

Enhancer Co-option During Vertebrate Evolution

A Thesis

submitted to

Indian Institute of Science Education and Research Pune in partial fulfilment of
the requirements for the BS-MS Dual Degree Programme

by

Apoorvaa Gopal



Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,

Pashan, Pune 411008, INDIA.

Date: March, 2026

Under the guidance of

Supervisor: Dr. Tatjana Piotrowski

Stowers Institute for Medical Research

From May 2025 to Mar 2026

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

Certificate

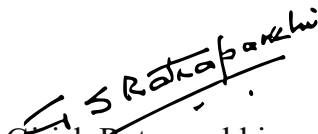
This is to certify that this dissertation entitled “ Enhancer Co-option During Vertebrate Evolution” towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Apoorvaa Gopal at the Stowers Institute for Medical Research under the supervision of Dr. Tatjana Piotrowski, Investigator, Stowers Institute of Medical Research, during the academic year 2025-2026.

Tatjana Piotrowski

Committee:



Tatjana Piotrowski



Girish Ratnaparkhi

This thesis is dedicated to *my beloved*

Declaration

I hereby declare that the matter embodied in the report entitled “ Enhancer Co-option During Vertebrate Evolution ” are the results of the work carried out by me at the Stowers Institute for Medical Research, under the supervision of Dr. Tatjana Piotrowski and the same has not been submitted elsewhere for any other degree. Wherever others contribute, every effort is made to indicate this clearly, with due reference to the literature and acknowledgement of collaborative research and discussions.



Name of Student: Apoorvaa Gopal

Roll #: 20211159

1. Table of Contents

2.	ABSTRACT	9
3.	ACKNOWLEDGMENTS.....	10
4.	CONTRIBUTIONS	11
CHAPTER 1 INTRODUCTION		12
1.	GENE CONTENT v/s GENE REGULATION IN EVOLUTION.....	12
2.	ENHANCERS AS SUBSTRATES FOR EVOLUTION.....	13
3.	ROUTES FOR ENHANCER EVOLUTION	14
4.	WATER TO LAND TRANSITION AS AN EVOLUTIONARY FRAMEWORK	15
5.	ZEBRAFISH AS A MODEL AQUATIC VERTEBRATE.....	16
6.	MOUSE AS A MODEL TERRESTRIAL VERTEBRATE	17
7.	THE LATERAL LINE AS A MODEL SYSTEM TO STUDY ENHANCER CO-OPTION	18
8.	EVOLUTIONARY SCENARIOS FOR CONSERVED LATERAL LINE ENHANCERS	19
9.	RESEARCH QUESTION AND HYPOTHESIS	20
10.	SCOPE OF THE STUDY.....	20
CHAPTER 2 MATERIALS AND METHODS.....		22
11.	ANIMAL MODEL AND MAINTENANCE	22
12.	GENOMIC ANALYSIS	22
13.	PRODUCTION OF TRANSGENIC ZEBRAFISH LINES.....	22
14.	PRODUCTION OF TRANSGENIC MICE	23
15.	FIXATION, DAPI STAINING, AND MOUNTING OF EMBRYOS FOR IMAGING	23
16.	GENOTYPING ZEBRAFISH PIGLET LINES.....	23
17.	MOUSE GENOTYPING	24
18.	IMMUNOFLUORESCENCE STAINING AND HISTOLOGICAL ANALYSIS.....	24
19.	HYBRIDIZATION CHAIN REACTION (HCR).....	24
20.	MICROSCOPY AND IMAGE ACQUISITION	25
CHAPTER 3 RESULTS		26
21.	ESTABLISHMENT OF STABLE ENHANCER TRANSGENIC LINES.....	26
22.	ZEBRAFISH ENHANCERS DRIVE REPORTER ACTIVITY IN THE LATERAL LINE AND INNER EAR	28
23.	CHARACTERIZATION OF ENHA ACTIVITY AND CANDIDATE TARGET GENES.....	30
1.	<i>Temporal and Spatial Activity of enhA During Development.....</i>	30
2.	<i>Identification of Putative Targets using Hi-C Data.....</i>	30
3.	<i>Single-Cell Transcriptomic Analysis Using Daniocell.....</i>	31
4.	<i>Validation of Enhancer–Gene Relationships Using HCR.....</i>	32
24.	COMPARISON OF ENHANCER ACTIVITY BETWEEN TOL2 AND PIGLET TRANSGENIC LINES.....	33
1.	<i>EnhA reporter activity in Tol2 and pIGLET lines</i>	33
2.	<i>EnhE Mouse Ortholog Reporter Activity in Tol2 and pIGLET Lines.....</i>	34
3.	<i>EnhH reporter activity in Tol2 and pIGLET Lines</i>	35
25.	SEQUENCE DIVERGENCE BETWEEN ORTHOLOGOUS ENHANCERS CAN INFLUENCE THEIR SPATIAL REGULATORY ACTIVITY	37
1.	<i>Spatial Activity of the orthologs of EnhA</i>	37
2.	<i>Spatial Activity of the Orthologs of EnhE</i>	40
26.	ACTIVITY OF ENHC IN ZEBRAFISH AND MOUSE EMBRYOS	42

1.	<i>Activity of the zebrafish ortholog of EnhC activity in zebrafish embryos</i>	42
2.	<i>Activity of the mouse ortholog of EnhC in transgenic mouse embryos</i>	43
27.	PROFILING OF ENHANCER ACTIVITY ACROSS MULTIPLE FOUNDERS	44
1.	<i>EnhA</i>	45
2.	<i>EnhB</i>	48
3.	<i>EnhC</i>	49
4.	<i>EnhE</i>	52
5.	<i>EnhF</i>	56
6.	<i>EnhG</i>	57
7.	<i>EnhH</i>	58
8.	<i>EnhJ</i>	58
CHAPTER 4 DISCUSSION		59
5.	REFERENCES	63

List of Figures

Figure 1: Multiple levels at which changes in gene regulation can occur.	12
Figure 2: Possible ways of emergence of new enhancers (Long et al. 2016)	14
Figure 3: Evolutionary transition from aquatic vertebrates to terrestrial tetrapods	15
Figure 4: Major stages of zebrafish embryonic development.	16
Figure 5: Structure and organization of zebrafish neuromasts in the lateral line system.	18
Figure 6 Comparison of Tol2-mediated random transgenesis and pIGLET-mediated targeted insertion in zebrafish.	26
Figure 7 Genotyping strategy and validation of pIGLET-mediated targeted insertions	27
Figure 8: Zebrafish orthologous enhancers drive reporter expression in the lateral line system.	28
Figure 9: Candidate enhancers drive activity in inner ear sensory organs and anterior lateral line neuromasts.	29
Figure 10: Spatiotemporal activity of enhA during zebrafish development.	30
Figure 11: Hi-C chromatin contact maps showing regulatory domains of enhA Hi-C interaction data along zebrafish chromosome 20, where enhA are located. The three enhancers lie within a single topologically associating domain (TAD) containing prkchb, six4b, six6b, six1b, and ppmlab, indicating these genes as potential regulatory targets within the same chromatin domain.	31
Figure 12: Single-Cell Daniocell dot plot shows that genes six4b and six1b are most highly expressed in the otic region and the lateral line, which overlaps with region of activity of enhA, making them the most likely regulatory targets	31
Figure 13: Validation of enhancer–gene relationships using HCR.	32
Figure 14: Comparison of EnhA-driven reporter activity in Tol2 and pIGLET transgenic zebrafish lines.	33
Figure 15: Comparison of EnhE-driven reporter activity in Tol2 and pIGLET transgenic zebrafish lines.	35
Figure 16: Comparison of EnhH-driven reporter activity in Tol2 and pIGLET transgenic zebrafish lines.	36
Figure 17: Spatial activity of the zebrafish EnhA enhancer and its mouse ortholog in zebrafish embryos at 5 dpf.	38
Figure 18: Developmental differences in the activity of zebrafish EnhA and the mouse ortholog.	39
Figure 19: Comparison of the spatial activity of the zebrafish EnhE enhancer and its mouse ortholog in zebrafish embryos.	40
Figure 20: Cell-type specific differences in neuromast activity driven by the zebrafish EnhE enhancer and its mouse ortholog.	41
Figure 21: Activity of the zebrafish ortholog of EnhC in zebrafish embryos.	42
Figure 22: Analysis of the mouse ortholog of EnhC in a transgenic mouse embryo.	43
Figure 23: EnhA Tol2 founder 2 established line exhibits stage-dependent activity in inner ear structures and lateral line neuromasts.	45
Figure 24: Mouse ortholog of EnhA (pIGLET24b) shows variable enhancer activity across independent founders at 5 dpf.	46
Figure 25: Mouse ortholog of EnhA (pIGLET24b, founder 1) exhibits dynamic enhancer activity across developmental stages	47
Figure 26: EnhB (Tol2, founder 1) stage-dependent enhancer activity in zebrafish.	48
Figure 27: EnhC (Tol2, founder 1) exhibits restricted activity in inner ear structures and anterior lateral line neuromasts across development.	49
Figure 28: EnhC (Tol2) shows reproducible enhancer activity across independent zebrafish founders at 2 dpf.	50
Figure 29: Mouse ortholog of EnhC (pIGLET24b) founder 1 progeny activity across developmental stages.	51
Figure 30: EnhE (zebrafish ortholog, pIGLET24b founder 1) exhibits neural and lateral line-associated activity across developmental stages.	52
Figure 31: Mouse ortholog of EnhE (Tol2, founder 1) shows robust activity in lateral line neuromasts, ear-associated regions, and neural tissues at 5 dpf.	53
Figure 32: Mouse ortholog of EnhE (Tol2, founder 2) enhancer activity at 5dpf	54
Figure 33: Mouse ortholog of EnhE (pIGLET24b, founder 1) enhancer activity at 5 dpf.	55

Figure 34: EnhF (Tol2, founder 1 progeny) stage-dependent activity	56
Figure 35: EnhG (Tol2, founder 1 progeny) exhibits weak and transient enhancer activity restricted to early developmental stages.	57
Figure 36: EnhJ (zebrafish ortholog, pIGLET14a, founder 1) exhibits inner ear-specific enhancer activity at 5 dpf.	58

2. Abstract

Phenotypic diversity across species is often driven by changes in gene regulation rather than the gene content itself. Cis-regulatory elements, including enhancers, form the cornerstone of regulating any developmental program. While enhancer sequences typically diverge across multiple species, a certain subset of these enhancers have their sequences conserved between zebrafish and mice. This project aims to investigate whether enhancers that remain conserved despite the water-to-land transition of vertebrates have been co-opted to carry out novel functions. The lateral line is a great system to study enhancer co-option because it was lost in terrestrial vertebrates, and despite this loss, some lateral line enhancer sequences are still present in the mouse. This gives us a unique set of enhancers that are conserved in a species that no longer possesses the organ itself. This raises compelling questions about what these enhancers are doing in the mouse. We hypothesize that these sequences have been evolutionarily conserved because they acquired new developmental functions associated with the emergence of novel phenotypes. To test this hypothesis, we conducted analyses of ATAC-seq data combined with sequence conservation, to identify a set of 72 putative lateral line enhancers whose sequences were conserved in both zebrafish and mice. To assess enhancer activity, we carried out enhancer-reporter assays, which revealed 12 enhancer candidates that could drive reporter expression. Comparative analysis of the activity of zebrafish and mouse enhancer orthologs in zebrafish revealed that while many enhancer orthologs exhibited a great overlap in their spatial activity, notable differences in tissue and cell-type-specific expression were observed. Additionally, comparison of the orthologs of the same enhancer in their native species, as well, revealed differences in their spatial activity, while some overlap in their regulatory outputs was still retained. Together, these findings suggest that conserved enhancers may have been repurposed across vertebrate species during evolution, while function in similar sensory structures was still retained. This work provides insights into how enhancer elements associated with lost organs may have been co-opted to contribute to novel developmental programs during vertebrate evolution.

