Molecular basis of sexual dimorphism in paradise fish (*Macropodus opercularis*)

36. OTDK Biológia Szekció

Kata Szabó

ELTE Eötvös Loránd University Faculty of Science

Biology Master's Program, 2nd year

Supervisor:

Dr. Máté Varga, Associate Professor

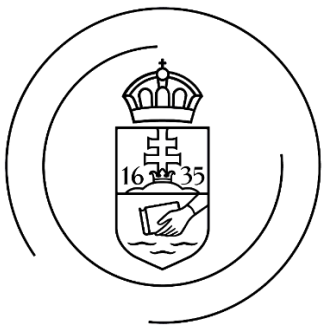

TABLE OF CONTENTS

Table of Contents 2
Abbreviations 3
Summary 4
Introduction 5
Aims 9
Materials and Methods 10
Fish maintenance 10
Multiple sequence alignment 10
Normalized body length measurement 10
Fin ray measurement 11
Total RNA isolation 11
cDNA synthesis 11
Primer design for qPCR 12
Making agarose gels 13
Cloning of the kcnb2l enhancer region 13
Transient transgenesis 14
Statistics and visualization 14
Multiple sequence alignment 15
Normalized tail length measurement 16
Fin ray measurements 16
Comparison of kcnb2l expression levels between paradise fish males and females 18
Cloning of the kcnb2l enhancer region and creation of zebrafish reporter lines 19
Discussion 22
Acknowledgements 23
References 24
Supplementary Information 27

ABBREVIATIONS

alf: another longfin

cDNA: complementary Deoxyribonucleic Acid

DNA: Deoxyribonucleic Acid

dnajb6l: dnaJ homolog subfamily B member 6-like

eGFP: Enhanced Green Fluorescent Protein

iTOL: Interactive Tree of Life

kb: Kilobase

Kcc4a: K^+-Cl^- Cotransporter Subfamily 4a

Kcnb2: Potassium Voltage-Gated Channel Subfamily B Member 2

Kcnb2l: Potassium Voltage-Gated Channel Subfamily B Member 2-like

Kcng2: Potassium Voltage-Gated Channel Modifier Subfamily G Member 2

Kcnk5b: Potassium Channel Subfamily K Member 5b

Kcnh8: Potassium Voltage-Gated Channel Subfamily H Member 8

Kcnh2a: Potassium Voltage-Gated Channel Subfamily H (ether-a-go-go-related) Member 2a

Kcnj13: Potassium Inwardly Rectifying Channel Subfamily J Member 13

Kcnj15: Potassium Inwardly Rectifying Channel Subfamily J Member 15

Kcns1: Potassium Voltage-Gated Channel Modifier Subfamily S Member 1

L: liter

log: longfin

mg: milligram

mRNA: messenger Ribonucleic Acid

MSA: Multiple Sequence Alignment

qPCR: Quantitative Polymerase Chain Reaction

RNA: Ribonucleic Acid

rpz: rapunzel

SB: Sodium Borate

TAE: Tris Base, Acetic Acid, EDTA

SUMMARY

During the last decades of the 20th century the paradise fish (*Macropodus opercularis*) became a popular model organism for behavioral studies. While recently it has faded from the limelight, the advent of novel genetic techniques and their application to this particular model species can help them become a relevant model again

Paradise fish display sexual dimorphism in the length of their tail, the appendage being longer in males than in females. As longer tails do not confer significant advantage to the males, this seems to be a classic case of sexual selection. Identifying the molecular mechanisms behind this phenomenon, therefore, will also help us to gain insights into the genetic basis and evolution of sexual dimorphic characters.

Length and patterning of fish fins has been recently linked to potassium channels and the bioelectricity regulated by them. To reveal the molecular basis of paradise fish tail dimorphism we combined phenotyping with bioinformatics and molecular biology experiments. First we provide a detailed quantification of the tail-differences between the two sexes. Using transcriptomic data, we also show that one of the most likely candidates for driving the difference tail length between males and females is *kcnb2l.*

We designed further experiments to validate our finding by quantitative PCR (qPCR). Lastly to study the enhancer region of this gene – and identify potential male-specific enhancers – we started the generation of transgenic zebrafish reporters.

INTRODUCTION

Ever since Darwin the origin of sexually dimorphic traits has been a main interest of evolutionary biologists. Such traits can have varied functions within the animal kingdom. In the case of the now extinct New Zealand huia males had short and stout bills, while females long and sickle-shaped ones, which gave access to different food sources, effectively creating isotopic niches for the two sexes (Figure 1A). Peacock and deer bucks are classic examples of sexual selection, as their tail feathers and antlers, respectively, are essential accessories during courtship behavior (Figure 1B, 1C) (1). Buck antlers, just like the crests that the common newt males develop (Figure 1D) are present only during mating season. Oftentimes these exaggerated traits reduce the fitness of the individual carrying them, which led to the discovery of the importance of the female sexual selection (1,2).

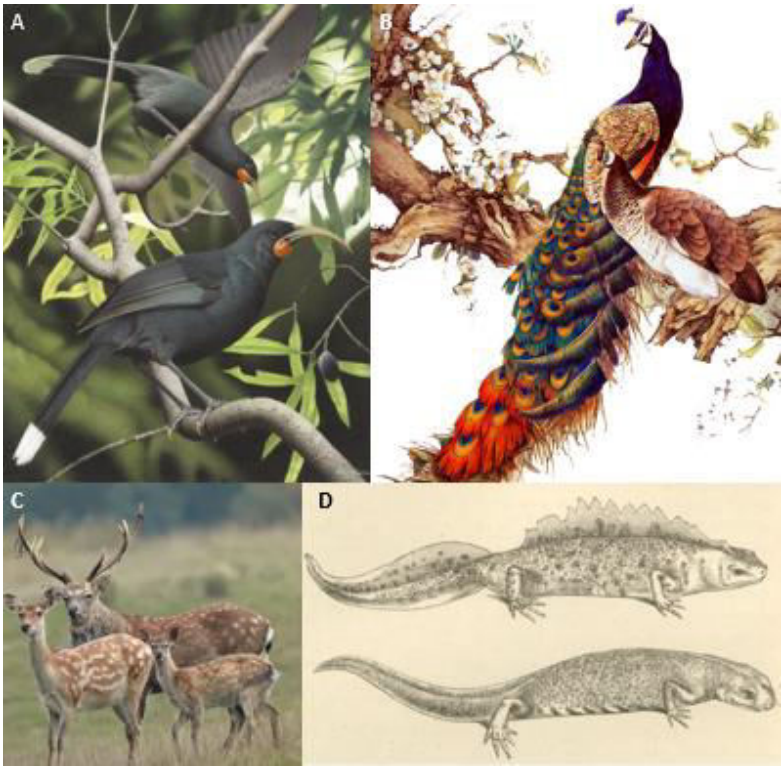

*Figure 1. Sexually dimorphic traits in tetrapods. (A) Female (bottom) and male (top) New Zealand huia (*Heteralocha acutirostris*) (i) (B) Male (left) and female (right) peacocks (*Pavo *sp.) (ii) (C) A family of Dybowski's sika deer (*Cervus nippon dybowskii*) (iii) (D) Common newt (*Triton cristatus*), male during mating season (top), female (bottom)* (1)

Several examples of sexual dimorphism can be found in fish as well (Figure 2). These range from differences in coloration (Figure 2C, 2F) to differences in body size (Figure 2E), but perhaps the most notable are the ornamental appendages that males of some species carry. For example, modified tails can be found in several species in the *Xiphophorus* (Greek for dagger bearer) genus (Figure 2D), where the size, shape and color of the lower margin of caudal fin of the males differs from the females. The tail is elongated resembling a sword (hence the name) and can reach the size of the fish. Swordtails affect the mobility of the fish, becoming a hindrance when facing a predator (1,3).

