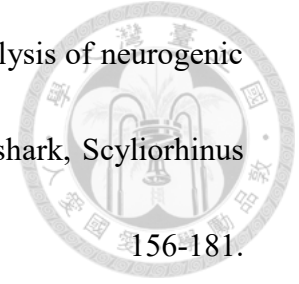
Mosimann, C., Puller, A. C., Lawson, K. L., Tschopp, P., Amsterdam, A., & Zon, L. I. (2013). Site-directed zebrafish transgenesis into single landing sites with the phiC31 integrase system. *Developmental Dynamics*, 242(8), 949-963.

<https://doi.org/10.1002/dvdy.23989>

Nakatani, Y., Shingate, P., Ravi, V., Pillai, N. E., Prasad, A., McLysaght, A., & Venkatesh, B. (2021). Reconstruction of proto-vertebrate, proto-cyclostome and proto-gnathostome genomes provides new insights into early vertebrate evolution. *Nat Commun*, 12(1), 4489. [https://doi.org/10.1038/s41467-021-](https://doi.org/10.1038/s41467-021-24573-z)

[24573-z](https://doi.org/10.1038/s41467-021-24573-z)





O'Neill, P., McCole, R. B., & Baker, C. V. (2007). A molecular analysis of neurogenic placode and cranial sensory ganglion development in the shark, *Scyliorhinus canicula*. *Dev Biol*, 304(1), 156-181.

<https://doi.org/10.1016/j.ydbio.2006.12.029>

Onimaru, K., Motone, F., Kiyatake, I., Nishida, K., & Kuraku, S. (2018). A staging table for the embryonic development of the brownbanded bamboo shark (*Chiloscyllium punctatum*). *Dev Dyn*, 247(5), 712-723.

<https://doi.org/10.1002/dvdy.24623>

Papadogiannis, V., Pennati, A., Parker, H. J., Rothbacher, U., Patthey, C., Bronner, M. E., & Shimeld, S. M. (2022). Hmx gene conservation identifies the origin of vertebrate cranial ganglia. *Nature*, 605(7911), 701-705.

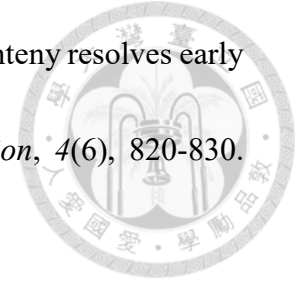
<https://doi.org/10.1038/s41586-022-04742-w>

Prummel, K. D., Nieuwenhuize, S., & Mosimann, C. (2020). The lateral plate mesoderm. *Development*, 147(12). <https://doi.org/10.1242/dev.175059>

Shubin, N., Tabin, C., & Carroll, S. (2009). Deep homology and the origins of evolutionary novelty. *Nature*, 457(7231), 818-823.

<https://doi.org/10.1038/nature07891>

Simakov, O., Marletaz, F., Yue, J. X., O'Connell, B., Jenkins, J., Brandt, A., Calef, R., Tung, C. H., Huang, T. K., Schmutz, J., Satoh, N., Yu, J. K., Putnam, N. H.,



Green, R. E., & Rokhsar, D. S. (2020). Deeply conserved synteny resolves early events in vertebrate evolution. *Nature Ecology & Evolution*, 4(6), 820-830.

<https://doi.org/10.1038/s41559-020-1156-z>

Stolfi, A., Ryan, K., Meinertzhagen, I. A., & Christiaen, L. (2015). Migratory neuronal progenitors arise from the neural plate borders in tunicates. *Nature*, 527(7578),

371-374. <https://doi.org/10.1038/nature15758>

Szeto, D. P., Rodriguez-Esteban, C., Ryan, A. K., O'Connell, S. M., Liu, F., Kioussi, C.,

Gleiberman, A. S., Izpisua-Belmonte, J. C., & Rosenfeld, M. G. (1999). Role of

the Bicoid-related homeodomain factor Pitx1 in specifying hindlimb

morphogenesis and pituitary development. *Genes Dev*, 13(4), 484-494.

<https://doi.org/10.1101/gad.13.4.484>

Tahara, N., Akiyama, R., Theisen, J. W. M., Kawakami, H., Wong, J., Garry, D. J., &

Kawakami, Y. (2018). Gata6 restricts Isl1 to the posterior of nascent hindlimb

buds through Isl1 cis-regulatory modules. *Dev Biol*, 434(1), 74-83.

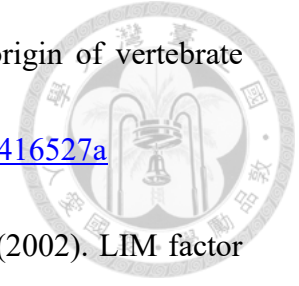
<https://doi.org/10.1016/j.ydbio.2017.11.013>

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary

Genetics Analysis Version 11. *Molecular Biology and Evolution*, 38(7), 3022-

3027. <https://doi.org/10.1093/molbev/msab120>

Tanaka, M., Munsterberg, A., Anderson, W. G., Prescott, A. R., Hazon, N., & Tickle, C.



(2002). Fin development in a cartilaginous fish and the origin of vertebrate limbs. *Nature*, 416(6880), 527-531. <https://doi.org/10.1038/416527a>

Thaler, J. P., Lee, S. K., Jurata, L. W., Gill, G. N., & Pfaff, S. L. (2002). LIM factor Lhx3 contributes to the specification of motor neuron and interneuron identity through cell-type-specific protein-protein interactions. *Cell*, 110(2), 237-249. [https://doi.org/Doi 10.1016/S0092-8674\(02\)00823-1](https://doi.org/Doi%2010.1016/S0092-8674(02)00823-1)

Thisse, C., & Thisse, B. (2008). High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nature Protocols*, 3(1), 59-69. <https://doi.org/10.1038/nprot.2007.514>

True, J. R., & Haag, E. S. (2001). Developmental system drift and flexibility in evolutionary trajectories. *Evolution & Development*, 3(2), 109-119. <https://doi.org/10.1046/j.1525-142x.2001.003002109.x>

Wong, E. S., Zheng, D., Tan, S. Z., Bower, N. L., Garside, V., Vanwalleghem, G., Gaiti, F., Scott, E., Hogan, B. M., Kikuchi, K., McGlenn, E., Francois, M., & Degnan, B. M. (2020). Deep conservation of the enhancer regulatory code in animals. *Science*, 370(6517). <https://doi.org/10.1126/science.aax8137>

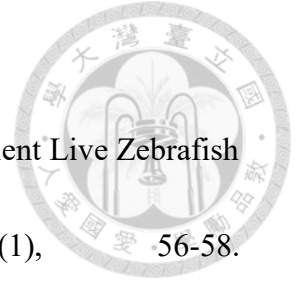
Yang, L., Cai, C. L., Lin, L., Qyang, Y., Chung, C., Monteiro, R. M., Mummery, C. L., Fishman, G. I., Cogen, A., & Evans, S. (2006). Isl1Cre reveals a common Bmp pathway in heart and limb development. *Development*, 133(8), 1575-1585.