3. Acknowledgments

I would first sincerely like to thank Dr. Tatjana Piotrowski for giving me the opportunity to join her lab and work on such an exciting project. Our meetings and discussions always motivated me to think deeply and strive to improve at every step. I had a great time interacting with her, and learning from her expertise.

My deepest gratitude and love goes to Dr. Aurélie Hintermann, who has been an exceptional mentor throughout the project. She has played a key role in shaping the way I think about science and the kind of scientist I want to become. I will never tire of admiring her, not only for her scientific skills, but also for her kind and patient guidance despite my shortcomings. When I was unsure of myself, she never let me feel discouraged. Conversations with her have always been insightful, engaging and would often lift my mood. I have learnt an immense amount from her and yet I feel there remains so much more to be learnt. She continues to be a great source of inspiration for me.

I am very grateful to the Piotrowski lab members for creating a very supportive environment. I am especially thankful to Daniela, who became my go-to anatomy expert, and was generous enough to let me ride with her which made the Kansas City winters bearable. I would like to thank Paloma, whose ATAC-seq data formed the foundation of my study, to Kelsey, who helped me getting started in the lab and Raman who was always so supportive and shared his pearls of wisdom with me.

I would also like to thank the wonderful core facilities at the Stowers Institute, especially the Aquatics facility, who ensured our fish were happy and the Mouse facility who supported the mouse work involved in this project. I would like to thank Shenping from histology and Cindy and Zulin from microscopy for their help and guidance in imaging of mouse embryos.

Finally I would like to thank the people who are closest to me, whose love and support made this journey possible. My parents, who have been my back bone and my constant source of encouragement, no matter where I am in the world. I am endlessly grateful for their love, blessings and their unwavering belief in me. My younger brother who would unconsciously lift my mood by his silly humour and love. I am also deeply thankful to my batchmates at Stowers, who became my family far from home and with whom I have created many cherished memories during my time here. To my cousins, whose weekly calls filled with jokes and laughter always brightened my days, you never let me feel lonely, and I love you all for that. My friends from IISER Pune, Deepika and Ishaan from whom I never felt apart, despite our distances. Finally I thank Deep, who was always there to listen, support and encourage me through every high and low.

4. Contributions

Contributor name	Contributor role
Aur�lie Hintermann, Apoorvaa Gopal	Conceptualization Ideas
Apoorvaa Gopal, Dr Aur�lie Hintermann	Methodology
Apoorvaa Gopal, Dr Aur�lie Hintermann	Software
Apoorvaa Gopal	Validation
Apoorvaa Gopal	Formal analysis
Apoorvaa Gopal, Dr Aur�lie Hintermann	Investigation
Stowers Institute for Medical Research	Resources
Apoorvaa Gopal	Data Curation
Apoorvaa Gopal	Writing - original draft preparation
Apoorvaa Gopal, Dr Aur�lie Hintermann, Dr Tatjana Piotrowski	Writing - review and editing
Apoorvaa Gopal	Visualization
Dr Tatjana Piotrowski	Supervision
Dr Tatjana Piotrowski	Project administration
Dr Tatjana Piotrowski	Funding acquisition

AI Usage

I declare that the usage of Generative AI in this thesis did not extend beyond what is described in Section 4.6.1 of the ‘Guidelines for Generative AI usage at IISER Pune’, and accordingly, no attribution is required.

This contributor syntax is based on the Journal of Cell Science CRediT Taxonomy¹.

¹ <https://journals.biologists.com/jcs/pages/author-contributions>

Chapter 1 Introduction

It is not the most intellectual of the species that survives; it is not the strongest that survives; but the species that survives is the one that is able to adapt to and to adjust best to the changing environment in which it finds itself

— Megginson's interpretation of Charles Darwin's "Origin of Species" (Megginson 1963)

1. Gene Content v/s Gene Regulation in Evolution

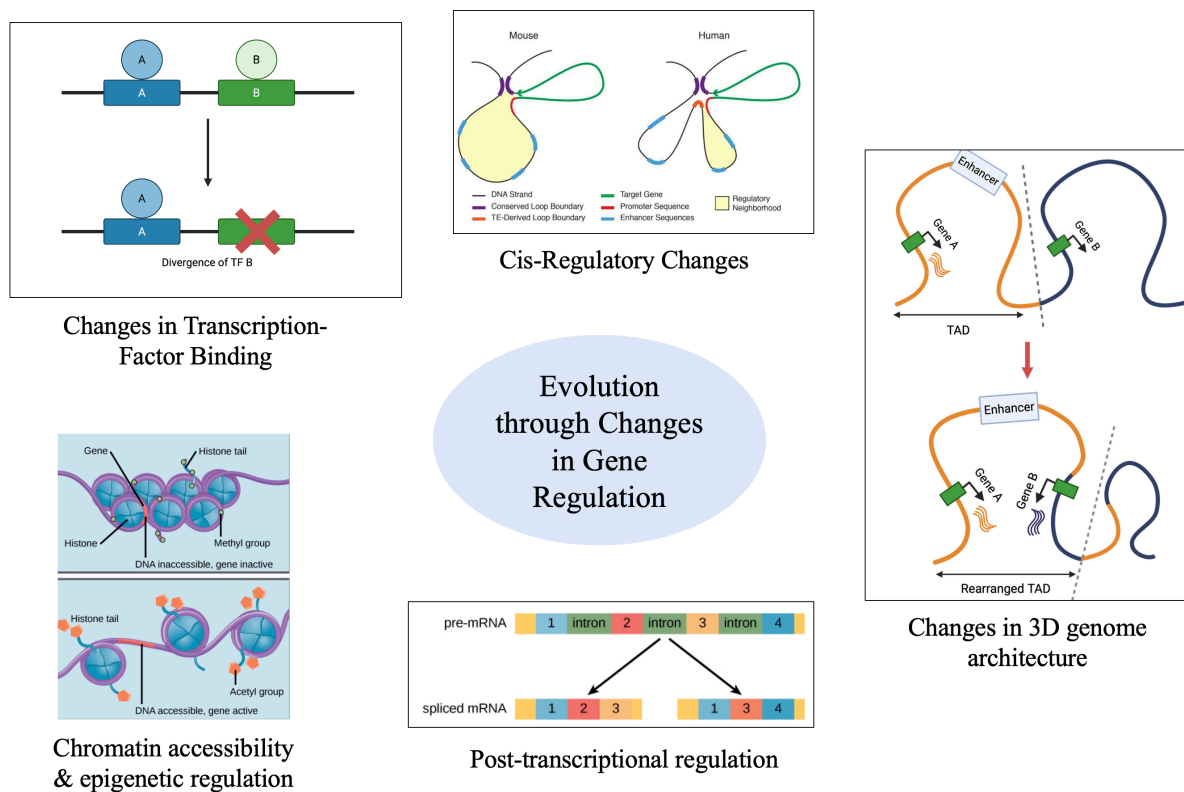


Figure 1: Multiple levels at which changes in gene regulation can occur.

Phenotypic diversification across species occurs often due to changes in gene regulation than in the gene content itself. There are several mechanisms by which regulatory evolution can occur. Changes in transcription factor binding motifs can occur through sequence mutations that can alter the binding affinity of transcription factors to their motifs. Cis-regulatory changes can change the spatio-temporal expression of target genes. Chromatin accessibility and epigenetic regulation influence whether regulatory regions are accessible to transcription factors through histone modifications and chromatin remodelling (Lumen Learning, n.d.). Post-transcriptional regulation, including alternative splicing, can generate multiple transcript isoforms from a single gene. Finally, changes in the 3D chromatin structure, such as alterations to topologically-associated domains (TADs) can alter regulatory neighbourhoods of regulatory elements.

One of the central questions in evolutionary biology is to understand the principles behind the phenotypic diversity we see across the tree of life. Earlier studies on comparative genomics have revealed that the protein-coding regions of most organisms are far more similar than the phenotypic diversity suggests. For example, the *hox* genes, widely known to establish the body plan are known to be broadly conserved, even across deeply diverged lineages of

metazoans (Gaunt 2022). While events such as gene duplication, loss and divergence have undoubtedly contributed to evolution, changes in gene content itself is not sufficient to explain the vast differences (Carroll 2008) in body plans, organ systems, and cell types that we see across these organisms.

This realization has led to a paradigm shift, with changes in gene regulation (Carroll 2008) emerging as a major driver of phenotypic evolution. Rather than inventing novel genes, evolution can occur via changes in the spatiotemporal regulation of genes, which might result in major morphological and functional diversifications. Thus, understanding the molecular mechanisms underlying the regulatory logic is essential to explain how complex traits arise and diversify.

Changes in gene regulation can occur at cis or trans-regulatory levels (Carroll 2008; Hill et al. 2021; Signor and Nuzhdin 2018). Trans-regulatory evolution occurs due to changes in the diffusible material, like transcription factors and co-factors. Changes in their abundance, spatiotemporal expression patterns and DNA-binding specificity can modify phenotypic outcomes. In contrast, cis-regulatory evolution involves changes in DNA elements that control gene expression. These changes can occur at the level of DNA sequence itself (Long et al. 2016), which could modify transcription factor binding motifs leading to alterations in the spatiotemporal enhancer activity. Additionally, epigenetic features such as DNA methylation and histone modification influence chromatin accessibility and transcriptional competence of regulatory elements, thereby modulating enhancer activity (Villar et al. 2015). (Cande et al. 2009; Lupiáñez et al. 2015)

Beyond sequence level changes, cis-regulatory evolution could also occur at the level of changes in the chromatin structure and its 3D organisation. Alterations at higher order chromatin level like the topologically-associated domains (TADs) determine the regulatory neighbourhood of enhancers. Genomic rearrangements which change the position of regulatory elements relative to their target genes, change their regulatory context, which in turn could generate novel gene expression patterns.

2. Enhancers as Substrates for Evolution

While changes in the above-mentioned regulatory levels are important for evolution, they do have major pleiotropic costs associated with them, as they affect multiple genes or their regulatory interactions simultaneously. In contrast, changes at the level of cis-regulatory elements, like enhancers, provide for rather flexible substrates for evolution due to their modular behaviour, making a significant proportion of enhancers, function in a cell-autonomous manner.

Enhancers are short DNA elements, which can bind to transcription factors and drive precise spatiotemporal expression of genes during development (Levine 2010; Spitz and Furlong 2012). The speciality of enhancers is that they function in a tissue-specific manner. A gene can be regulated by multiple enhancers, where each enhancer can be active in a different tissue, cell-type or developmental stage. This modular behaviour of enhancers allows for alterations in their function in one context, without perturbing core gene functions in other contexts (Carroll 2008). This property makes enhancers, amenable for “experimentation” by evolution.