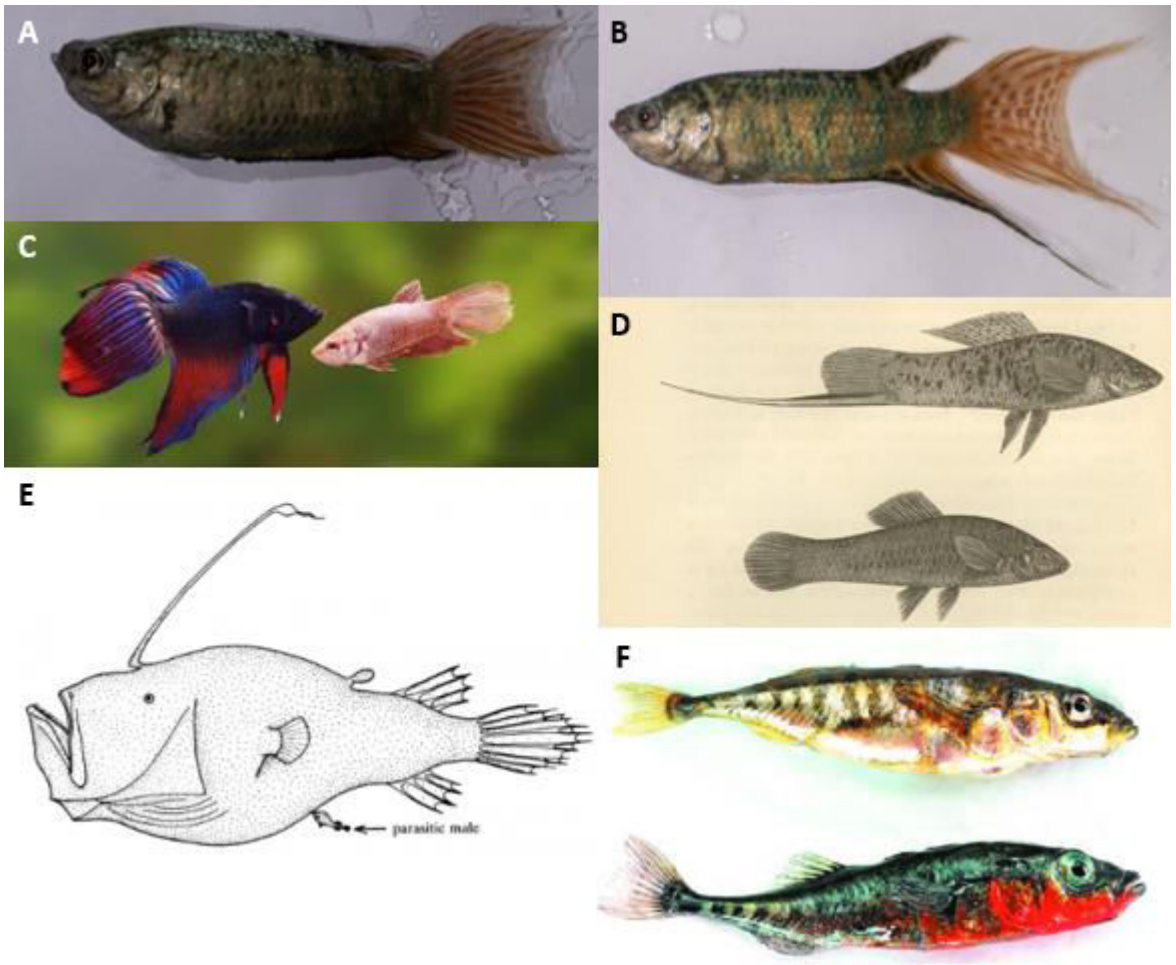

Figure 2: Sexually dimorphic traits in fish.
A,B) Female (A) and male (B) paradise fish (photo by Dávid Czimer) C) Male (left) and female (right) Betta splendens *(iv) D) Male (top) and female (bottom)* Xiphophorus hellerii (1) *E) Female and parasitic male angler fish (*Lophius piscatorius*) F) Three-spined stickleback (*Gasterosteus aculeatus*) female (top), male in breeding dress (bottom)* (4)

The development of different parts of the body is achieved in a complex highly regulated manner (5). This growth can occur in an allometric way for different organs during the development (i.e. they can develop at a different speed at different times) (6). While classical developmental models concentrate on morphogen fields and morphogenetic movements, lately bioelectrical signals have also been connected to patterning, limb development and wound healing (7). Such signals can also be responsible for the previously mentioned allometric growth as well (7).

Bioelectric signaling plays a role in the development of an organism trough electric currents, fields and action potentials. One of the key player in the aforementioned process, in fish, is the two-pore domain potassium-channel family (6). For example, the development of the *Xiphoporus'* sword can be linked to one of these potassium-channels, Kcnh8 (Potassium Voltage-Gated Channel Subfamily H Member 8) (2,3,8). Mutations in potassium channels also lead to different tail phenotypes in other fishes as well, for example the long-tail phenotype in goldfish (*Carassius auratus*) (9), or the long-fin (8) and elephant-ear (10) in Siamese Fighting fish (*Betta splendens*) have all been linked to such ion channels.

The connection between the bioelectricity and the development of the fins was also studied in zebrafish (*Danio rerio*) (6), as the causative mutations for several fin-mutant phenotypes, like *longfin* (*log*), *another longfin* (*alf*) and *rapunzel* (*rpz*) were mapped to genes encoding potassium-channels (6,11,12). In these mutants the long fin phenotype often arises as the membrane potential cannot be maintained due to gain-of-function mutations in the respective potassium-channels (6,11).

The main focus of our laboratory is related to the molecular genetics of zebrafish. These animals are very popular model organisms and have many beneficial properties, but – like all the currently popular model species – they have their own limitations. In the last few years, we also began to use paradise fish (*Macropodus opercularis*) in our work. These fish have been traditionally studied at the Ethological Department of the ELTE Eötvös Loránd University Biological Institute as they have highly complex social behavior.

In paradise fish a pronounced sexual dimorphism of the caudal fin can also be observed: males have longer tail fin compared to the female fish (Figure 2A,B) (13). To uncover the molecular basis of this dimorphism, the transcriptomes of male and female fish have been sequenced and compared. The results suggested that one of the best candidate genes, which shows differential

expression between the tails of the two sexes was *kcnb2l (potassium voltage-gated channel subfamily B member 2-like),* a potassium-channel coding gene (Figure 3). Given the well documented involvement of other potassium-channels in fin size in other fish species, our group decided to further study the potential role of this gene in the development of paradise fish sexual dimorphisms.