<https://doi.org/10.1242/dev.02322>

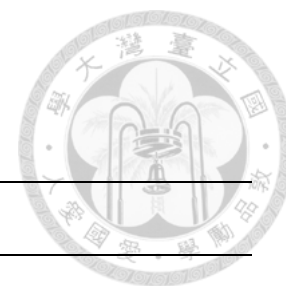
Zhang, X., Zhang, Z., Zhao, Q., & Lou, X. (2020). Rapid and Efficient Live Zebrafish

Embryo Genotyping. *Zebrafish*, 17(1), 56-58.

<https://doi.org/10.1089/zeb.2019.1796>



## Tables



**Table 1. List of primers used**

Name	Sequence (5' – 3')
F-isl1 ish	CAGTCCAGAGAGACACGACG
R-isl1 ish	ACGTTGCGCACAGAAGTAGA
F-isl1b ish	AGCAAGTGTGCAAAAATGCGA
R-isl1b ish	TGATGTCCGTTGGACTTGCT
F-isl2a ish	CAGTTCGACAGCCRCCTCAT
R-isl2a ish	AAACCCCAAAGTCGCTTGGA
F-isl2b ish	TTGCAGGCAGCCCAATAAGA
R-isl2b ish	AGCCTAAAACAAATGCGCAAA
F-Cpisl1 ish	ACGAGAAGCAGCTCCATACG
R-Cpisl1 ish	TTTGGCAAAGCTGACCCACT
F-Cpisl2 ish	CGGACCTGCTATGCAGCTAA
R-Cpisl2 ish	CCCCAGGCGGTAATCATCTC
F-Cpisl11 ish	GGTTCACATGTCAGGGGTCA
R-Cpisl11 ish	TGGCATGGGGGTCTATCCT
F-Cpgdf11 ish	ACCCACTGTCCTCTCCTCTC
R-Cpgdf11 ish	AGATTTGCGCCGGGTACGTTT
F-Cpmstn ish	CCATATCCCAGCCCAACAGG
R-Cpmstn ish	ACCGTTGATGTCAGACGCAT
F-Lcis11 ish	GATCCGCGACCCCTTCATC
R-Lcis11 ish	CTGCTTCTCGTTCAGCACCG
F-Lcis12 ish	CCTTCTGCAAGCGCGACTA
R-Lcis12 ish	GGATGCTGCGTTTCTTGTC

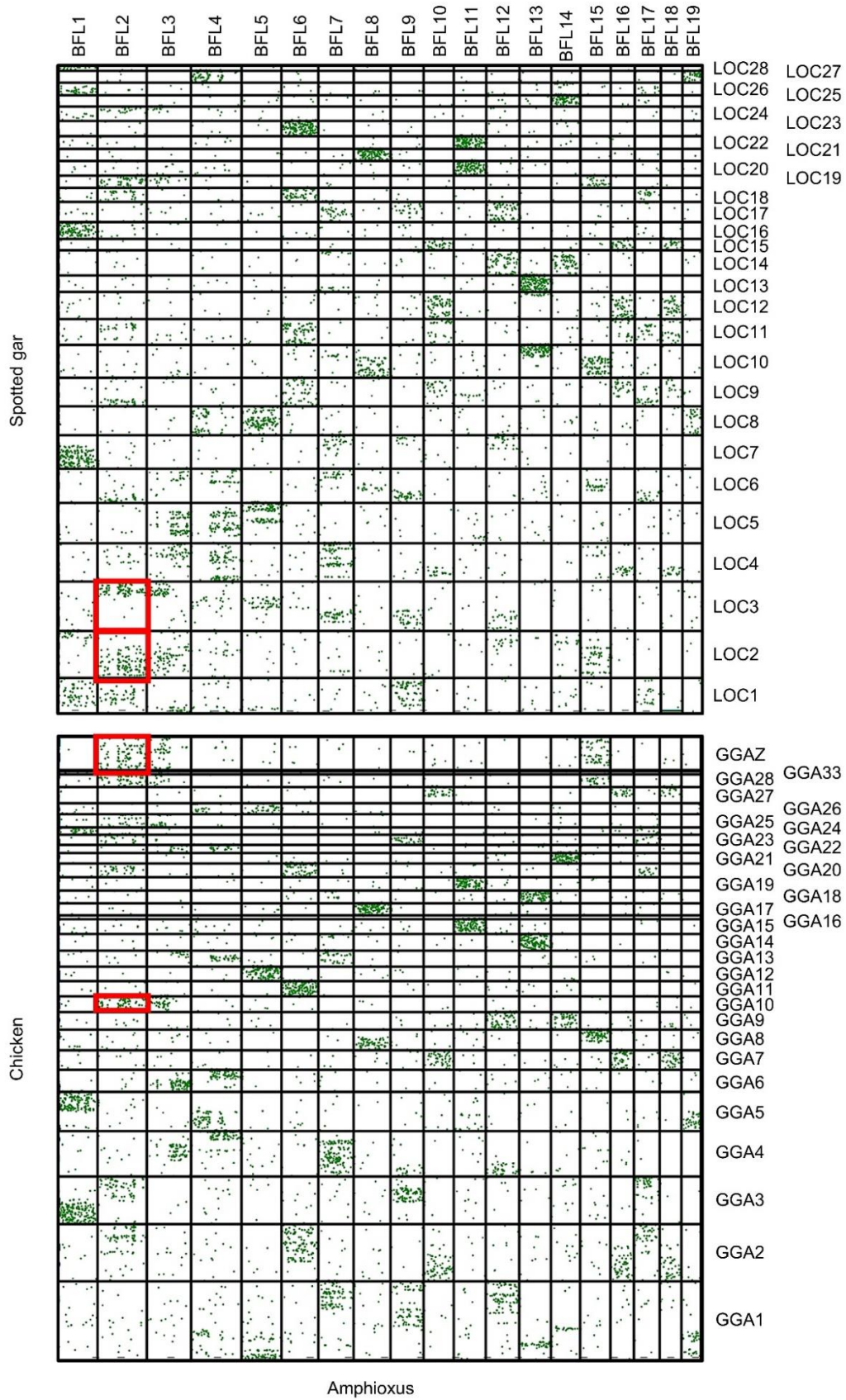
F-Lcis111 ish	CTGCTCGCTGGGTCTGAG
R-Lcis111 ish	CTCCTTCAGGAGGGCGTCG
F-and1 ish	GAGGGGTTTGGCTCACCTTT
R-and1 ish	ATTGACCAGAACGTGGGTCC
F-shha ish	AGGATGAGGAGAACACGGGA
R-shha ish	TGGACCCCTCCTGTTGTAA
F-pitx1 ish	CCGCTGGAGAGGCTTTTCTAA
R-pitx1 his	AGATTCAGCCGTAGCGTGG
F-tbx4 ish	GGTTCCAAGAACACGGCCTA
R-tbx4 ish	GGGCTGTGAAAGGCTGGTAT
F-gdf11 ish	GAATTCCTGATATTTAACAACATGAAAAGGTA
R-gdf11 ish	AGTGGATAGCGACAACAGCG
F-mstna ish	TCCTTTAGCACGCCTTGAA
R-mstna ish	TCCCTCCGGATTCGTTTTGG
F-Hs Isl1 enh-pT2	AACACAGGCCAGATGTACGTCGCAACAGGGCAAA
R-Hs Isl1 enh-pT2	ATCGATAACGTCGAGTCACAGAATAGAACCTGCACAA G
F-Dr isl1 enh-pT2	AACACAGGCCAGATGTAAACAGCACAGCAGGCCAA
R-Dr isl1 enh-pT2	ATCGATAACGTCGAGTGGAATGACTTAAGGGAGCGTT
F-Lo isl1 enh-pT2	AACACAGGCCAGATGATTCAGGTTACCGGCCTTGAG
R-Lo isl1 enh-pT2	ATCGATAACGTCGAGTGCCTAGCTCCAACATGCTTT
F-Dr isl2a enh-pT2	AACACAGGCCAGATGCTAAAGATGCGTGACACACGG
R-Dr isl2a enh-pT2	ATCGATAACGTCGAGTTGGACAAAAATGAGCCCCCT
F-isl1-pT2	GCTACCGGTCGCCACGGGCCTTCTGTCCGGTTTTA
R-isl1-2A	AGAGAAGTTCGTGGCGGCCTCTATAGGACTCGCTACC