Importantly, the activity of enhancers is impingent upon two factors; its sequence and its context. While the sequence determines its affinity and specificity towards transcription

factors, its context, whether cell-type, developmental or environmental, determines the availability of transcription factors. Slight changes in either of these factors can generate meaningful phenotypic consequences without disrupting core developmental programs.

3. Routes for Enhancer Evolution

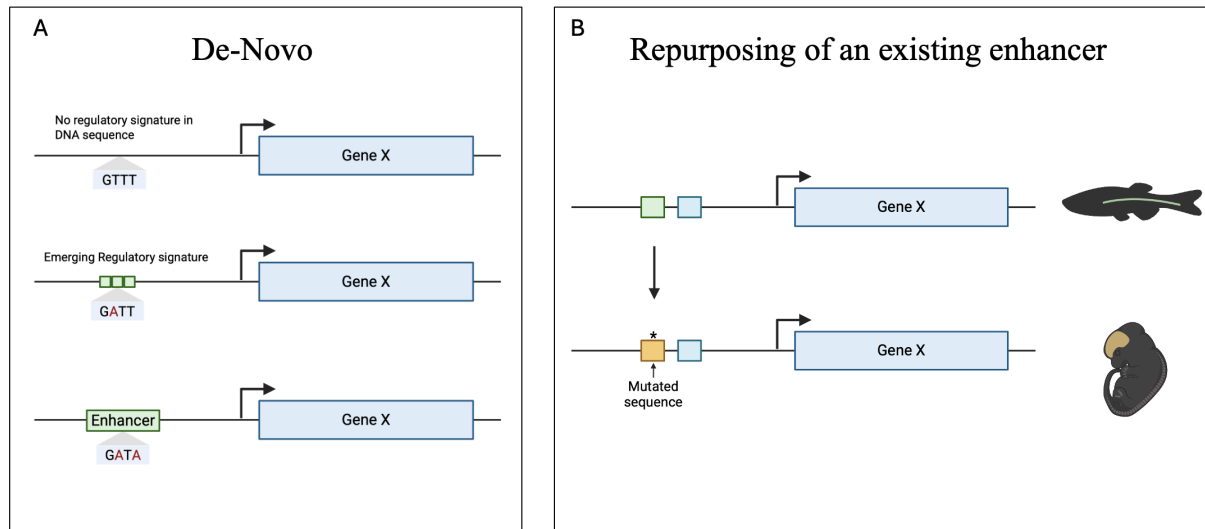


Figure 2: Possible ways of emergence of new enhancers (Long et al. 2016)

The establishment of new expression patterns relies on the emergence of new regulatory activities. New regulatory activities can emerge either de novo, by spontaneous mutations or insertion of transposable elements, or by duplication or divergence of existing regulatory elements..

De-novo emergence is fairly common in nature, where previously non-regulatory elements can gain regulatory signatures by random mutations. There have been experimental studies which show that synthetic random sequences are capable of driving reporter function with different elements capable of driving expression in unique spatiotemporal domains, which also suggests that a lot of the non-regulatory elements might be just few mutations away to acquire a regulatory signature (Galupa et al. 2023; McDonald and Reed 2024). Thus, de-novo evolution of enhancers is relatively common over evolutionary time scales.

An alternative route is **repurposing of ancestral enhancers**, where an existing enhancer, originally active in one tissue or developmental context, acquires a novel activity elsewhere, while it could still retain aspects of its ancestral function. This route is increasingly being argued in favour of, as ancestral enhancers already contain transcription factor binding motifs in accessible regions, compatible with the existing regulatory framework, making it a rather efficient route of evolution (Xin et al. 2020). Yet there is a significant lack of number of mechanistic case studies, which show enhancer co-option during vertebrate evolution. (Xin et al. 2020).

4. Water to Land Transition as an Evolutionary Framework

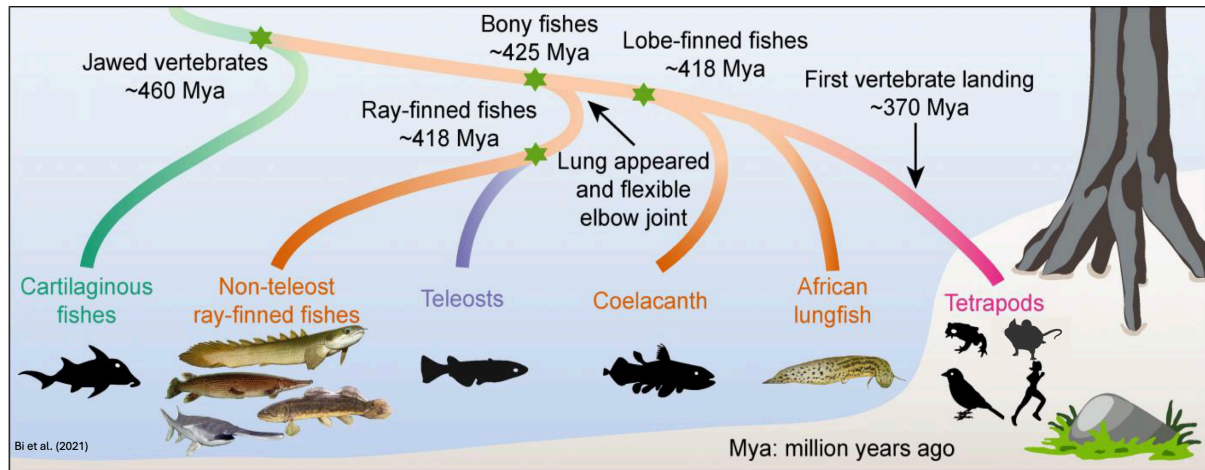


Figure 3: Evolutionary transition from aquatic vertebrates to terrestrial tetrapods

Schematic representation of major vertebrate lineages illustrating the water-to-land transition of vertebrates. Jawed vertebrates first diversified around ~460 million years ago, giving rise to cartilaginous fishes and bony fishes. Within bony fishes, ray-finned fishes and lobe-finned fishes diverged (~418–425 Mya). Lobe-finned fishes later gave rise to lineages such as coelacanths, lungfish, and ultimately tetrapods.

The water to land transition refers to the evolutionary period during the Late Devonian, when vertebrates first moved from aquatic environments to terrestrial habitats, and involved **dramatic phenotypic innovation**, including modifications to sensory systems, locomotion, physiological processes, and integumentary structures, to adapt to the new environment. This framework provides for a fertile ground to study enhancer co-option as, such large-scale morphological changes could not have occurred solely by changes in gene content, but rather through extensive rewiring of the existing regulatory networks.

5. Zebrafish as a model Aquatic Vertebrate

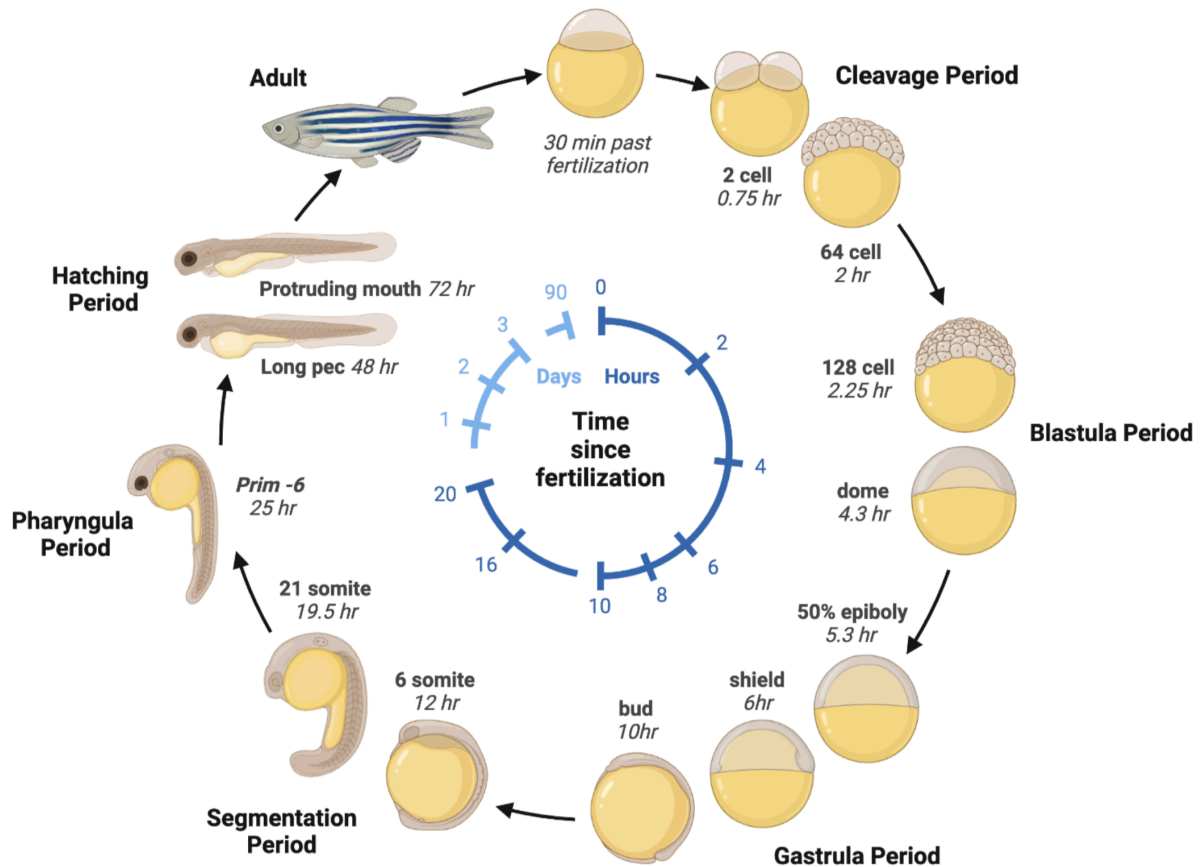


Figure 4: Major stages of zebrafish embryonic development.

Schematic overview of the major stages of zebrafish development from fertilization to adulthood. The embryos undergo rapid cleavage divisions following fertilization until the 128 cell stage. This is followed by epiboly during the blastula period, when the blastoderm spreads over the yolk sac. During gastrulation, major body axes are established and germ layers begin to form. Somite formation occurs during the segmentation period. Organ systems start to differentiate during the pharyngula period. Finally, hatching marks the emergence of larva from the chorion leading to their development to freely swimming juvenile. The embryo matures into a sexually mature adult by 90 days post fertilization.

The zebrafish (*Danio rerio*) is a freshwater teleost, originally found in the floodplains of South Asia, particularly India, Bangladesh and Nepal (Spence et al. 2008). In the wild, they are commonly found in shallow vegetative waters which are subject to frequent fluctuation in temperatures, oxygen levels and water flows. This ecological context demanded for a life history trait characterised by **rapid development, high fecundity and remarkable physiological resilience**. Zebrafish exhibits external fertilization, with females capable of laying hundreds of eggs weekly, under laboratory conditions (Westerfield 2000). The fertilized eggs undergo rapid embryogenesis with major organ systems establishing within 24 hours of fertilization. The larva hatch within 2-3 days of fertilization (Kimmel et al. 1995) and reach sexual maturity in approximately 3 months. These traits have made zebrafish a powerful vertebrate model organism.