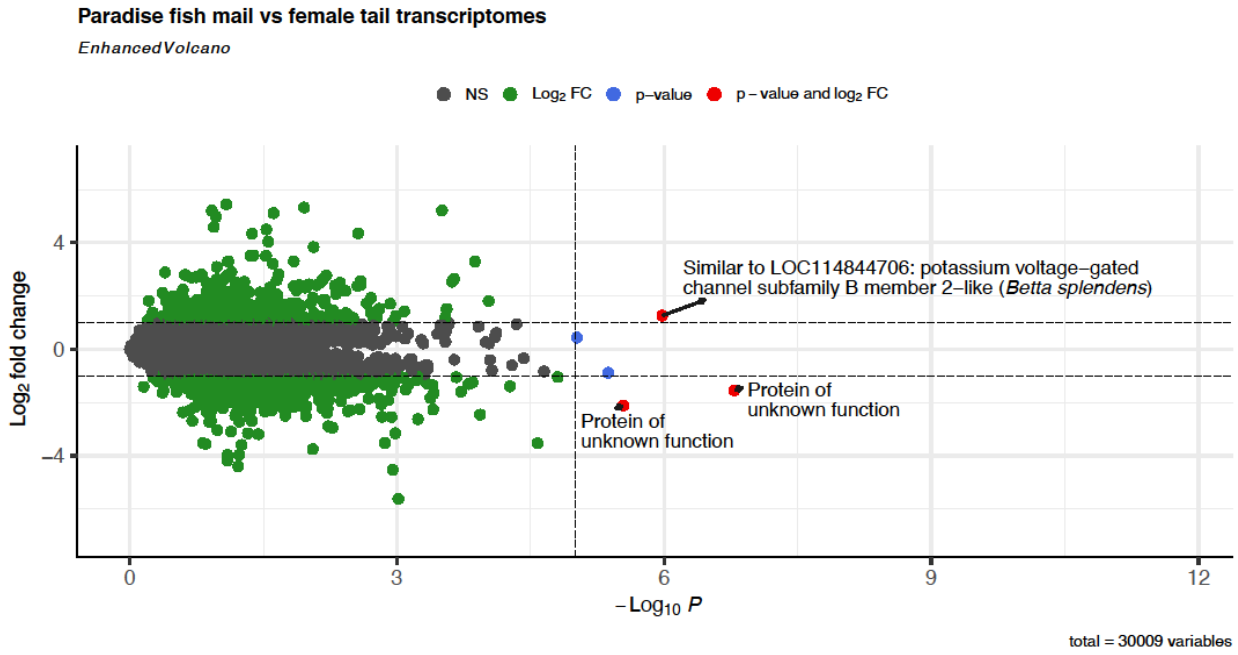

*Figure 3. Volcano plot of the transcriptome analysis of male and female paradise fish tail transcriptomes. A gene encoding a potassium-channel (*kcnb2l*) shows up as one of the differentially expressed genes that is upregulated in males. (Figure from Máté Varga)*

AIMS

The main goal of my research is to quantitatively characterize the differences between the caudal fins of male and female paradise fish and to study the potential role of *kcnb2l* in the emergence of this sexual dimorphism in paradise fish.

Firstly, I performed an *in silico* analyses to compare different potassium channels that have been linked to tail phenotypes in different fish species.

I also wanted to quantitatively validate our observations about the tails of the male paradise fish being longer than females. For this we took photos of the fish, measured their tails and calculated their relative length, normalized to their body length.

To confirm our hypothesis of *kcnb2l* being the responsible for the longer tails in the males we wanted to perform quantitative PCRs to see if there is a difference there. For an inner control β-*actin* is used as it is a protein expressed in all cells, with the same amount in cells of the same origin.

Lastly, to study potential male-specific enhancer sequences upstream of *knc2bl*, the putative enhancer region of our target gene was cloned in the front of a green fluorescent reporter gene (*eGFP*). Using a two-component transgenesis system, based on the Tol2 transposon, I injected the reporter construct into zebrafish embryos. We plan to raise these transgenic fish and follow the expression pattern of the reporter gene as they develop. If males will show transgene expression in their fins, which would confirm that this putative enhancer region indeed contains a sequence that is activated in males (likely due to the presence of male-specific hormones, such as testosterone).

MATERIALS AND METHODS

Fish maintenance

Paradise fish used in our experiments were all maintained in the Animal Facility of ELTE Eötvös Loránd University, under standard conditions (Permit Number: PE/EA/406—7/2020) (13).

Multiple sequence alignment

Amino acid sequences of potassium channels from diverse species were aligned using the Clustal Omega program, with standard parameters. The fish studied were turquoise killifish (*Nothobranchius furzeri*), Amazon molly (*Poecilia formosa*), medaka (*Oryzias latipes*), three-spined stickleback (*Gasterosteus aculeatus*), Siamese fighting fish (*Betta splendens*), paradise fish (*Macropodus opercularis*), goldfish (*Carassius auratus*), zebrafish (*Danio rerio*). The protein sequences used were downloaded from the Uniprot database. Paradise fish orthologs were extracted from a previously assembled, unpublished *de novo* transcriptome database, by my supervisor.

Normalized body length measurement

Photographs were taken of ten male and ten female fish, all of them previously anaesthetized with 120 mg/L final concentration of Tricaine (Sigma Aldrich, MS222) as required by standard protocols. I analyzed the images with the help of ImageJ, by measuring their total length, fork length and standard length (Figure 4). To determine whether the difference in tail length between males and females is significant, the results were normalized to their total length, as males tend to be larger.

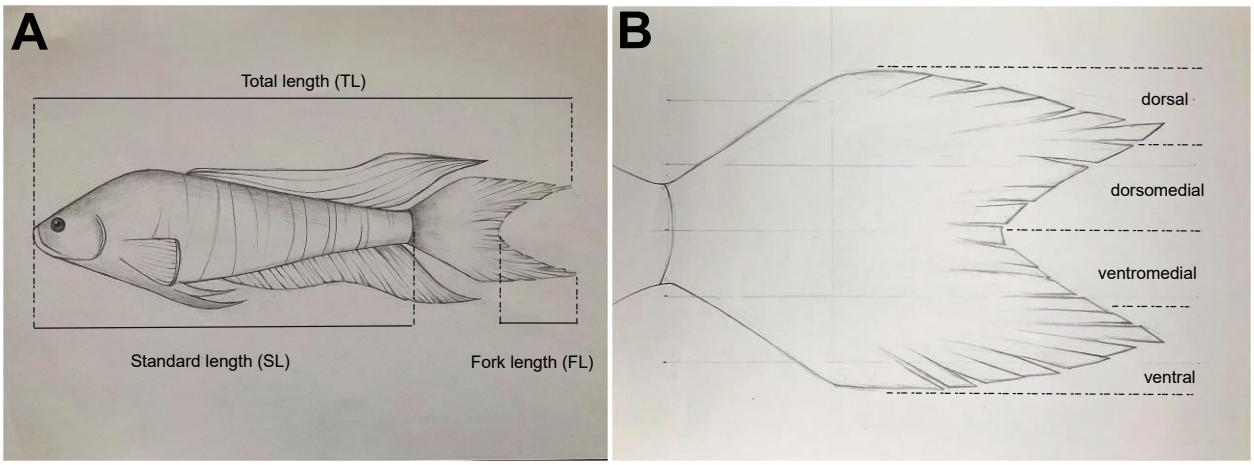

Figure 4: (A) Different types of length measurements. Total length (TL), Fork length (FL) and Total length (TL) of a fish. (B) Tail regions sampled. (Drawings from Júlia Szabó)

Fin ray measurement

The measurement of the fin rays occurred in a similar manner to the body length measurement. For this experiment we used the same fish as mentioned in the previous paragraph. I counted the number of their fin rays and measured their length by ImageJ. For rays that contained bifurcations I registered the length on the longest fin ray. To obtain objective results the lengths were also normalized to the body length.