F-2A-nlsEGFP	GCCACGAACTTCTCTCTG
R-2A-nlsEGFP-pT2	CTAGATTAGTTACCCCTTGTACAGCTCGTCCATG ACAGGGATCGGCGGCCGCTTTACTGCGCTCGGCATCA
F-CR2 mut	CCTGC GCAGGTGATGCCGAGCGCAGTAAAGCGGCCGCGAT
R-CR2 mut	CCCTGT AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATA
Constant oligo	ACGGACTAGCCTTATTTAACTTGCTATTTCTAGCTCT AAAAC
gRNA-1	ggTATCTAAATATCGCGCCGACCGG
gRNA-2	GGTTGATCACTACTACGGGAGGG
gRNA-3	ggAAAGGGAGTCCCGTTAAGAGAGG
gRNA-4	GGGGCAGGAGTTTTAGCCGCTGG
gRNA-5	GGCGGCCGCTTATCTGCGCTCGG



# Figures





**Figure 1.** *Conserved synteny between amphioxus and vertebrates.*



Oxford plot showing genome correspondence between amphioxus (*Branchiostoma floridae*, BFL) and two vertebrate species: spotted gar (*Lepisosteus oculatus*, LOC)

and chicken (*Gallus gallus*, GGA). The X and Y-axis represents chromosomal

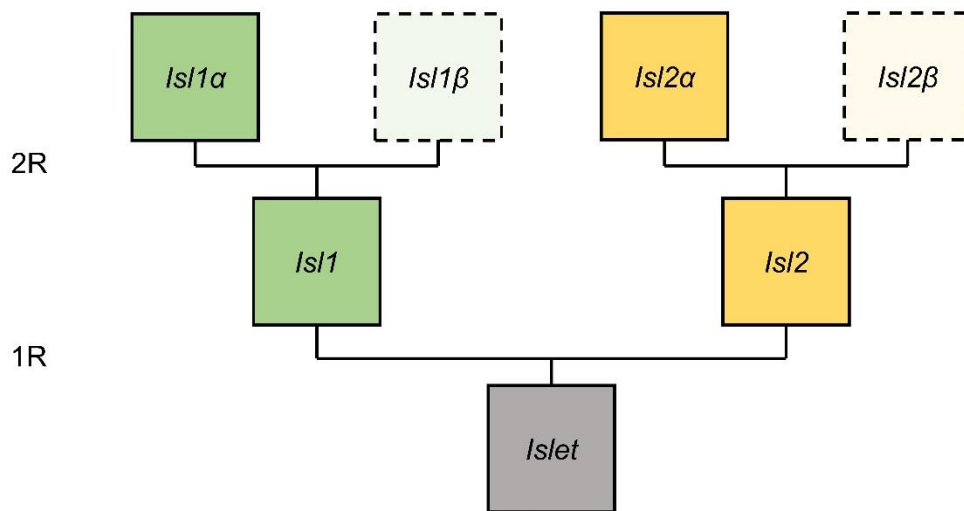
location. The number after each species abbreviation stands for the chromosome

number. Each dot represents an ortholog that is mapped in the species comparison.