Being a teleost fish places zebrafish at a unique phylogenetic position, where changes relative to the ancestral regulatory programmes can be studied while it still maintains strong genomic conservation with tetrapods. The zebrafish genome, annotated in detail (Howe et al. 2013), has largely conserved sets of developmental genes that complement with mammals despite undergoing an additional round of **whole-genome duplication** in the teleost lineage (Postlethwait et al. 1998). Importantly, even though the gene families are broadly conserved, the cis-regulatory landscape could have undergone rewiring.

One of the defining features of zebrafish as a model organism is the easy accessibility to the embryonic development, given the eggs undergo external fertilization. At the earlier stages of development, the embryos are optically transparent, making them suitable for in-vivo visualization of reporter activity. Additionally, zebrafish have benefited from the increasing number of genomic and epigenomic data, facilitated by ATAC-seq, ChIP-seq, Hi-C (Wike et al. 2021; Hintermann et al. 2025) and single-cell multi-ome (Farrell et al. 2018) datasets, giving us high-resolution maps for chromatin accessibility and transcriptional states. These datasets allow us to carry out a multi-layered and comprehensive study to characterise regulatory elements and compare their regulatory contexts.

6. Mouse as a Model Terrestrial Vertebrate

The mouse (*Mus musculus*) serves as a primary model for **mammalian systems**. As an amniote and a terrestrial vertebrate, it has undergone major morphological and ecological shifts compared to its aquatic ancestors. Their genome is extensively annotated and exhibits high levels of synteny with the human genome (Mouse Genome Sequencing Consortium et al. 2002). The regulatory landscapes of the mouse genome have undergone high-level mapping, including the chromatin accessibility profiles, histone modifications and the 3D chromatin structure.

The developmental stages of mouse embryos allow us to study early morphogenesis of systems, central to vertebrate development. The water-to-land transition involved extensive remodelling of the respiratory system, appendages and sensory structures crucial for terrestrial adaptation of vertebrates. Previous studies have demonstrated the regulatory changes underlying these innovations in the context of limb formation and reorganization of the regulatory landscape (Petit et al. 2017; Shubin et al. 2009).

Importantly, the mouse system allows us to functionally test non-coding elements in a terrestrial developmental context. Transgenic reporter assays and chromatin conformation assays enable us to map enhancer-gene interactions. These methods allow us to assess the absence or presence of activity in the orthologous non-coding elements in mouse relative to zebrafish and test whether the activity of a non-coding element was co-opted or silenced.

7. The Lateral Line as a Model System to study Enhancer Co-option

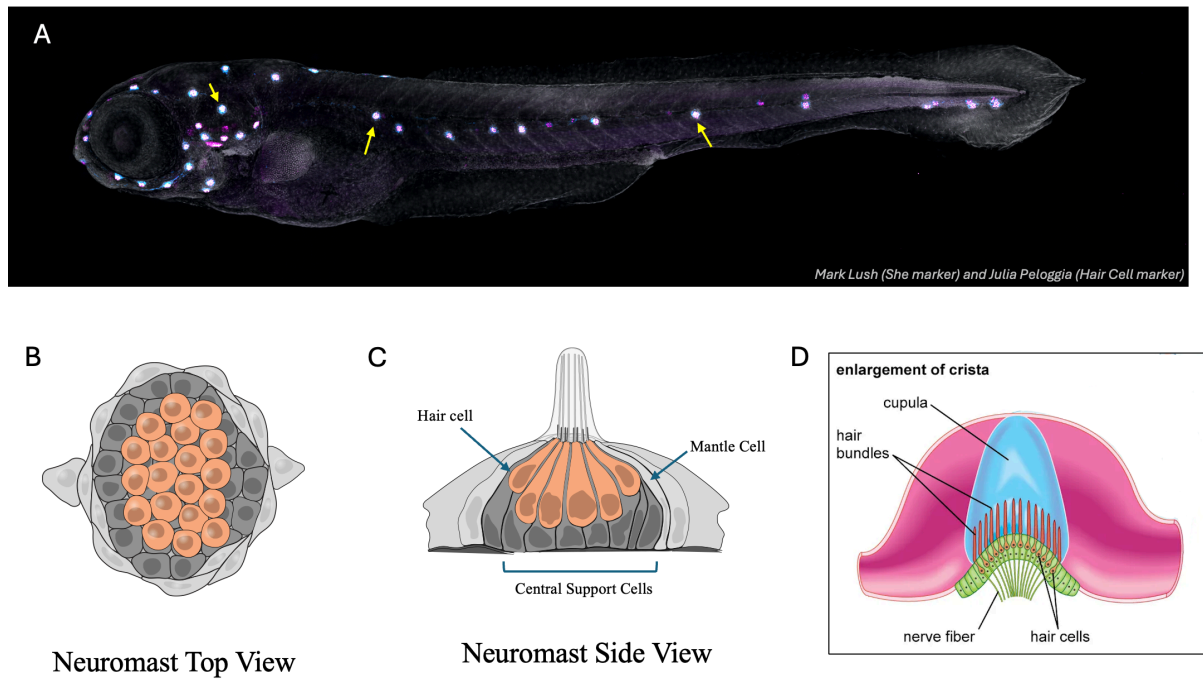


Figure 5: Structure and organization of zebrafish neuromasts in the lateral line system.

(A) Confocal image of a 5dpf zebrafish larva with the lateral line labelled in cyan and hair cells labelled in magenta (B) Schematic top view of a neuromast illustrating the rosette-like arrangement of centrally located hair cells (orange) surrounded by supporting and mantle cells. (C) Side view of a neuromast showing the organization of hair cells in the centre, flanked by supporting cells and mantle cells that form the outer layer. (D) Diagram of the mechanosensory structure of hair cells found in the vertebrate inner ear which detect vestibular fluid motion

The **lateral line of zebrafish** is a **mechanosensory organ system** widely present in aquatic vertebrates, consisting of the sensory organs called **neuromasts**, deposited along the antero-posterior body-axis, and is used to detect water motion. The posterior lateral line primordium migrates along the trunk, depositing neuromasts in a stereotypical pattern (Thomas et al. 2015). This is tightly regulated by signalling pathways and transcription factor networks.

From an evolutionary perspective, the lateral line represents a unique system because it is present in the aquatic vertebrates but absent in amniotes. This allows us to study regulatory evolution from three scenarios– loss of regulatory function, conservation of regulation with functional divergence and co-option of regulatory function into other tissues.

Interestingly, the hair cells of the neuromast are structurally and functionally similar to the hair cells found in the vertebrate inner ear (Teresa Nicolson 2005; T. Nicolson et al. 1998). In the inner ear, the hair cell detect the motion of endolymphatic fluid and in semicircular canal sand otolithic organs, enabling the perception of balance and sound. Both these sensory organs rely on mechanotransduction from hair cells that convert motion into electrical signals. Their similar structural and functional architecture reflects their shared developmental origin from cranial placodes and also suggests a deep evolutionary relationship between the two mechanosensory systems

Since the gene content utilised for sensory placode development remains largely conserved across vertebrates (Patthey et al. 2014), the presence or absence of the lateral line is unlikely to lead to a wholesale loss of genes. Thus difference in regulatory deployment becomes more likely, with enhancers active in the zebrafish lateral line representing regulatory modules

either lost or repurposed in amniotes. Studying this can offer us insights into remodelling of the regulatory landscape during macroevolutionary transitions.

8. Evolutionary Scenarios for Conserved Lateral Line Enhancers

The ATAC-seq data generated in our laboratory revealed 72 candidate enhancers of the lateral line of the zebrafish, which were conserved in mouse. This leads to an evolutionary paradox as lateral line is an aquatic sensory system, absent in amniotes, yet the regulatory sequence is maintained across 450 million years of evolutionary distance, suggesting that there might be either functional constraint due to the regulatory reuse. Several non-mutually exclusive scenarios could explain this.

One of the possibilities is **functional conservation in homologous or partially homologous contexts**. Even though the lateral line is lost in terrestrial vertebrates, there are aspects of sensory placode development which remain deeply conserved across vertebrates.

Developmental regulators responsible for specification, migration and differentiation remain conserved in the mouse. It makes for a plausible explanation that conserved enhancers regulate genes within related cranial sensory territories, for example, the otic placodes (Venkataraman et al. 2025). In this scenario, sequence conservation could mean preservation of ancestral regulatory function despite morphological divergence.

Alternatively, a particularly compelling possibility is sequence conservation due to regulatory divergence by **enhancer co-option** (Cvekl and Piatigorsky 1996; Rebeiz et al. 2011), where the regulatory competence of the enhancer is retained, but it drives expression in different/novel tissues in mouse. The advantage of this evolutionary strategy is that since transcription factor binding motifs are already present within accessible regions in chromatin, they become amenable substrates for redeployment, as the developmental and ecological contexts shift. Even slight changes in transcription factor availability, motif affinity, chromatin accessibility and enhancer-promoter contexts can shift the spatial output even though the core sequence features remain preserved.

A third explanation is that there is an **architectural constrain** within conserved regulatory landscapes (Ohmori 2026). Vertebrate genomes are arranged into regulatory neighbourhoods called topologically associated domains (TADs), where many are evolutionarily conserved. If some of the conserved elements in the list of 72 candidates reside within such regulatory neighbourhoods which are syntenically preserved in mouse, their retention reflects constraints imposed by the 3D chromatin organisation rather than the strict conservation of enhancer activity.

To be able to distinguish between these scenarios requires assessment of function, as one cannot solely rely on comparative sequence analysis alone. Following identification of the 72 conserved elements, functional assessment of their activity as enhancers is necessary, to be able to further assess the evolutionary trajectories of enhancers. These cohort of elements with a potential function in sensory structures prevents us from relying on isolated examples to study regulatory repurposing. This study will allow us to evaluate whether loss, maintenance, divergence or co-option are the predominant outcomes after the morphological loss during evolution.

9. Research Question and Hypothesis

The ATAC-seq profiling revealed 72 candidate enhancers of the lateral line of the zebrafish which have a conserved counterpart in the mouse genome. The fundamental question that arises is: **why are the zebrafish lateral line-specific candidate enhancers conserved in an organism which has lost the organ system?** We hypothesise that a subset of the conserved candidate enhancers have undergone **regulatory co-option** in vertebrates.

Sequence conservation could mean functional constraint, but not necessarily conservation of the regulatory output. If our hypothesis is correct, functional assays of the orthologous mouse sequences corresponding to those candidate elements that are functionally validated as active enhancers in the zebrafish should reveal reproducible enhancer activity *in vivo*. If we see enhancer activity in new spatial domains, it may correspond to tissues sharing components of ancestral regulatory elements or integration into novel mammalian developmental programs. Conversely, if enhancer loss is prevalent, most elements would lack activity, while, if strict conservation of activity predominates, we will find that the expression of the orthologous enhancers is driven in homologous territories.

10. Scope of the Study

The primary objective is to assess co-option of enhancer activity in the conserved zebrafish lateral line enhancers. The scope of this study operationally defines repurposing as a change in the regulatory potential of enhancers across species, rather than validation of direct target genes. Shift in enhancer deployment at the level of tissue or cell-type still provides for a tractable and informative readout for evolutionary divergence (Long et al. 2016).