Total RNA isolation

For transcriptome analysis and quantitative PCR (qPCR) assays we extract RNA from the amputated caudal fins of three male and three female fish. Before amputation animals were anesthetized with 120 mg/L final concentration of Tricaine and 4-5 mg/L final concentration of Lidocaine (Sigma Aldrich, L7757). The excised sections were separated into dorsal, dorsomedial, ventromedial and ventral sections (Figure 4B). Total RNA was extracted using Trizol (ThermoFisher Scientific, 15596026), according to the manufacturer's protocol. My supervisor, Máté Varga previously isolated RNA from brain samples, which were also used in our experiments.

cDNA synthesis

The RNA extracted from the samples was transcribed to cDNA using the SuperScript™ IV VILO™ Master Mix kit (ThermoFisher Scientific, 11756050), following the manufacturer's protocol. The reverse transcription reaction contained a mixture of random primers and oligo(dT)s.

Primer design for qPCR

We designed the primers (Table 1) for quantitative polymerase chain reaction using the primer design function of Benchling (http://benchling.com). The primers must be of the optimal length, they need to have the correct GC contents and they need to satisfy all the criteria for qPCR primers as it is described in the Maxima SYBR Green qPCR Master Mix (2X) (ThermoFisher Scientific, K0221) protocol. We designed primers both for a housekeeping gene that we plan to use as control, *β-actin*, and for our gene of interest, *kcnb2l*. The effectiveness of the primers was tested first by standard PCR, using different annelation temperatures.

Gene/genomic region	**Primer**	**Sequence (5'-3')**
kcnb2l	Mo-kcnb2l_qPCR-F1	AACCCAGCTCTTCTGTGGCTGC
	Mo-kcnb2l_qPCR-F2	ATGAGGAGCTGAGGAGGGAGGC
	Mo-kcnb2l_qPCR-F3	AGAAGGGGAGGAGGAGGAGCAC
	Mo-kcnb2l_qPCR-F4	TGTGGCTGCAAAGATCCTGGCC
	Mo-kcnb2l_qPCR-R1	CTGCAGCTCTGGCAACGTGTTG
	Mo-kcnb2l_qPCR-R2	GCTTCCACATGAGCCAGACTGGG
	Mo-kcnb2l_qPCR-R3	GACCACCTGCAGCTCTGGCAAC
β-*actin*	Mo-actb2_qPCR-ex1-F	CGGATTCACTCTGAGCGCCGTC
	Mo-actb2_qPCR-ex2-F	CGGATCCGGTATGTGCAAAGCCG
	Mo-actb2_qPCR-ex2-R	TCACTCCCTGATGTCTGGGGCG
	Mo-actb2_qPCR-ex3-F	CACAATACCGTGCTCGATGGGG
kcnb2l - enhancer	Mo-kcnb2l-enh-F	TGATGGAATCAGGTGTGTCACAC
	Mo-kcnb2l-enh-R	AAATGCGTCAGACTGGCAGGGC
	kcnb2l -enh-R	GCCCTGCCAGTCTGACGCATTT
	kcnb2l -enh-F1	TGAGAAGATGCCGCTAGGATGAG
	kcnb2l -enh-R1	TGCTTCGTGGGTGTTGCAGTCC

Table 1: qPCR primers designed for the kcnb2l and β-actin genes and the kcnb2l enhancer region

For the PCRs we made 12,5 µl reaction mixes containing 10,375 µl of RNase free water, 1,25 µl 10x DreamTaq Green Buffer (ThermoFisher Scientific, EP0713), 0,25 µl dNTP, 0,25 µl cDNA, 0,25 µl primer mix and 0,125 µl DreamTaq DNA Polymerase (ThermoFisher Scientific, EP0713). Then these were run using the following PCR protocol (Table 2):

Initial heating	**Denaturation**	**Annealing**	**Extension**	**Final extension**	**Hold temp.**
95.0 C^o	95.0 C^o	62.0 C^o	72.0 C^o	72.0 C^o	4.0 C^o
1:00 min	0:30 min	0:20 min	0:20	5:00 min	∞
	30 cycles				

Table 2: The PCR protocol used for cDNA synthesis

To check whether they worked or not, we used gel electrophoresis. Running the samples for 20 minutes on 120 Volts.

Making agarose gels

Agarose gel electrophoresis was used to determine the presence of RNAs and the correct size of the DNAs. This required SeaKem LE Agarose (Lonza, 50004), TAE/SB buffer and ethidium bromide (Invitrogen™, 15585011). In our experiments we used 0.8%-1% agarose gel and the nucleic acids detected under UV light.

Cloning of the kcnb2l enhancer region

The ~3.6 kb long genomic region upstream from *kcnb2l* and the neighboring *dnaJ homolog subfamily B member 6-like* (*dnajb6l*) was identified in the (unpublished) draft paradise fish genome by my supervisor.

In order to amplify this region, likely to contain the putative enhancers of *kcnb2l* we extracted high quality genomic DNA from the tail of a male paradise fish. To achieve this, we followed the corresponding protocol from Westerfield's the zebrafish book (14).

The primers (Table 1) used for amplification contained digestion sites of the rare cutting restriction enzymes AscI (ThermoFisher Scientific, ER1891) and FseI (New England BioLabs, R0588S) in their 5' ends. For amplification we used the Phusion™ Plus Green PCR Master Mix (Thermo Fisher Scientific, F632S). The end product was purified using standard phenol-chloroform purification (v). To confirm the amplicon, we sequenced the purified PCR product by Sanger Sequencing.

After the validation the PCR product was digested using AscI and FseI and ligated into a Tol2 transgenesis plasmid in front of a reporter gene encoding an enhanced green fluorescent protein (eGFP) – a gift from my lab colleague, Tamás Annus. The ligated construct was transformed into DH5α competent cells and these were grown on ampicillin-containing medium. Plasmids containing our insert were then recovered using ZymoPURE™ Plasmid Miniprep Kit (Zymo Research, D4212). Lastly, plasmids were test-digested with BamHI restriction enzyme (ThermoFisher Scientific, ER0051) and Sanger sequenced.

Transient transgenesis

In order to create transgenic zebrafish line, Tol2 vectors containing the enhancer region for *kcnb2l,* and transposase protein were coinjected into the yolk of one-cell stage zebrafish embryos. For the injection we filled special glass capillaries with the aforementioned mixture and the tip of the needle was broken off with a fine tip tweezer under a light microscope. We optimized the size of the drops by adjusting the size of the tip of the capillary and the pressure of the nitrogen. Then we lined up the embryos on a Petri-dish containing a glass microscopy slide and injected them.

Statistics and visualization

Multiple sequence alignment (MSA) on sequences downloaded from the Uniprot database was performed using Clustal Omega (13). Phylogenetic tree visualization was performed using the Interactive Tree of Life (iTOL) script (16). Statistical analysis and visualizations were performed in R (17) with the help of the ggplot2 package (18).

RESULTS

Multiple sequence alignment

MSA is a tool that helps us study the homology between the sequences of our proteins of interest. It is important to note that even for proteins that originated from of common ancestral sequence, structure and function can diverge over time. However, by aligning their sequences we can study the homology between them, and infer to their structure, function and evolution.