The chromosomes of *Islet* orthologs are outlined in red. The two *Islet* orthologs in

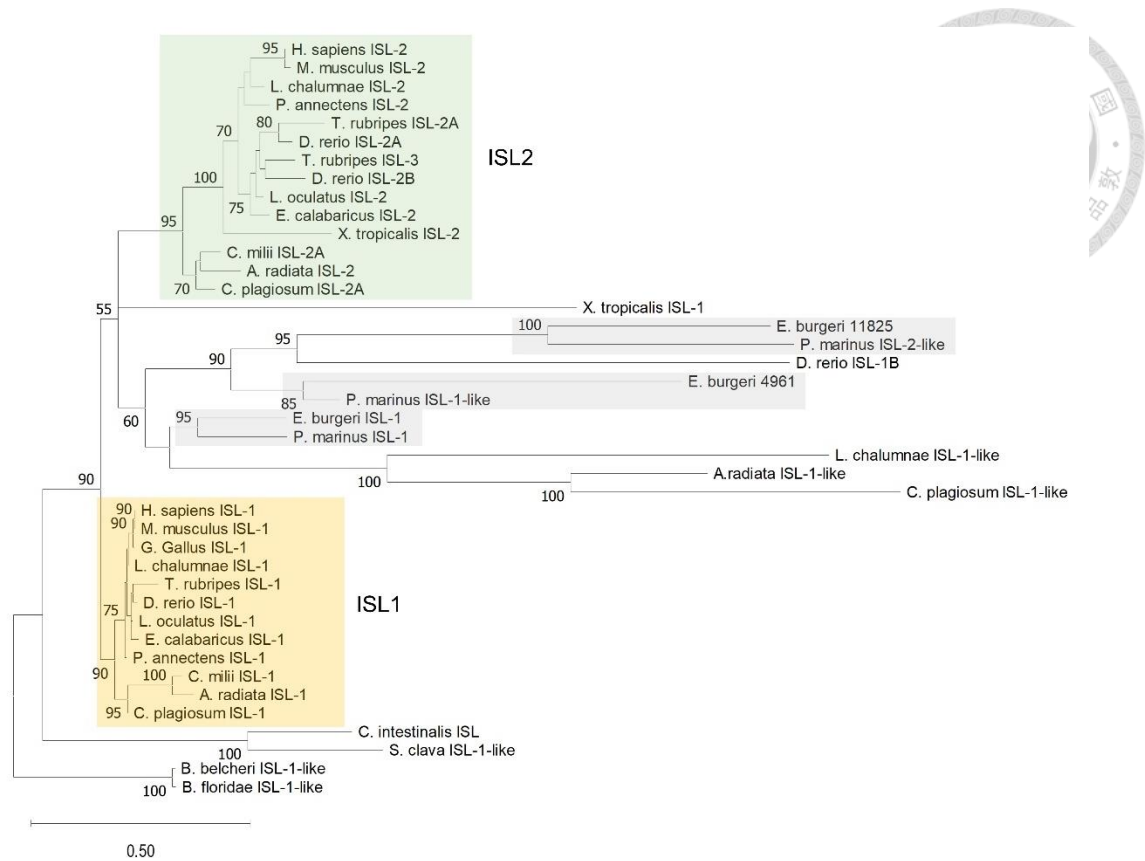
both vertebrate species are located on the conserved amphioxus–vertebrate syntenic

blocks, where the dense dots form a block.



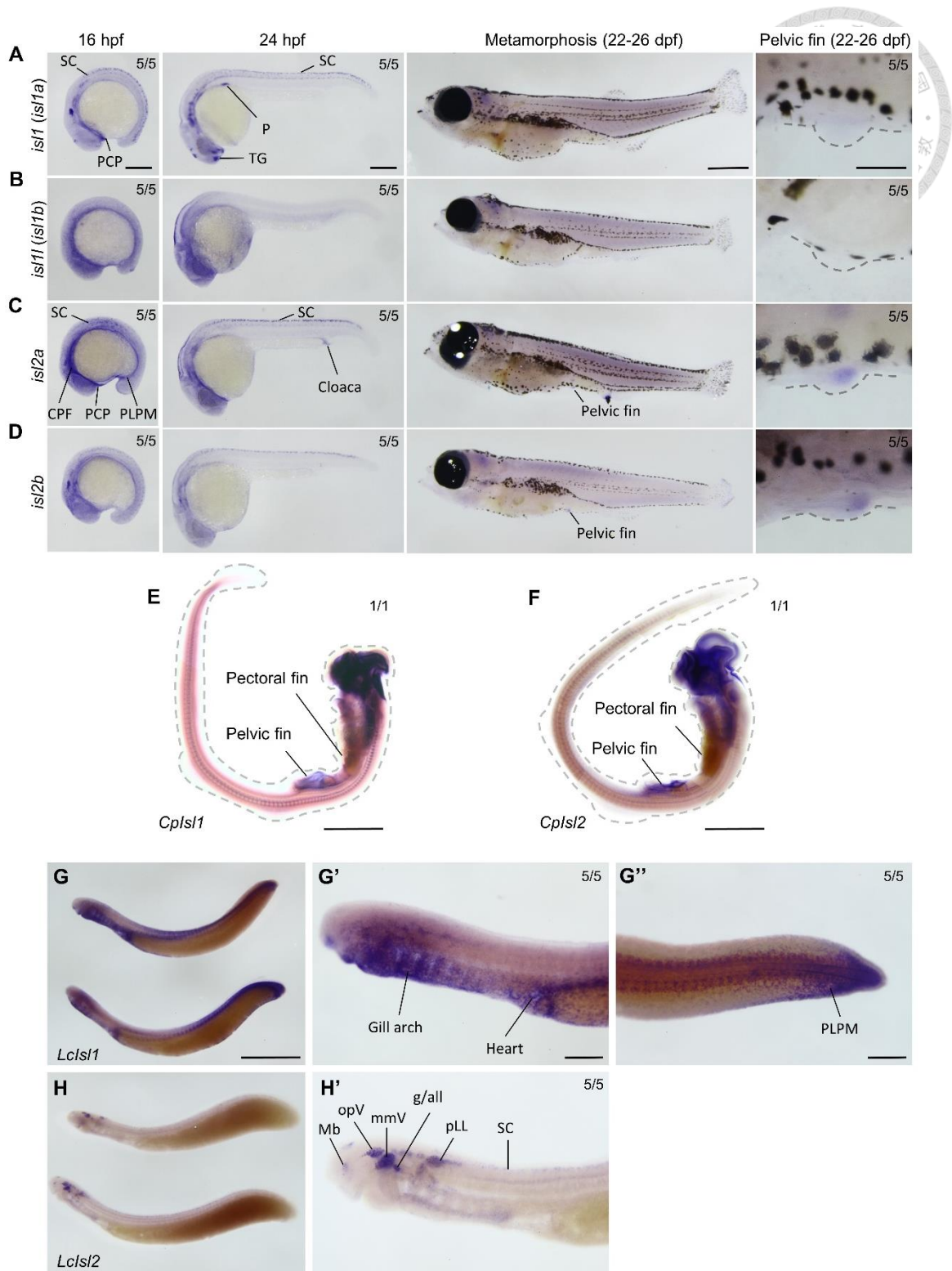
**Figure 2.** *Inferred Islet gene duplication tree.*

According to the organization of bony vertebrate chromosomes by Simakov et al. (2020), two duplicates in the spotted and chick were lost after the second genome duplication.