The starting point of this study is with the identification of 72 candidate enhancer elements identified by ATAC-seq profiling of the zebrafish lateral line. Since chromatin accessibility does not necessarily equate to enhancer function, the first step is to validate the candidate elements as functional enhancers by conducting enhancer-reporter assays. We use the **Tol2 transgenesis system** as it allows for a successful and efficient enhancer-reporter integration (Kawakami 2007). In this operational framework, enhancer activity is defined by the ability to drive detectable reporter expression.

To get insights into the potential regulatory targets we look at the genes lying in the same topologically-associated domain (TAD) as the candidate enhancer element. TADs represent structural units of genome organization, that constrains enhancer-promoter communication. We examine whether there is an overlap between the spatial expression of the genes located within the same regulatory neighbourhood as the enhancer, and the domains of spatial activity of the enhancer, as assessed by the reporter expression.

A central conceptual question of this study is whether the sequence alone is sufficient to encode regulatory repurposing, or change in context is necessary. Cross-species functional assays allow us to disentangle their contributions. To examine this, conserved mouse orthologous sequences are inserted into the zebrafish genome. If regulatory repurposing is encoded by the sequence, we should see an altered spatial activity of the mouse ortholog compared to the zebrafish ortholog, in the developmental context of zebrafish.

The initial functional testing of enhancer activity is carried out using Tol2-mediated transgenesis, allowing us to rapidly assess enhancer activity. To reduce variability due to random integration in the Tol2 transgenesis system and ensure reproducibility, the spatial

activity of the confirmed enhancers is assessed using the pIGLET transgenesis system (Lalonde et al. 2024) which allows for a targeted insertion in zebrafish lines with the attP landing site.

Finally, direct assessment of the activity of mouse orthologues of the conserved lateral line enhancer candidates which were functionally validated in zebrafish, is carried out by in vivo enhancer reporter assay in mouse.

Chapter 2 Materials and Methods

11. Animal Model and Maintenance

The model aquatic invertebrate used for this study is *Danio rerio* (zebrafish). They were maintained at the temperature conditions of 28°C. Wild-type fish used were TU. The various transgenic and wild-type strains used in this study were housed and cared for by the Aquatics Facility at the Stowers Institute for Medical Research. Embryos and larvae were raised in the 0.5X E2 (ZIRC) Embryo Media. The night/light cycle length is kept at 11hrs/13hrs, respectively.

The model terrestrial vertebrate used for this study was *Mus musculus* (mouse). All mouse experiments were carried at the Stowers Institute for Medical Research in an AAALAC-accredited animal facility. Animal care and housing was provided by the Stowers Institute Rodent Department, following the Guide for the Care and Use of Laboratory Animals and in accordance with the Animal Welfare Act. C57BL/6J mice were used for all experiments and maintained in an in-house breeding colony. Mice were provided ad libitum access to water and the standard laboratory chow. The set point of the rodent macroenvironment was maintained at 72 °F, with a 12-hour light/dark cycle.

12. Genomic Analysis

Putative enhancer candidates were previously identified by analyzing ATAC-seq data from the zebrafish lateral line mapped on danRer11. Putative enhancer candidates were considered conserved if they could be lifted to the mouse genome (mm10) with the liftover tool from the UCSC genome browser, using the default threshold for sequence similarity (10%) [find on the website how to cite it]. We used the Integrative Genome Viewer (IGV) to visualize conserved regions, coordinates of the candidate enhancers and their surrounding genes (Thorvaldsdottir et al. 2013) and pyGenomeTracks (Lopez-Delisle et al. 2021) for HiC contact maps for our region of interest. To identify putative target genes for each enhancer, we plotted the expression of the surrounding genes by cell types and tissues using either the lateral line scRNA-seq data of the laboratory [unpublished] or the Daniocell Seurat object (Sur et al. 2023). The lateral line scRNA-seq data was processed using standard Seurat workflows, including quality control, normalization, dimensionality reduction, clustering, and visualization.

13. Production of Transgenic Zebrafish Lines

All injections were carried out using the Tol2 and pIGLET transgenesis systems. Microinjection needles were prepared using borosilicate glass capillary tubes (Cat. No. GC100TF-10 (filamented), on a Model P-97 Flaming/Brown Micropipette Puller from Sutter Instrument Co. with the following parameters: heat = 533, pull = 150, velocity = 80, time = 30. Injection dishes were prepared using 2% agarose in E2 embryo medium. Injection mixes were prepared fresh prior to microinjection. We aimed the plasmid concentration at 15 ng/μL. Additionally, the transposase mRNA, Nuclease Free Water (NFW) and Phenol Red (as a visual tracer) were added to the injection mix. We aimed at performing microinjections at the 1-cell stage zebrafish embryos produced by a wild-type TU in-cross.

14. Production of Transgenic Mice

3-4 week old C57BL/6J females were given hormonal treatment for superovulation to generate fertilized mouse embryos. Females received pregnant mare serum gonadotropin (PMSG) via subcutaneous injection at noon, followed by human chorionic gonadotropin (HCG) administration 48 hours later, intraperitoneally to induce ovulation. Upon HCG administration, the females were paired with C57BL/6J stud males overnight. The next morning, females were checked for copulation plugs and were used for embryo collection. Once embryos were harvested, 25 fertilized embryos were placed in 25 μ L drops of adeno-assisted virus (AAV) from Vector Builder. AAV preparations were supplied in 25 μ L aliquots and diluted in Advanced KSOM medium to a final concentration ranging from 1.36×10^9 to 2.81×10^{10} viral genomes, depending on the lot titre. Embryos were co-cultured for 4-5 hours, then washed through before electroporation. After washing, the embryos are electroporated with the mix provided by the Genome Engineering Core, using the Nepa Gene electroporator.

After electroporation, embryos were transferred into pseudopregnant recipient females. These females carried the embryos to term and gave birth to pups which were weaned and sequenced by GE to find Founders.

15. Fixation, DAPI Staining, and Mounting of Embryos for Imaging

Transgenic zebrafish embryos expressing GFP were fixed in 4% paraformaldehyde (PFA) containing 0.1% Tween-20 (referred to as PFAT). The PFA stock aliquots were stored at -20°C until use.

Embryos were dechorionated manually, if they had to be fixed at a stage $<2\text{dpf}$.

Embryos were fixed in PFAT overnight, at 4°C on a gentle rocking platform. The following day, PFA was removed, and embryos were washed and transferred into $1\times$ phosphate-buffered saline with 0.1% Tween-20 (PBST).

DAPI (4',6-diamidino-2-phenylindole) was used at a 1:1000 dilution in PBST for nuclear staining. The embryos were incubated with DAPI overnight at 4°C , although a 1-hour incubation at room temperature can also be used.

16. Genotyping Zebrafish pIGLET Lines

Genotyping was performed on whole zebrafish embryos to confirm the enhancer-reporter construct insertion at the attP landing site. The presence of attP site was confirmed by DNA extraction from fin clips or sperms of zebrafish adults. The genomic DNA was isolated in 100 μ L of classic lysis buffer with 2 μ L of 50X Proteinase K per sample. Samples were incubated at 55°C for 15–24 hours to ensure complete digestion, followed by enzyme inactivation at 98°C for 10 minutes in a thermocycler. Samples were subsequently cooled at 4°C to be used as DNA template for PCR.

PCR amplification was carried out using GoTaq DNA polymerase (Promega), following the thermocycler instructions mentioned in the manual. Primer stocks were maintained at 100 μM and diluted to a 10 μM working concentration prior to use.

17. Mouse Genotyping

Mouse genotyping was performed on ear clips from an adult mouse. The ear-clip tissue was incubated in 50 μ L QuickExtract™ Lysis Buffer for DNA extraction at 65°C for 2215 minutes, followed by enzyme inactivation at 98°C for 2 minutes. The resulting lysate was then used as a DNA template for PCR.

The amplification reactions were performed using TaKaRa LA Taq DNA polymerase according to the recommendations in the manufacturer's manual. Each 21 μ L reaction contained LA Taq polymerase, 10 $\times$ LA PCR buffer, 2.5 mM dNTPs, forward and reverse primers (10 μ M working concentration), nuclease-free water, and 1 μ L genomic DNA template.

18. Immunofluorescence Staining and Histological Analysis

Mouse embryos were fixed at embryonic day 10.5 (E10.5) in 4% paraformaldehyde (PFA) for 1.5 hours on ice. The mouse embryos were cryo-embedded and sectioned using a cryostat model Thermo Scientific Cryostar NX70. 10 μ m thick sections were created and collected on glass slides and stored at -20°C. We used High-Profile Disposable Blades DB80HS for sectioning the embryos. The embryos were embedded in O.C.T embedding medium. The sectioning was performed at -11°C while the specimen temperature was maintained at -10°C.

To locate GFP signal specifically, we performed immunofluorescence staining. All the washing steps were carried out in phosphate-buffer saline (PBS) or PBS supplemented with 0.1% tween (PBST). Antigen retrieval was carried out using PBST for 30 minutes to facilitate the penetration of the antibody. To block non-specific binding, we incubated the sections in 5% normal donkey serum in PBST for 1 hour at room temperature. Primary antibody incubation was performed overnight at 4°C using anti-GFP antibody diluted in 1% serum in PBST. This was followed by 3 washes of 5 minutes each, in PBST. We next carried out secondary antibody incubation using Alexa Fluor 647–conjugated secondary antibody diluted in PBST. This was chosen to be able to avoid autofluorescence from the surrounding tissue, commonly observed in the 488 nm channel. DAPI was used for nuclear staining.

19. Hybridization Chain Reaction (HCR)

Standard HCR Gold protocol was followed. Fixed embryos were dehydrated in 100% methanol at -20 °C for at least overnight and up to several weeks. Before hybridization, samples were rehydrated stepwise through 66% and 33% methanol containing 0.1% Tween-20 and finally rinsed in ultrapure H₂O + 0.1% Tween-20. 80% acetone was used for embryo permeabilization, with an incubation time of 30 minutes at -20 °C, then washed twice for 5 minutes in PBST (1 $\times$ PBS + 0.1% Tween-20) at room temperature. A 1:1000 DAPI in 5x SSCT wash is done 2x for 60 minutes each.

20. Microscopy and Image Acquisition

All the zebrafish embryos were imaged using Nikon Eclipse Ti2 spinning disk confocal microscope at 10X and 40X magnifications. All samples were embedded in 0.8% Low Melting Point (LMP) agarose was prepared in embryonic E2 media on a 30 mm petri dish. LMP agarose aliquots of 1 mL were kept at -20°C and were heated at 80°C for at least 5 minutes, then transferred to 43°C prior to embedding. Larvae were mounted laterally and ventrally. For imaging of more internal tissues, their jaw was dissected, after which the embryos were mounted laterally.

All mouse samples were acquired on a Nikon Eclipse Ti2 microscope equipped with a Yokagawa CSU W1 10,000 rpm Spinning Disk Confocal with 50um pinholes.

23 Samples were illuminated with 488nm(8.5mw) and 405nm(3.9mw) lasers. A Nikon Plan Apochromat Lambda 10X objective lens and Plan Apochromat Lambda S 40X C Silicone objective lens was used to acquire the image.

Similar acquisition parameters were used for comparable samples. The images were processed using FIJI-ImageJ software (Schindelin et al. 2012).