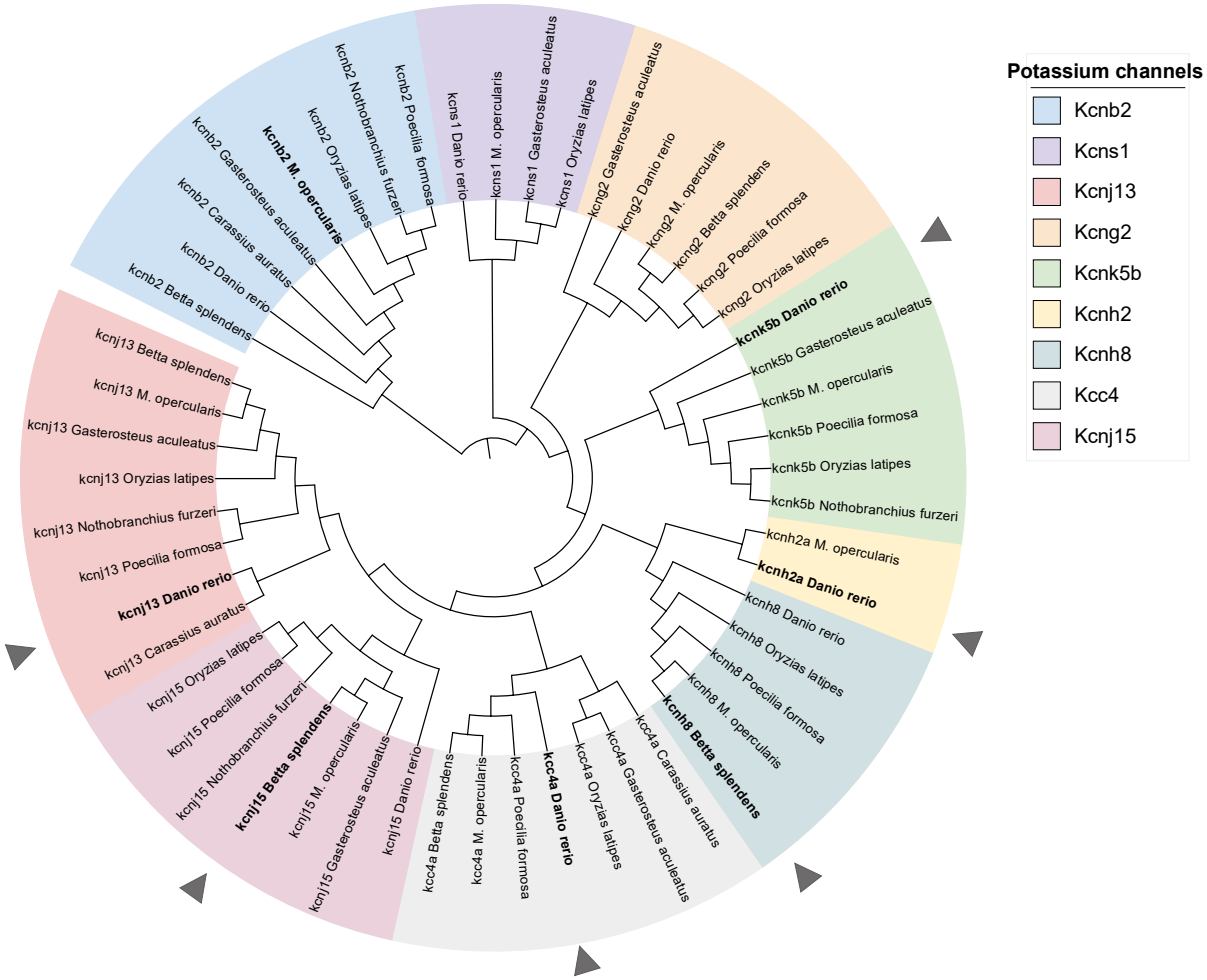

Figure 5: Circular phylogenetic tree showing the studied inwardly rectifying potassium channel proteins in different colors. The genes that are known to cause fin phenotypes are highlighted in bold and by the grey arrowhead. (Note: Kcns1 and Kcng2 are thought to be involved in the development of the swordtail, but the respective Xiphophorus *sequences were too fragmented to be used in a MSA).*

During our literature review we found multiple different fish species/varieties (namely the green swordtail, the Siamese fighting fish, some goldfish varieties and a number of zebrafish mutants) that showed a differential phenotype in their caudal fin length due to a mutation in one of their potassium channels. We decided to further study them by aligning the orthologous sequences (when available) of these proteins from multiple fish species with well annotated draft genomes (see Supplementary Information for a full list).

As a result of MSA a phylogenetic tree can be constructed, which shows us the evolutionary distances between the sequences and consequently the species (Figure 5). We can see that a large variety of potassium channels can be involved in regulating the bioelectric signals related to tail development. This distribution also suggests that the only common theme in the long-tail phenotypes is the breakdown of membrane potential due to the gain of function mutations of these potassium channels.

Normalized tail length measurement

To determine that the tails of the male fish are indeed longer than the tails of the females we first photographed them and measured their lengths in ImageJ. However, since the males tend to have larger bodies we normalized the outcomes with their standard length compared to their total length, as described in the Methods section. The results confirmed our hypothesis, the caudal fin of the males is effectively larger than that of their female counterparts (Figure 6A).

Fin ray measurements

Fish tails contain multiple fin rays, the number of which varies between individuals (in general paradise fish males and females have 13 fin rays, with some individuals having 14). By carefully measuring the length of individual fin rays (Figure 6B and C) I was able to show that the length of fin rays 3 to 11 is significantly longer in the average male than in the average female and rays 4, 5, 8, 9 and 10 are especially elongated in males.