**Figure 3.** *Phylogenetic relationship of ISLET orthologs.*

The phylogenetic trees for ISLET orthologs were generated. ISL1 and ISL2 are clustered individually. The three ISL paralogs in cyclostomes are clustered together, along with Isl1-like proteins in several species, potentially due to independent duplication.



**Figure 4.** Whole-mount *in situ* hybridization (WISH) staining of Islet orthologs in vertebrates.

(A – D) Staining of *DrIsl1a*, *DrIsl1b*, *DrIsl2a*, and *DrIsl2b* on zebrafish embryos and

larvae at the designated stages (hpf: hours post-fertilization, dpf: days post fertilization).

Dash lines outline the pelvic fins. **(E – F)** Stage 27-28 bamboo shark embryos. **(G – H)**

Stage 26 lamprey embryos. I performed WISH on only one bamboo shark for each

probe. g/all: geniculate/anterior lateral line ganglionic complex, opV: ophthalmic

trigeminal, Mb: midbrain, mmV: maxillomandibular trigeminal, p: pancreas, PCP:

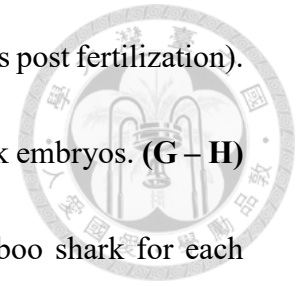
prechordal plate, PLPM: posterior lateral plate mesoderm, pLL: posterior lateral line

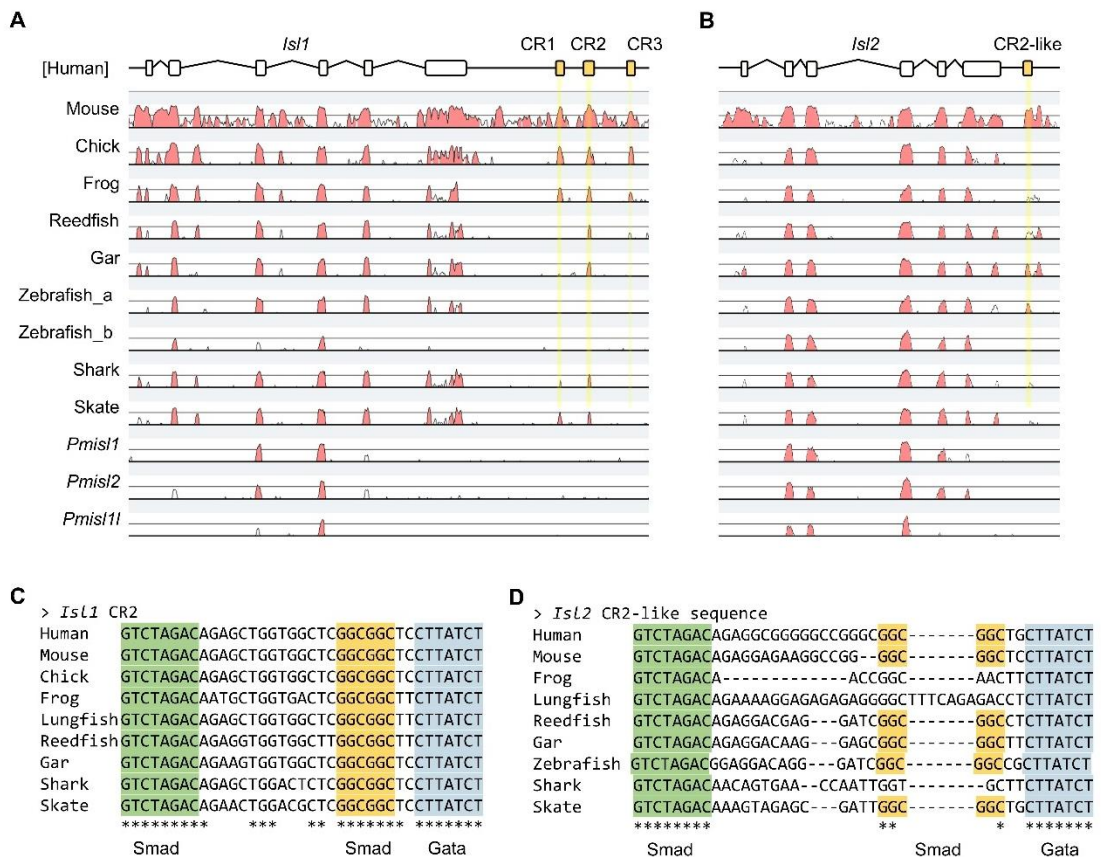
ganglion, SC: spinal cord, TG: trigeminal ganglia. Scales are the same in the same

column, and the scale bar is only shown on the top panel of each column. Scale bar size:

1 mm for whole metamorphosing zebrafish larvae and whole lamprey larvae, 5mm

for bamboo shark, and 100  $\mu\text{m}$  for all other panels.