Chapter 3 Results

21. Establishment of Stable Enhancer Transgenic Lines

To functionally validate the enhancer activity of the candidate enhancers, we generated stable zebrafish transgenic lines using both Tol2 (Kawakami 2007) and pIGLET transgenesis systems (Lalonde et al. 2024). While both methods lead to successful genomic integration events, they differ fundamentally in their genomic integration strategy. The Tol2 system leads to random genomic insertion, possibly at multiple locations while the pIGLET system as a dedicated attP landing site, allowing us to compare activity across different constructs.

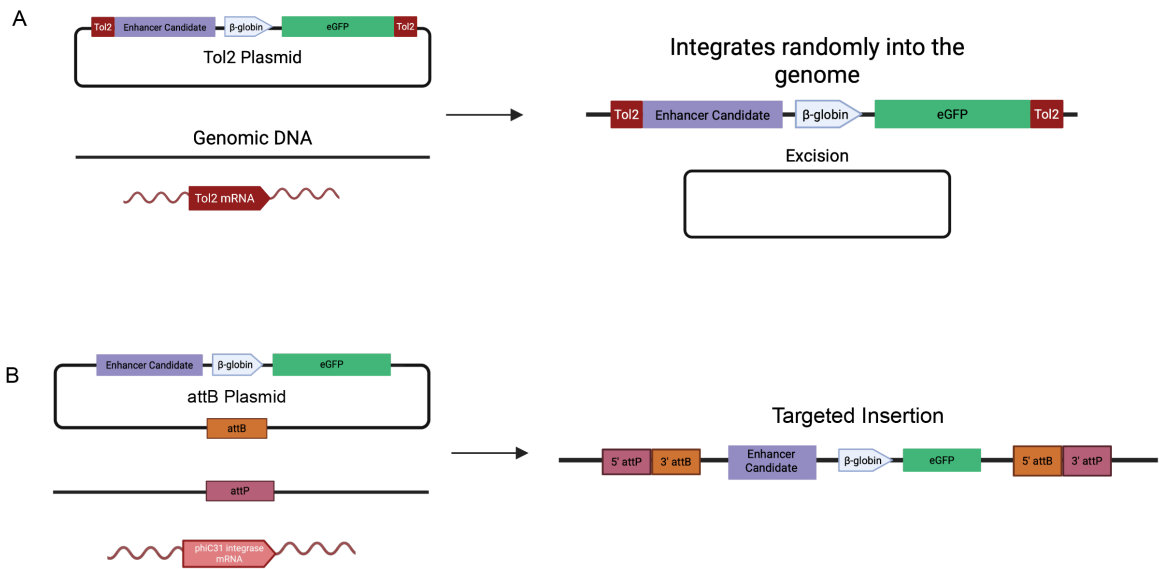


Figure 6 Comparison of Tol2-mediated random transgenesis and pIGLET-mediated targeted insertion in zebrafish.

(A) Tol2 transgenesis. The enhancer-reporter construct, consisting of the candidate enhancer upstream of a minimal β -globin promoter driving eGFP, is flanked by Tol2 inverted terminal repeats (ITRs). Upon co-injection with Tol2 transposase mRNA, the transposase mediates excision of the transgene cassette from the plasmid backbone and its integration into the zebrafish genome. Integration occurs at random genomic locations (B) pIGLET-mediated targeted insertion. The enhancer-reporter cassette is cloned into an attB-containing plasmid. In embryos carrying a genomic attP landing site, co-injection of ϕ C31 integrase mRNA enables recombination between attB and attP sites, resulting in site-specific integration of the construct. This targeted insertion minimizes position-effect variegation and allows direct comparison of enhancer activity across constructs.

Following microinjection in wild-type TU embryos at the one-cell stage, the embryos showing GFP expression between 1 day post fertilization (dpf) to 5dpf were raised to maturity, as F0s to be screened. Each F0 parent is mosaic since their germ cells are not necessarily derived from a single integration event. To resolve this mosaicism, the F0 founders are outcrossed with wild-type TU to give rise to the F1 generation. The F0 parents that are able to transmit the transgene are the “Founder” parents. The F1 progeny arising from F0 founders possess a homogenous germ cell population because each F1 fish arises from a single germ cell of the F0 parent.

The F1 progeny from each F0 founder was screened for reporter expression. However, the F1 progeny derived from the same F0 founder, may still carry multiple insertions in the Tol2 transgenesis system. To establish single insertion lines, we outcrossed the F1 individuals to the wild-type TU fish and evaluated their transmission ratios in the F2 generation. A single Mendelian insertion is expected to yield approximately 50% reporter-positive embryos upon outcrossing to wild type.

In cases where the F2 generation produced ~50% GFP positive embryos with identical expression patterns, it indicated a stable, single insertion and hence such lines were said to be established. These established lines were used for further downstream imaging and analysis. Conversely, if an F1 parent produced >50% reporter expressing F2 generation, it suggested multiple insertions. The F2 individuals from such individuals were raised to maturity and subsequently outcrossed to assess transmission ratios. They were sequentially outcrossed until a stable ~50% positive Mendelian segregation pattern was obtained, at which point the line was designated as an established transgenic strain.

Since the pIGLET transgenesis system enables targeted genomic integration, a successful integration event was validated by genotyping, using locus specific primers in lines producing approximately 50% reporter-positive progeny.

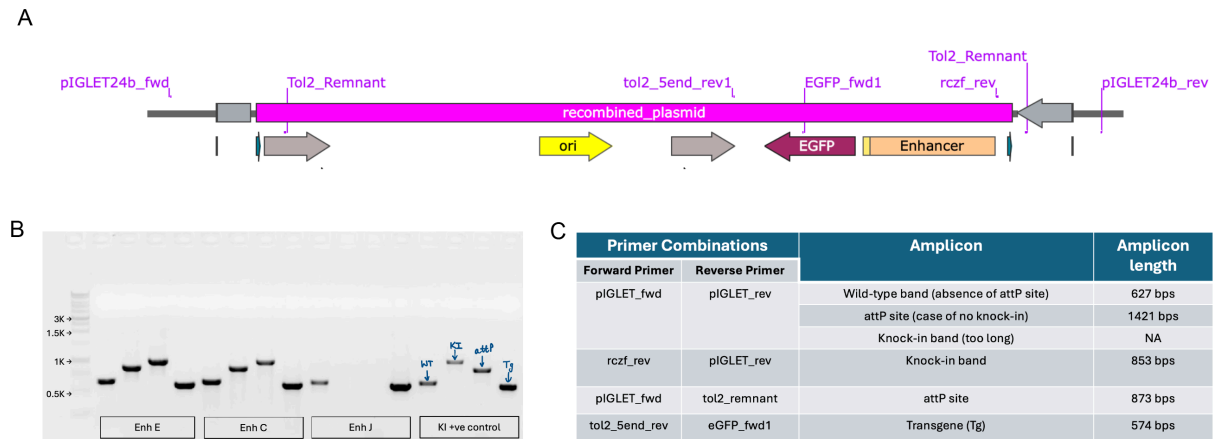


Figure 7 Genotyping strategy and validation of pIGLET-mediated targeted insertions

(A) Schematic representation of the recombined plasmid following ϕC31-mediated recombination at the genomic attP landing site. Primer binding sites used for genotyping are indicated in purple.

(B) Representative agarose gel images showing PCR products for enhancer constructs (enhE, enhC, enhJ) and KI+ control. For enhE and enhC, the WT band (~627 bp) confirms heterozygosity, absence of the attP band (~1421 bp) indicates recombination, the knock-in band (~853 bp) confirms targeted insertion, and the transgene band (~574 bp) verifies presence of the enhancer-reporter cassette. For enhJ, WT and transgene bands are detected, but the knock-in band is absent, indicating integration outside the attP landing site.

(C) Table summarizing primer combinations, expected amplicons, and fragment sizes used for genotyping targeted insertion events.

To validate the site-specific insertion of the enhancer-reporter constructs at the attP site in the pIGLET zebrafish lines, we carried out PCR genotyping using primer combinations that spanned the attP locus, recombination junctions and the transgene cassette (Fig. 7A). In the representative gel image shown above (Fig. 7B), each enhancer has four lanes dedicated to the 4 primer combinations used for this validation. The primer combinations used for genotyping appear in the same serial as shown in the table (Fig. 7C). For the enhE and enhC blocks, the first lane shows the amplification of the wild-type (WT) genomic allele (~627 bp), which is consistent with the heterozygosity expected for the attP site. As expected, we don't see any intact band for the attP site itself because the site has undergone plasmid recombination. The second lane displays the knock-in band (~853 bp), confirming correct recombination at the downstream junction of the attP site. The band in the third lane demonstrated correct integrations at the upstream recombination junction. Both the junction-specific bands together indicate that the enhancer-reporter cassette is correctly integrated, intact, and in the proper orientation at the attP locus. The fourth lane confirms the presence of the enhancer-reporter transgene (~574 bps). Together, these results demonstrate precise site-specific insertion of enhE and enhC constructs at the attP landing site.

In contrast, for **enhJ**, the WT and transgene bands were detected, but the bands showing knock-in at the correct genomic locus were absent. Furthermore, there is no band corresponding to the intact attP site (~ 1421 bps if the plasmid has not recombined at this locus), indicating the absence of the attP site altogether. The presence of the transgenic band implies that genomic integration of the enhJ-reporter construct has occurred randomly. Thus, genotyping confirmed that enhE and enhC represent correctly targeted insertions suitable for line establishment, whereas enhJ reflects an insertion event outside the intended landing site.

Such an analysis was done for the remaining enhancer constructs as well, followed by confocal imaging to assess their spatial activity.

22. Zebrafish enhancers drive reporter activity in the lateral line and inner ear

To functionally validate the activity of the candidate elements and assess their spatial activity, we carried out confocal imaging on the transgenic lines explained in the results section 21. The transgenic lines contain sequences of the candidate enhancer clones upstream of a minimal promoter driving a green fluorescent protein reporter. We analysed reporter expression during different larval stages to determine whether these elements can drive tissue-specific transcriptional activity.

Out of the 72 enhancer candidates, 12 were able to drive the reporter expression. They exhibited distinct but partially overlapping activity patterns within the components of the lateral line and the inner ear. Representative examples of three enhancers—EnhA, EnhF, and EnhE—are shown (Fig. 8).