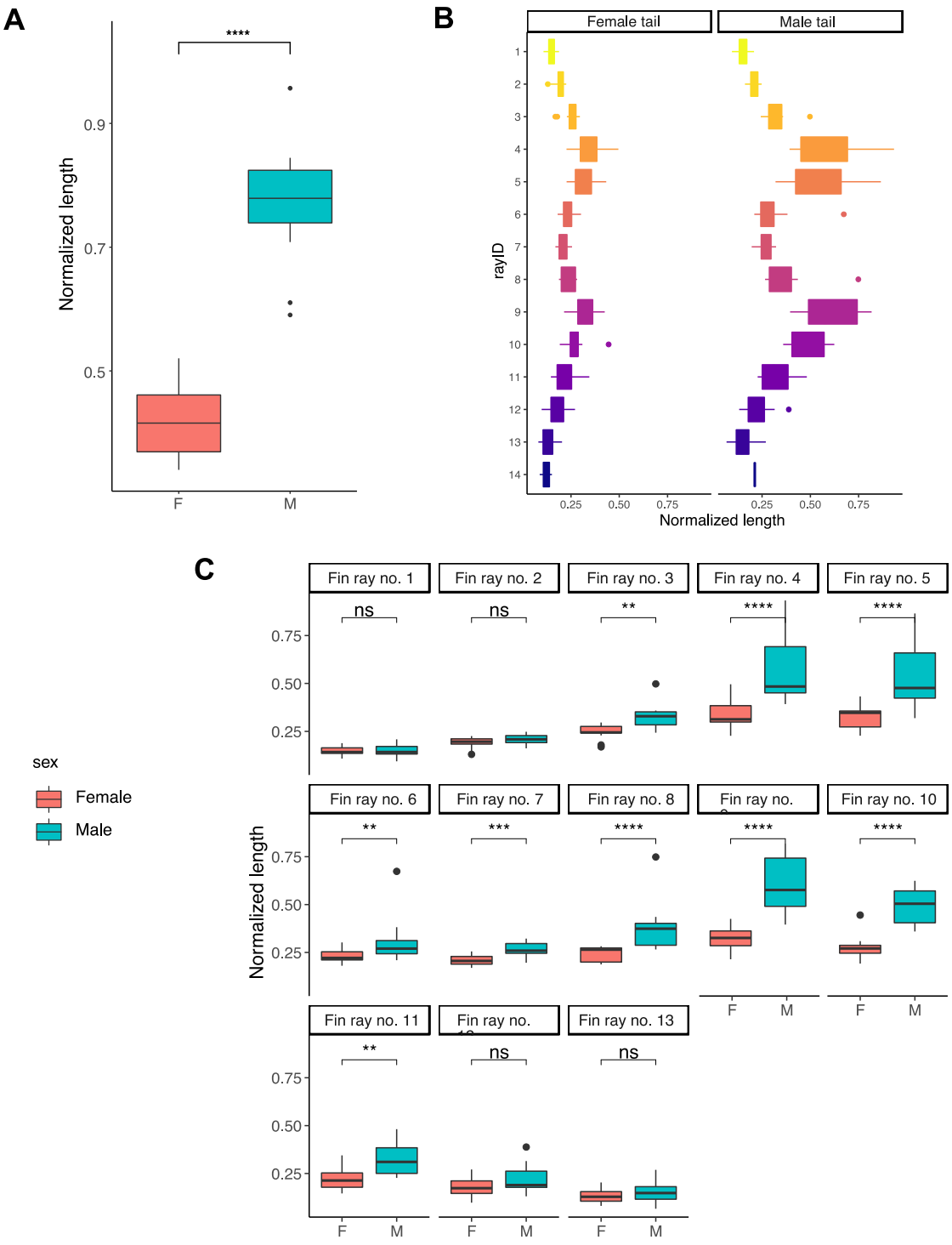

Figure 6: The length of some caudal fins and fin rays is sex-dependent in paradise fish. (A) Box plots showing the full normalized tail length of paradise fish. (B) Normalized length of the fin rays in the tail of paradise fish, numbered in descending order from dorsal to ventral. (C) Statistical comparison of the length of individual fin rays in females (F) and males (M). (Note as the 14th ray is absent in most individuals, we compared only fin rays 1 to 13. Pairwise p-values were calculated using Welch t-test; ns = not significant.)

Comparison of kcnb2l expression levels between paradise fish males and females

To study whether there is a difference in the mRNA levels of *kcnb2l* in the tail fin between male and female paradise fish, we wanted to perform quantitative PCR. To achieve this, we first had to isolate RNA from these tail fragments. The result of the RNA extractions are shown on Figure 7. In addition, we also used RNA from zebrafish caudal fin samples as control, since they are an already established model organism on which it works.

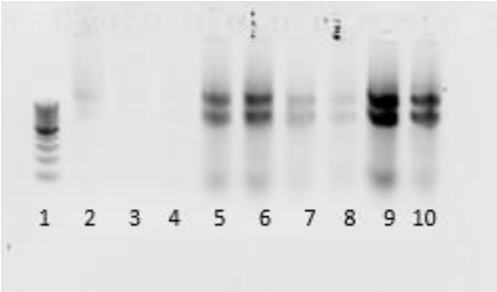

Figure 7: Isolated total RNA from zebrafish and paradise fish. The columns are: 1: GeneRuler 1 Kb Plus DNA Ladder (ThermoFisher Scientific, SM1331), 2-4: Zebrafish total RNA, 5: Paradise fish dorsal sample, 6: Paradise fish dorsal tail sample total RNA, 7: Paradise fish dorsomedial tail sample total RNA, 8: Paradise fish ventromedial tail sample total RNA, 9-10: Paradise fish ventral tail sample total RNA

Purified total RNA samples run on a standard agarose gel show two bands typical for ribosomal RNA species. Our results show that RNA isolation was successful in one of the male paradise fish. (Similar results were obtained for the other fish as well, but cannot be shown due to lack of space.)

Since our gene of interest was only recently identified from *de novo* sequencing, we needed to design novel primers suitable for qPCR analysis. This posed some challenges, as the primers for qPCR need to be really specific and fit several criteria.

First, we needed a reliable inner control for our analysis and we have chosen β-*actin* as it is present in all cells of animals. Our new primer pair amplified the expected 194 bp long amplicon in every sample that we tried them on (Figure 8D). We were able to confirm the identity of the amplicon by Sanger sequencing.

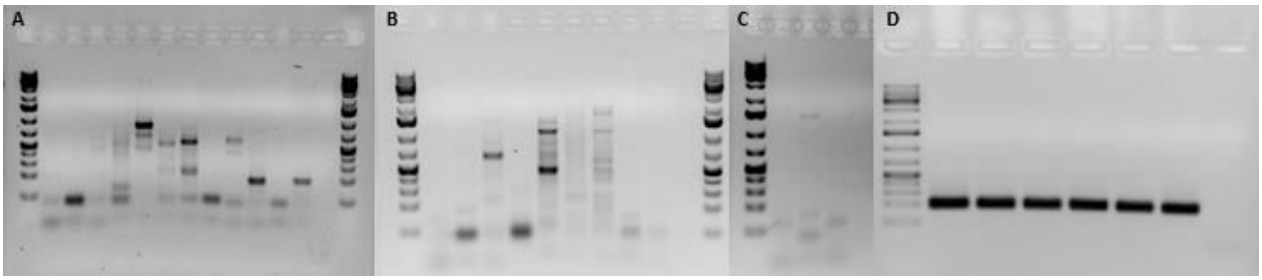

Figure 8: qPCR primers tested using cDNA isolated from paradise fish tail (A) and brain (B, C). There were 12 samples from different primer pairs loaded on agarose gel in the same order. D) qPCR primer for β-actin tested on several samples of tail specific paradise fish cDNA

Up to this point, however, we have not yet been able to design suitable primers for *kcnb2l*. Based on the orthologous sequence from the draft genome we have designed x forward and y reverse primers, where the amplicon would span the two (predicted) exons of the gene. We tested all the possible combinations of these primers (Table 1) however, we were unable to find (yet) the conditions that would provide us specific amplification (Figure 8A).

One possibility was, that our differential analysis based on transcriptomic datasets was erroneous and *kcnb2l* is not expressed in the tail of the fish. For this reason, we also tested our PCR primers on cDNA samples obtained from the brain since we expect *kcnb2l* to be expressed there. These reactions showed a similar pattern, thus confirming that the problem is with our primers (Figure 8B).

Next we tested if this gene is polymorphic in paradise fish and whether the draft genome sequence is contains numerous polymorphic sites compared to our samples. To check this, we sequenced larger fragments of the gene from novel cDNA samples as well. This revealed numerous polymorphisms compared to the draft genome sequence that likely affected the binding of our previous primers (not shown).

Cloning of the kcnb2l enhancer region and creation of zebrafish reporter lines

Since there is a difference in tail length between males and females, we suspected that there might be a male-specific regulatory sequence in the enhancer region of the gene, maybe related to testosterone. To check this, we first have to map the putative regulatory regions.

As a first step, we applied high quality genomic DNA (Figure 9A) extraction with specific primers to amplify the enhancer sequence of the *kcnb2l* gene. This fragment is 3643 base pairs long, as shown in Figure 9B.