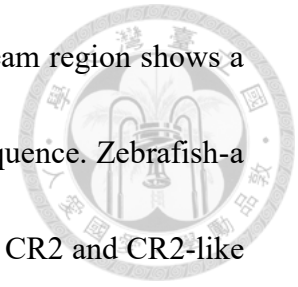


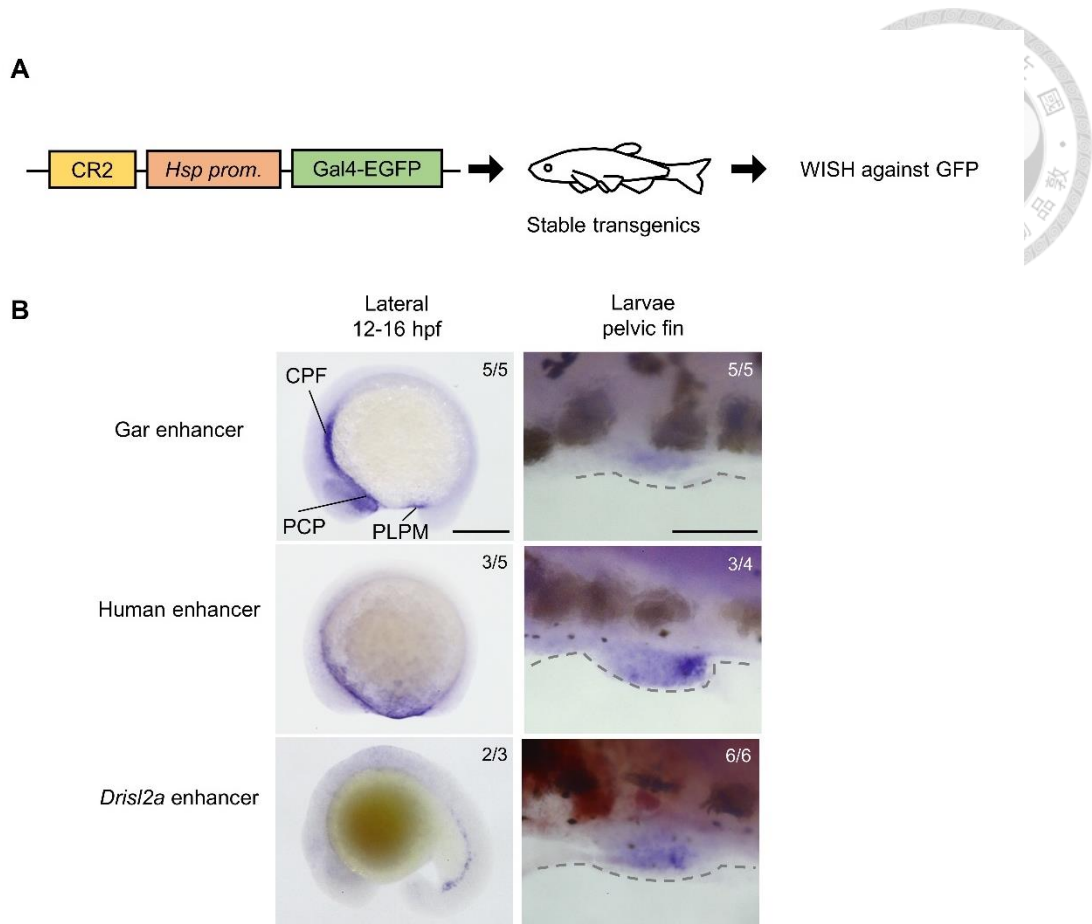


**Figure 5.** Sequence comparison identifies conserved putative enhancers of *Islet* genes.

(A) *Is11* and its downstream region from selected animals were aligned with the mVista SLAGAN program (Frazer et al., 2004) using humans as a reference. Sequences > 50% identity is marked in red, and the conserved regions (CR1-3) are highlighted in yellow. Zebrafish\_a and zebrafish\_b represent the a and b paralogs of *Is11/2*, respectively. As orthology mapping of the three *Islet* genes in the sea lamprey (*Pmis11*, *Pmis12*, and *Pmis11*) is not feasible, all three genes were included in the analysis to ensure comprehensive coverage. The peak for CR1 is not seen in the selected actinopterygians (reedfish, gar, and zebrafish), and CR2 is absent in zebrafish. Conservation of CR1-3 is

not seen in lamprey. **(B)** mVista alignment of *Isl2* and its downstream region shows a sequence conserved in gnathostomes, which I call the CR2-like sequence. Zebrafish-a and zebrafish-b represent *Drisl2a* and *Drisl2b*, respectively. **(C, D)** CR2 and CR2-like sequences share a high degree of sequence identity, harboring binding sites for Smad (highlighted in green and yellow) and Gata (highlighted in blue) transcription factors.