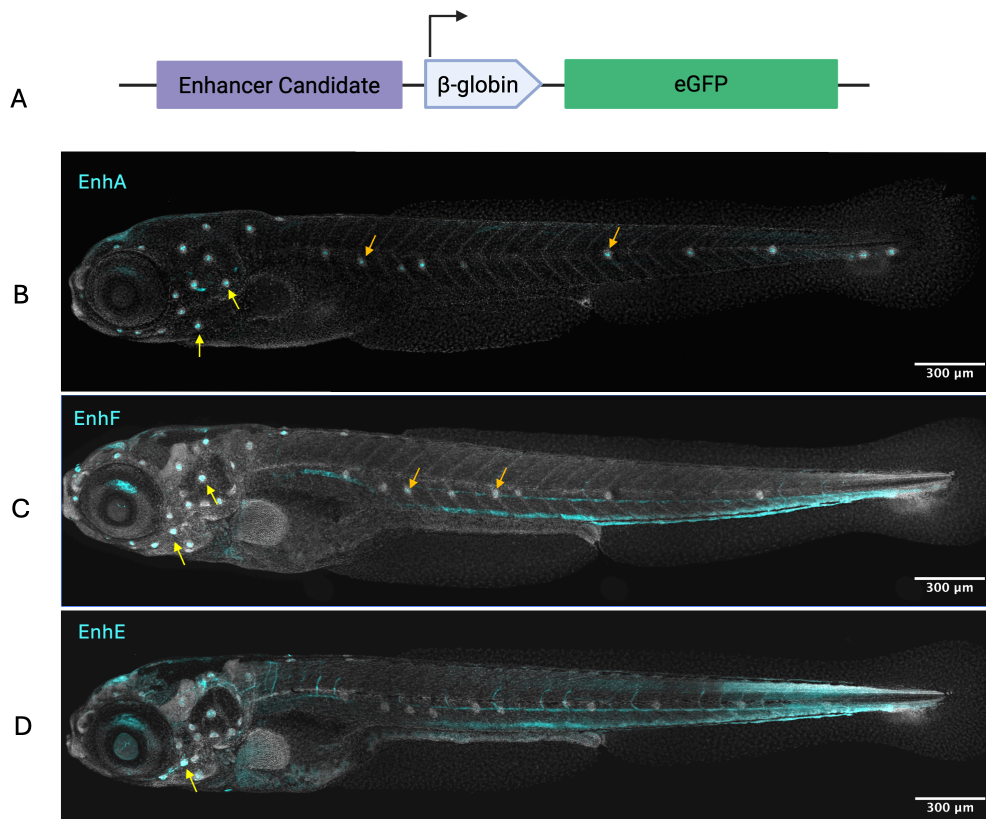


Figure 8: Zebrafish orthologous enhancers drive reporter expression in the lateral line system.

(A) Schematic representation of the enhancer–reporter construct used in the *in vivo* assay. Candidate regulatory sequences were cloned upstream of a minimal β-globin promoter driving eGFP expression. Reporter activity reflects the spatial regulatory activity of the tested enhancer sequence. (B–D) Representative confocal images showing reporter expression driven by three zebrafish candidate enhancers in transgenic embryos. Endogenous GFP reporter signal is shown in cyan with

nuclei counterstained in grey. (B) EnhA drives expression in both the anterior lateral line (ALL) and posterior lateral line (PLL) neuromasts, as well as in inner ear sensory structures. (C) EnhF drives strong expression throughout the lateral line, including along the trunk and tail neuromasts. (D) EnhE shows a more restricted pattern, with reporter activity primarily in anterior lateral line neuromasts. Yellow arrows indicate representative neuromasts along the anterior lateral line. Orange arrows indicate representative neuromasts along the posterior lateral line. Scale bars: 300 μ m.

EnhA drives reporter activity in both the anterior lateral line (ALL) and posterior lateral line (PLL) neuromasts, as well as in inner ear sensory structures (Fig. 8A). EnhF shows activity along the entire lateral line as well (Fig. 8B). In contrast, we see a more restricted pattern in case of EnhE, with activity in the ALL and neurons.

Taking a closer look at the head region reveals that some enhancers, including EnhA are active in the **inner-ear sensory organs** as well, including the anterior, lateral, and posterior cristae, as well as the utricle and saccule (Fig. 9B).

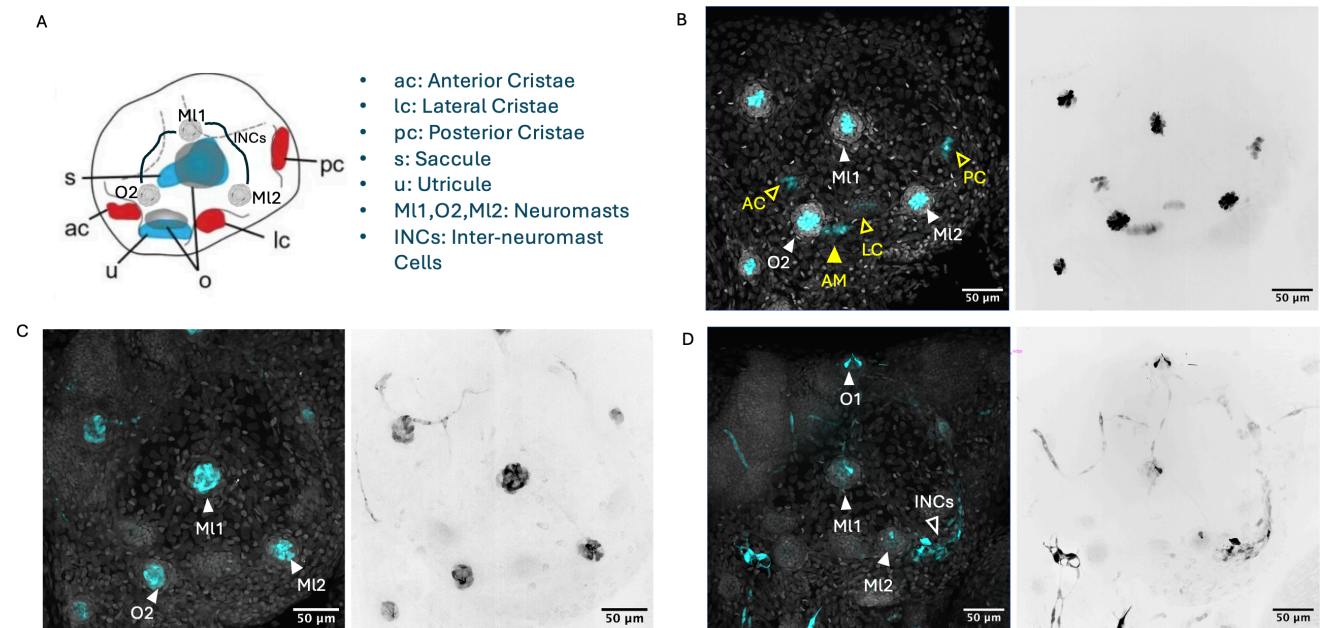


Figure 9: Candidate enhancers drive activity in inner ear sensory organs and anterior lateral line neuromasts.

(A) Schematic representation of the zebrafish inner ear highlighting the locations of sensory structures analysed. ac: anterior cristae; lc: lateral cristae; pc: posterior cristae; s: saccule; u: utricle; MI1, MI2, O2: anterior lateral line neuromasts; INCs: inter-neuromast cells. (B–D) Confocal images showing enhancer-driven reporter expression in neuromasts and inner ear. Reporter signal (cyan) is overlaid with nuclear staining (grey). (B) Reporter expression in case of EnhA is observed in anterior lateral line neuromasts (MI1, MI2, O2 indicated by white arrowheads for representation), anterior, lateral and medial cristae (AC, LC and PC) and Maculae (AM) (C) EnhF drive expression in the neuromasts of the anterior lateral line (MI1, MI2 and O2 represented by white arrowheads) (D) EnhE drives reporter expression in anterior lateral line neuromasts (MI1, MI2, O2) and inter-neuromast cells. Right panels show the inverted reporter channel alone for clarity. Scale bars: 50 μ m.

The observation that multiple candidate enhancers drive reporter expression in the lateral line is consistent with the ATAC-seq peaks from the lateral line tissue, supporting the conclusion that these elements might be active and functional enhancers in the lateral line. Interestingly, several of these enhancer candidates are also active in the inner-ear sensory system, which shares structural and functional similarity with the neuromasts of the lateral line system. Both systems rely on fluid movement for detection by the sensory hair cells, indicating that these enhancers may regulate genes involved in shared sensory cell programs.

Together, these results suggest that a subset of the candidate enhancer elements identified through profiling of chromatin accessibility might function as enhancers *in vivo*, driving transcription within the sensory structures of zebrafish lateral line and inner ear.

23. Characterization of EnhA Activity and Candidate Target Genes

Following the validation of the candidate elements as active enhancers, we sought to assess their functional role, by identifying the genes within the same regulatory neighbourhood as our validated enhancers. Ideally, functional characterization would also involve mutational perturbation studies and see their effect on the expression of target genes. However, such experiments were beyond the scope of the study. Instead, we utilized, Hi-C to examine the regulatory neighbourhood of the enhancers and examine the cellular contexts in which potential target genes are expressed by studying single-cell datasets and in-situ hybridization.

Here, we illustrate this strategy using EnhA as an example to identify the potential target genes and gain insight into the regulatory context of this enhancer.

1. Temporal and Spatial Activity of EnhA During Development

The reporter line Tg(enhA:β-globin:eGFP) was imaged at different developmental stages. The first GFP signal was detected at the 17-somite stage (< 1 day post fertilization (dpf)) kept at 23°C to slow down the embryo's development. The enhancer activity could be seen in the olfactory epithelium, marking the onset of the activity of enhA.

We next characterized its activity at later developmental stages. By 2dpf, expression was evident in the otic vesicle and the lateral line primordium and by 5 dpf, the GFP signal could be seen very clearly in the inner ear and the neuromasts of the head and the trunk.

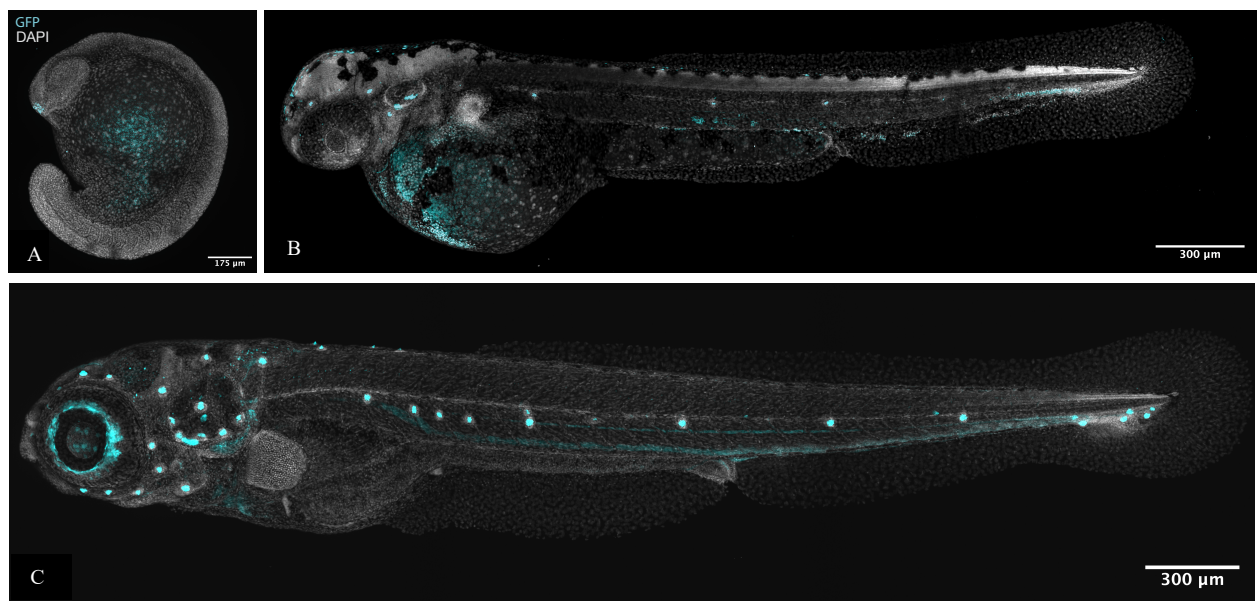


Figure 10: Spatiotemporal activity of enhA during zebrafish development.