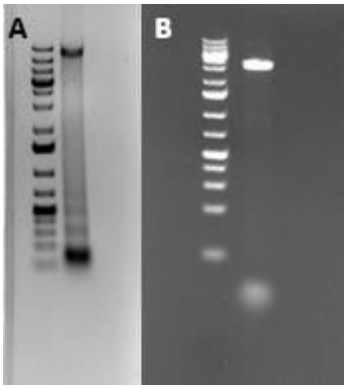

Figure 9: A) High quality genomic DNA extraction from paradise fish tissues. B) kcnb2l *enhancer region, amplified by Phusion polymerase on the right, gene ladder on the left*

The genomic fragment was PCR amplified and cloned into a Tol2 transgenesis plasmid. The two-component Tol2-system (derived from the Tol2 transposon) is the reference transgenesis tool in zebrafish (19). We intend to create transgenic zebrafish where the putative promoter and enhancers of the paradise fish *kcnb2l* drive the expression of a fluorescent reporter (eGFP), therefore we could visualize tissue- and sex-specific expression patterns (Tg(*Mo_kncb2l:eGFP*)). Our null hypothesis is that if *kcnb2l* is indeed expressed at higher levels in the tail of males, this difference could be due to hormonal differences between the two sexes. If our hypothesis is true, that could result in differential regulation of the reporter in the tails of male and female transgenic zebrafish as well.

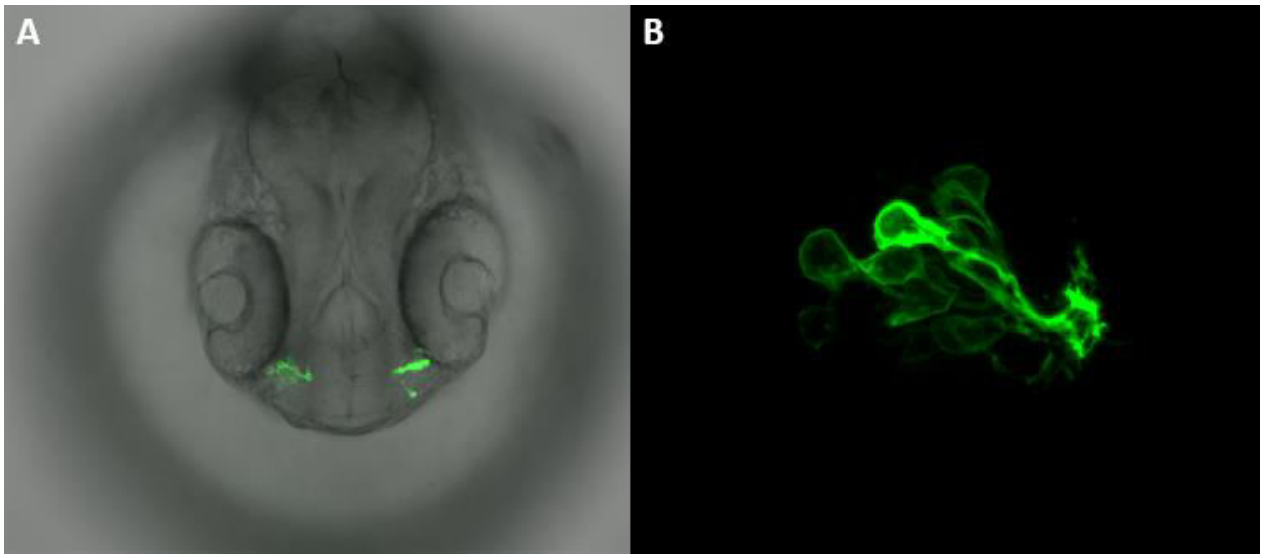

*Figure 10: A) 1 day old transgenic zebrafish embryo injected with Tg(*Mo_kncb2l:eGFP*) from top view, with eGFP expressing cells in the olfactory bulb. B) Close-up view of the eGFP expressing transgenic cells*

After injection, images of one day-old zebrafish were taken under an eGFP filter using an Apotome microscope. Specific signals were seen in the images, probably indicating the olfactory bulb (20,21) in 13 out of 16 fish (Figure 10). We confirmed this hypothesis using the zebrafish brain atlas, where we find the *y394* Gal4 allele , which is an enhancer-trap insertion of the Gal4 coding sequence into the *kcnb2* gene of zebrafish, is also expressed in the olfactory bulb (22). Based on these results, we could conclude that our work has indeed led to the cloning of endogenous enhancers of the paradise fish *kcnb2l* gene, hence the specific expression.

DISCUSSION

Paradise fish show sexual dimorphism in the length of their tail fin and the molecular mechanism behind this phenomenon is still unclear. Bioelectric signaling created by inwardly-rectifying potassium channels, however, has been linked to the development of gain-of-function tail phenotypes in multiple fish species. Interestingly, the respective potassium channels are quite different (Figure 5), suggesting their – shared – molecular function is the key for the development of the phenotype.

One of our primary aims was to confirm that the length of the caudal fin of male and female paradise fish differs. Analyses show that the caudal fin of males is generally longer than that of females, and that certain fin rays are significantly longer in males than in the other sex. After a transcriptome analysis showed that the expression of the *kcnb2l* gene, encoding a putative potassium channel, differs between the two sexes, we planned to compare the amount of *kcnb2l* mRNA present in the caudal fin of males and females by qPCR. As our work showed that our primers were strikingly aspecific, we sequenced a large part of the gene. This revealed that our gene contains several polymorphisms that affect the regions of the qPCR primers. In the future, we will design primers based on the sequencing results, hoping to amplify the appropriate sequence

With regard to the study of the *kcnb2l* enhancer region, our work has led to the cloning of endogenous enhancers. Injection a transgenic construct (Tg(*Mo_kncb2l:eGFP*)), where the putative *kcnb2l* regulatory region drives the expression of an eGFP reporter, resulted in fluorescence in the olfactory bulb, which was validated with a zebrafish brain atlas. The nearly identical expression of the transgene in multiple animals, plus the documented expression of the endogenous zebrafish *kcnb2* gene in the olfactory bulb, suggests that the cloned genomic region upstream of the paradise fish *kcnb2l* indeed contains enhancers of the gene (Figure 10). Furthermore, conserved expression in the olfactory bulb also demonstrated the presence of enhancer regions with evolutionary highly function in both species. In the future, we would like to follow the development of our transgenic zebrafish and see whether the enhancer region has a testosterone dependent effect in the caudal tail. Of note, the *y394* Gal4 reporter shows expression in the zebrafish pectoral fin, thus *kcnb2* might be predisposed to be expressed in fins (22).

ACKNOWLEDGEMENTS

In the first place I would like to thank my supervisor Máté Varga for the opportunity to work on this topic, and his coordination and insight in every stage of the work. I am very grateful to Nóra Szabó for her tireless help with the experiments and her patience and guidance during the writing process. I am also thankful to Dávid Czimer for taking photos of the fish, Juli Szabó for the beautiful fish drawings and Barnabás Németh for his help, even without me needing to ask for it. Lastly, I would like to thank every member of the ELTE Danio lab for their support, as well as every unnamed contributor.

REFERENCES

1. Darwin, C. R. 1871. The descent of man, and selection in relation to sex. London: John Murray. vol. 1. [Internet]. [cited 2022 Oct 17]. Available from: http://darwin-online.org.uk/content/frameset?pageseq=1&itemID=F937.1&viewtype=text

2. Powell DL, Payne C, Banerjee SM, Keegan M, Bashkirova E, Cui R, et al. The Genetic Architecture of Variation in the Sexually Selected Sword Ornament and Its Evolution in Hybrid Populations. Curr Biol. 2021 Mar 8;31(5):923-935.e11.

3. Schartl M, Kneitz S, Ormanns J, Schmidt C, Anderson JL, Amores A, et al. The Developmental and Genetic Architecture of the Sexually Selected Male Ornament of Swordtails. Curr Biol. 2021 Mar 8;31(5):911-922.e4.