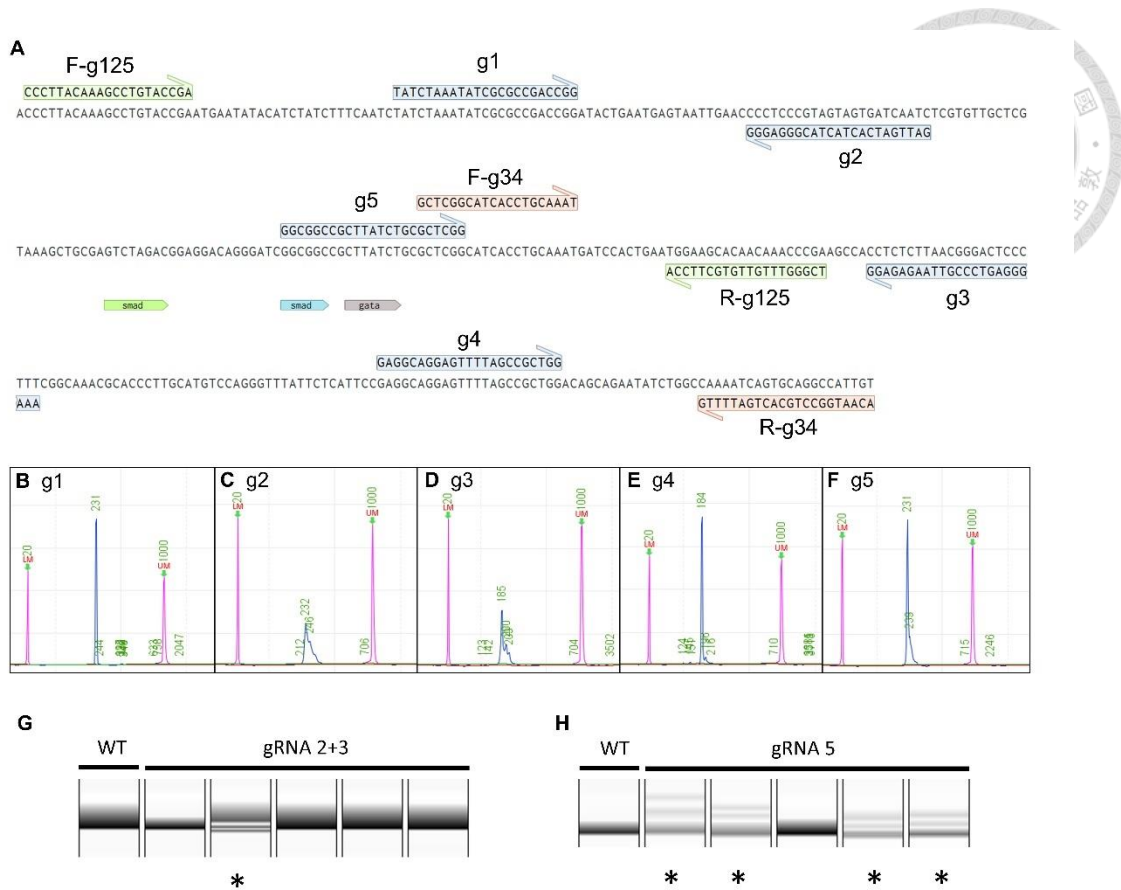




**Figure 6.** *CR2* drives reporter expression in the pelvic fin.

(A) Structure of reporter construct used to generate transgenic animals. (B) F1 transgenic embryos were stained with Whole-mount in situ hybridization (WISH) against EGFP at designated stages. Enhancer regions from gar, human, and *Drisl2a* drove expression in the cardiopharyngeal field (CPF), prechordal plate (PCP), and posterior lateral plate mesoderm (PLPM) and their derivatives. GFP signal is also expressed in the pelvic fin of metamorphosing larvae. Scale bars represent 100  $\mu\text{m}$ .

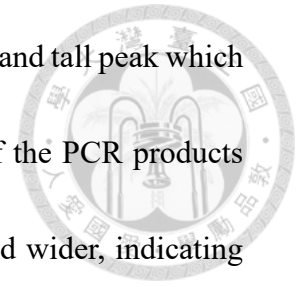


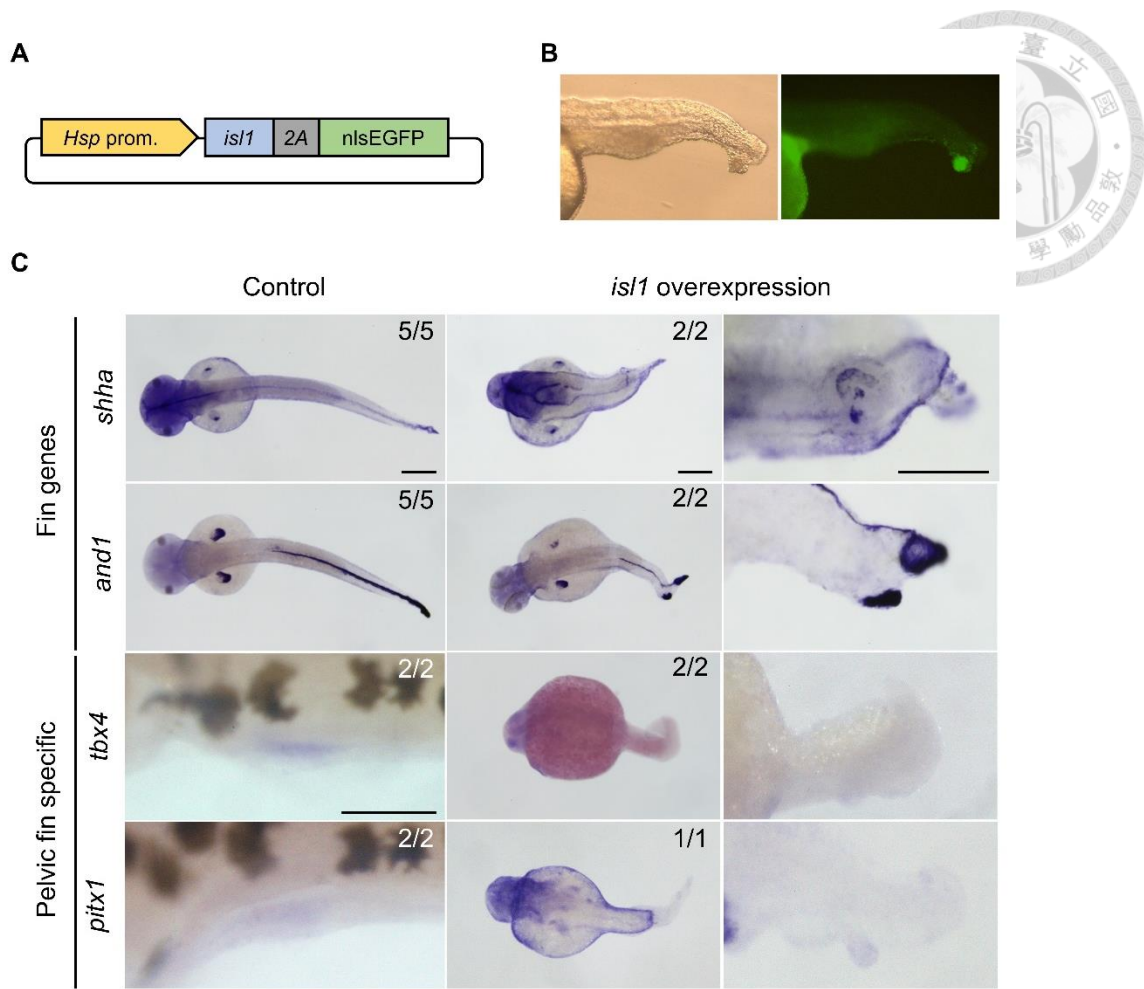


**Figure 7.** CRISPR/Cas9 design to knockout CR2.

(A) Five gRNAs (g1 to g5) were designed targeting different positions of *is12* CR2 region. Embryos injected with the designated gRNA/Cas9 mixture were lysed at 24 hours post-fertilization to collect respective genomic DNAs and subjected to PCR and subsequent genotyping using capillary electrophoresis. Two pairs of PCR primers were used for PCR: F-g125/R-g125 and F-g34/R-g34. F-g34 and R-g34. The gnathostome-conserved transcription factor binding sites for Smad and Gata are annotated in tags. (B – D) The signal distribution of PCR products from designated samples is shown. The size is annotated on the top of each peak. The 20 bp and 1000 bp peaks in pink represent the low and high alignment marker peaks, respectively. The sample peaks are in blue.