4. Gene Expression in the Three-Spined Stickleback (Gasterosteus aculeatus) of Marine and Freshwater Ecotypes - PubMed [Internet]. [cited 2022 Oct 29]. Available from: https://pubmed.ncbi.nlm.nih.gov/29713520/

5. Lui JC, Baron J. Mechanisms limiting body growth in mammals. Endocr Rev. 2011 Jun;32(3):422–40.

6. Perathoner S, Daane JM, Henrion U, Seebohm G, Higdon CW, Johnson SL, et al. Bioelectric Signaling Regulates Size in Zebrafish Fins. PLOS Genet. 2014 Jan 16;10(1):e1004080.

7. Levin M. Large-scale biophysics: ion flows and regeneration. Trends Cell Biol. 2007 Jun;17(6):261–70.

8. Zhang W, Wang H, Brandt DYC, Hu B, Sheng J, Wang M, et al. The genetic architecture of phenotypic diversity in the Betta fish (Betta splendens). Sci Adv. 2022 Sep 21;8(38):eabm4955.

9. Kon T, Omori Y, Fukuta K, Wada H, Watanabe M, Chen Z, et al. The Genetic Basis of Morphological Diversity in Domesticated Goldfish. Curr Biol. 2020 Jun 22;30(12):2260-2274.e6.

10. Wang L, Sun F, Wan ZY, Ye B, Wen Y, Liu H, et al. Genomic Basis of Striking Fin Shapes and Colors in the Fighting Fish. Mol Biol Evol. 2021 Aug 1;38(8):3383–96.

11. Silic MR, Wu Q, Kim BH, Golling G, Chen KH, Freitas R, et al. Potassium Channel-Associated Bioelectricity of the Dermomyotome Determines Fin Patterning in Zebrafish. Genetics. 2020 Aug 1;215(4):1067–84.

12. Harris MP, Daane JM, Lanni J. Through veiled mirrors: Fish fins giving insight into size regulation. WIREs Dev Biol. 2021;10(4):e381.

13. Rácz A, Adorján G, Fodor E, Sellyei B, Tolba M, Miklósi Á, et al. Housing, Husbandry and Welfare of a "Classic" Fish Model, the Paradise Fish (Macropodus opercularis). Animals. 2021 Mar;11(3):786.

14. ZFIN Publication: Westerfield, 2000 [Internet]. [cited 2022 Oct 28]. Available from: https://zfin.org/ZDB-PUB-101222-52

15. How to use phenol / chloroform for DNA purification: Thermo Fisher Scientific - US. How to Use Phenol / Chloroform for DNA Purification | Thermo Fisher Scientific - US. (n.d.). Retrieved October 28, 2022, from https://www.thermofisher.com/hu/en/home/references/protocols/nucleic-acid-purification-and-analysis/dna-extraction-protocols/phenol-chloroform-extraction.html.

16. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation - PubMed [Internet]. [cited 2022 Oct 28]. Available from: https://pubmed.ncbi.nlm.nih.gov/33885785/

17. R Core Team (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org.

18. Wickham H. Data Analysis. In: Wickham H, editor. ggplot2: Elegant Graphics for Data Analysis [Internet]. Cham: Springer International Publishing; 2016 [cited 2022 Oct 28]. p. 189–201. (Use R!). Available from: https://doi.org/10.1007/978-3-319-24277-4_9

19. Kawakami K. Tol2: a versatile gene transfer vector in vertebrates. Genome Biol. 2007 Oct 31;8(1):S7.

20. Zebrafish Brain Browser [Internet]. [cited 2022 Oct 28]. Available from: http://vis.arc.vt.edu/projects/zbb/

21. Gravagna NG, Knoeckel CS, Taylor AD, Hultgren BA, Ribera AB. Localization of Kv2.2 protein in Xenopus laevis embryos and tadpoles. J Comp Neurol. 2008;510(5):508–24.

22. Tabor KM, Marquart GD, Hurt C, Smith TS, Geoca AK, Bhandiwad AA, et al. Brain-wide cellular resolution imaging of Cre transgenic zebrafish lines for functional circuit-mapping. Raman IM, Stainier DY, Wyart C, editors. eLife. 2019 Feb 8;8:e42687.

i. *Huia: New Zealand Birds Online*. New Zealand Birds Online - The digital encyclopaedia of New Zealand birds. (n.d.). Retrieved October 29, 2022, from https://nzbirdsonline.org.nz/species/huia

ii. *Multicolored peacock on top tree branch, China peafowl glass feather, Peacock, animals, Peacock Feather, Galliformes PNG free download*. PNGWing. (n.d.). Retrieved October 29, 2022, from https://www.pngwing.com/en/free-png-bybxv/download

iii. *Male and female deer images – browse 29,801 stock photos, vectors, and video*. Adobe Stock. (n.d.). Retrieved October 29, 2022, from https://stock.adobe.com/search?k=male+and+female+deer

iv. Female Betta Fish. (n.d.). Retrieved October 29, 2022, from https://home.adelphi.edu/~ve21375/Female%20Betta%20Fish.html

v. . How to use phenol / chloroform for DNA purification: Thermo Fisher Scientific - US. How to Use Phenol / Chloroform for DNA Purification | Thermo Fisher Scientific - US. (n.d.). Retrieved October 28, 2022, from https://www.thermofisher.com/hu/en/home/references/protocols/nucleic-acid-purification-and-analysis/dna-extraction-protocols/phenol-chloroform-extraction.html.

SUPPLEMENTARY INFORMATION

Species	**Protein**	**UniProt code**
Betta splendens	Kcnj13	A0A6P7P6D7
	Kcnj15	G3QCH9
	Kcc4a	A0A6P7KY93
	Kcnh8	A0A6P7NY48
	Kcng2	A0A6P7NWR0
	Kcnb2	A0A6P7KV24
Carassius auratus	Kcnj13	A0A6P6LDG5
	Kcc4a	A0A6P6JB50
	Kcnb2	A0A6P6J8V2
Danio rerio	Kcnj13	Q0KIZ7
	Kcnj15	F1QQH0
	Kcc4a	X1WHL7
	Kcnh8	F1QTV6
	Kcnh2a	D9IW08
	Kcnk5b	Q5RGB0
	Kcng2	E7FFQ6
	Kcns1	UPI0000F1E400
	Kcnb2	A1L238
Gasterosteus aculeatus	Kcnj13	G3N5F7
	Kcnj15	G3QCH9
	Kcc4a	G3PNE3
	Kcnk5b	G3NYL0
	Kcng2	G3P1H2
	Kcns1	G3NYL0
	Kcnb2	G3NE98
Nothobranchius furzeri	Kcnj13	A0A1A8AUY3
	Kcnj15	A0A8C6L1U6
	Kcnk5b	A0A1A8AZL4
	Kcnb2	A0A1A8AYW9
Oryzias latipes	Kcnj13	A0A3B3DYD0
	Kcnj15	A0A3B3BA67
	Kcc4a	A0A3B3DL29
	Kcnh8	A0A3B3DQX4
	Kcnk5b	A0A3B3CIY5
	Kcng2	A0A3B3C038
	Kcns1	UPI0000EA1C69
	Kcnb2	A0A3B3BW59

Poecilia formosa	Kcnj13	A0A087XJI5
	Kcnj15	A0A087YS89
	Kcc4a	A0A087Y5L3
	Kcnh8	A0A087XG10
	Kcnk5b	A0A096MH82
	Kcng2	A0A087Y3N5
	Kcnb2	A0A087Y4P9

Table S1: UniProt codes of the proteins used for the MSA, the orthologues for the paradise fish were extracted from a previously assembled, unpublished de novo *transcriptome database*