PCR products from embryos injected with g1 and g4 exhibit a sharp and tall peak which is not different from uninjected siblings. The signal distribution of the PCR products from embryos injected with g2, g3, and g5 was greatly reduced and wider, indicating the presence of PCR fragments with different sizes. (G-H) Live embryo genotyping of F1 fish by capillary electrophoresis revealed multiple bands transforming from fluorescent signals in several individuals (asterisk). This indicates the presence of indels.





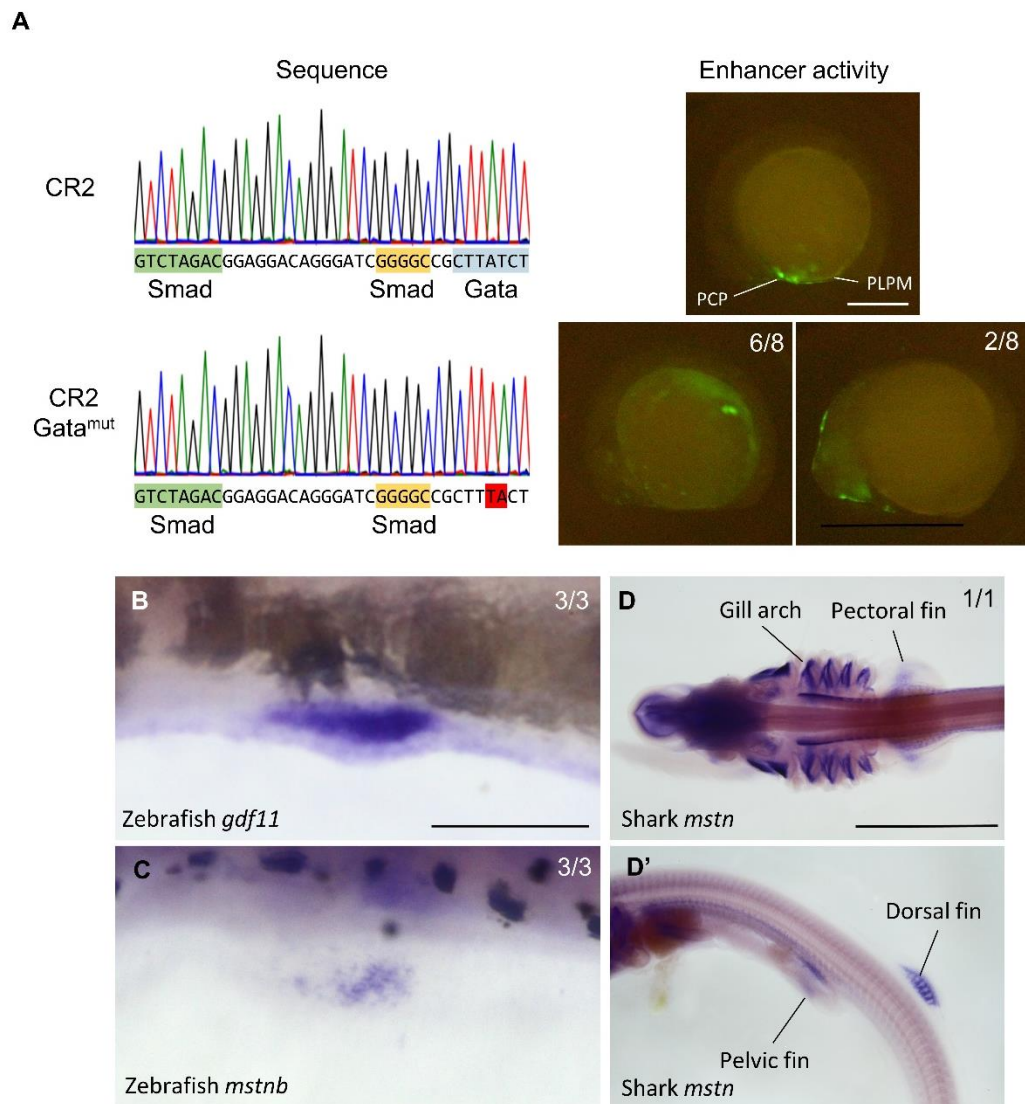
**Figure 8.** *Isl1* overexpression induces the outgrowth of a fin with an unspecified identity.

(A) Schematic of the heat shock-inducible construct. (B) Injected embryos heat-shocked at different stages showed phenotypes ranging from normal, deformed to protrusion outgrowth. Three repeats were performed. (C) Representative embryos exhibit various degrees of phenotype, including trunk bending and protrusion outgrowth (arrow). (D) Representative embryos are exhibiting fluorescent signal associated with the ectopic protrusion. Left: bright field; Right: dark field in the green channel. (E) Whole-mount *in situ* hybridization against designated genes on injected

embryos showed fin-associated markers expressed in the ectopic protrusion. Control embryos express *shha* in the pectoral fins and the median fin; In *isll* overexpressing embryos, *shha* is expressed in the protrusion, albeit without a particular polarization.

Control embryos express *andl* in the fin fold; In *isll* overexpressing embryos, *andl* is expressed at the distal part of the protrusion and is unconnected from the tail fin.

Wildtype larvae express *tbx4* and *pitx1* in the pelvic fin during metamorphosis; *Isll* overexpression embryos did not express *tbx4* or *pitx1* in the ectopic protrusion. Scale bars represent 100  $\mu\text{m}$ .



**Figure 9.** Probing the transcription factors of CR2.

(A) Sequencing results of the wildtype CR2 construct and Gata-mutated construct (CR2 Gata<sup>mut</sup>). The mutated site is highlighted in red. (B-C) Expression of *gdf11* and *mstnb* in zebrafish pelvic fin. (D) Expression of *mstn* in bamboo shark embryos. In both zebrafish and bamboo shark, *gdf11* or *mstn* expression is identified in the pelvic fin. Scale bars represent 100 μm for zebrafish embryos and larvae pelvic fin, and 5mm for bamboo shark.