Expression of enhA-driven GFP at various developmental stages: (A) early onset at the 17-somite stage (<1 dpf) in the olfactory epithelium (B) otic vesicle and lateral line primordium by 2 dpf, and (C) persisting in otic and neuromast structures at 5 dpf.

The spatiotemporal pattern of activity seen in enhA is consistent with the role of a developmental enhancer, becoming active early on, in the mechanosensory organ, and driving expression all throughout the larval stages.

2. Identification of Putative Targets using Using Hi-C Data

To identify the potential gene targets of these enhancers, which might lie in the same regulatory domain associated with enhA, enhB and enhC, Hi-C chromatin contact maps were analyzed (Franke et al., 2021; Wike et al., 2021; Hintermann, Bolt et al., 2024). The data

were visualized using pyGenomeTracks, which allowed the mapping of chromatin interaction domains along zebrafish chromosome 20, where the three enhancers are located.

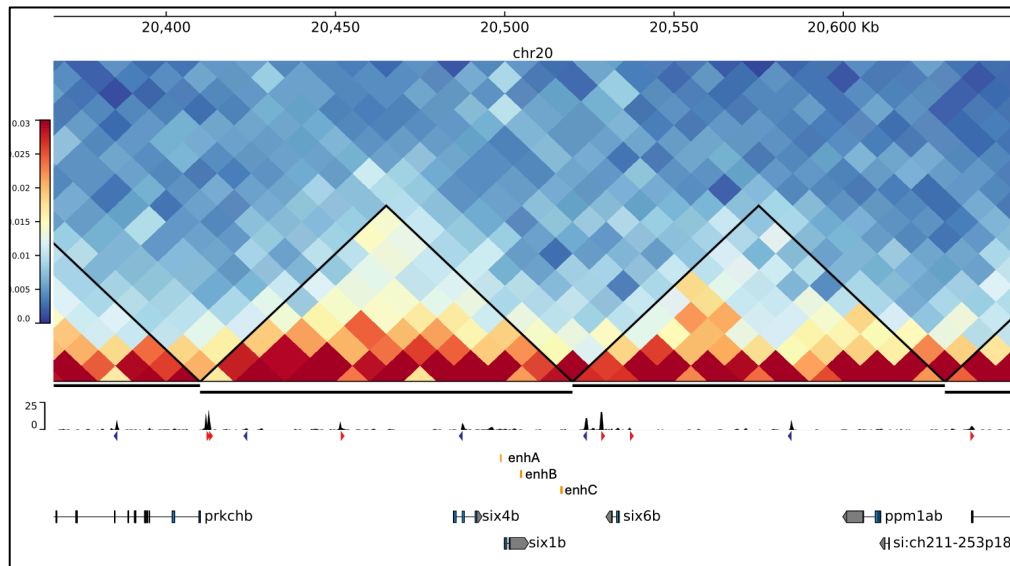


Figure 11: Hi-C chromatin contact maps showing regulatory domains of enhA

Hi-C interaction data along zebrafish chromosome 20, where enhA are located. The three enhancers lie within a single topologically associating domain (TAD) containing *prkchb*, *six4b*, *six6b*, *six1b*, and *ppmlab*, indicating these genes as potential regulatory targets within the same chromatin domain.

The Hi-C map revealed that genes from *prkchb* to *ppmlab* lie in the same, single **topologically associating domain (TAD)** as enhA, enhB and enhC. Because enhancers typically regulate genes within the same TAD, the genes in the boundary being ***prkchb*, *six4b*, *six6b*, *six1b*, and *ppmlab*** become the likely regulatory targets of these enhancers.

To refine these predictions further, single-cell expression profiles from the Daniocell atlas were analyzed.

3. Single-Cell Transcriptomic Analysis Using Daniocell

To assess the expression profiles of the likely targets of our enhancers from the Hi-C data, single-cell RNA-Seq data from the Daniocell zebrafish (Sur et al., 2023) were examined.

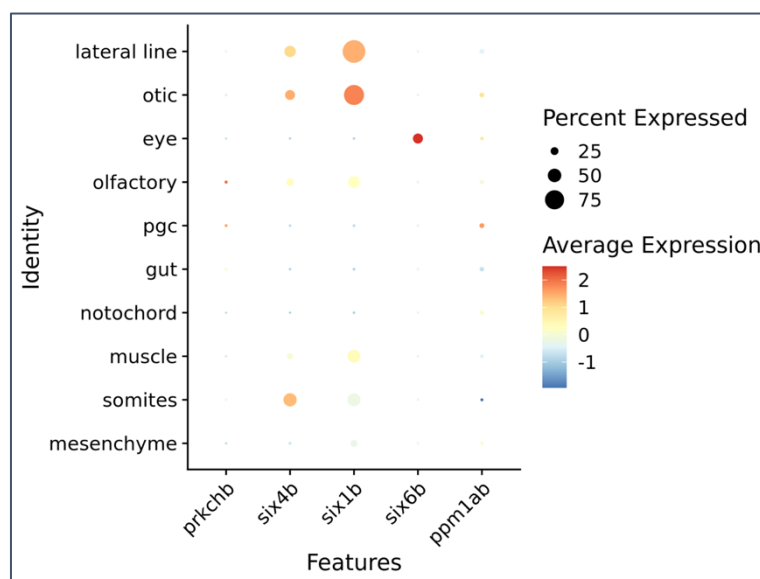


Figure 12: Single-Cell Daniocell dot plot shows that genes *six4b* and *six1b* are most highly expressed in the otic region and the lateral line, which overlaps with region of activity of enhA, making them the most likely regulatory targets

The dot-plots showed that the regions of expression of genes *prkchb*, *six6b* and *ppmlab* do not overlap with the regions where EnhA shows activity. This makes them unlikely regulatory targets of our enhancers. On the other hand, genes *six4b* and *six1b* showed strong expression in the lateral line and the otic region, which makes them the most likely transcriptional targets of these enhancers.

4. Validation of Enhancer–Gene Relationships Using HCR

To test these predictions experimentally, Hybridization Chain Reaction (HCR) was performed using probes specific for *six1b* and *six4b*.

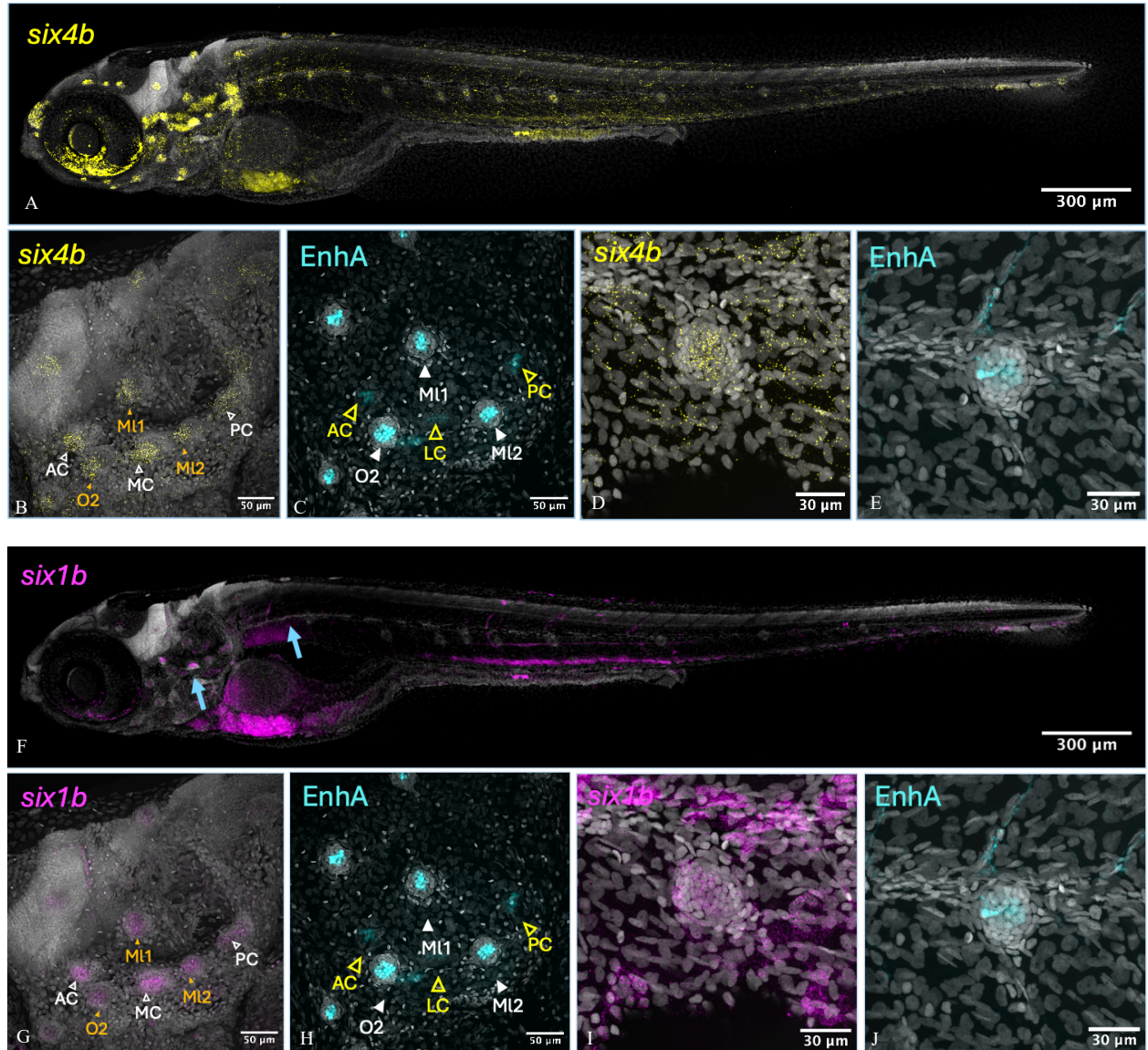


Figure 13: Validation of enhancer–gene relationships using HCR.

Fluorescent HCR in situ hybridization for (A–10X whole fish B–otic region, D–L1 neuromast) *six4b* and *six1b* (F–10X whole fish G–otic region, I–L1 neuromast) transcripts in zebrafish larvae. Confocal images show strong expression in otic sensory patches and neuromasts, overlapping with enhancer-driven GFP activity, confirming that *enhA* likely regulates *six1b* and *six4b* in mechanosensory structures.

The results showed that *six1b* and *six4b* transcripts were strongly expressed in neuromasts and otic sensory structures, with spatial patterns that closely matched enhancer-driven GFP expression. These findings confirm that the likely regulatory targets of *EnhA* are *six4b* and

six1b and support their roles as enhancers active in the mechanosensory structures of the ear and the lateral line.

24. Comparison of Enhancer Activity between Tol2 and pIGLET Transgenic Lines

After determining the enhancers identified by Tol2 mediated transgenesis, we wanted to determine how comparable and reproducible the enhancer activity is across Tol2 and pIGLET transgenic lines. In the sections below are shown such comparisons for three enhancers—zebrafish ortholog of EnhA in zebrafish, mouse ortholog of EnhE in zebrafish and zebrafish ortholog of EnhH in zebrafish, in Tol2 and pIGLET transgenesis backgrounds. This comparison enabled us to assess what system could be used reliably to assess the spatial activity of our candidate enhancers.

1. EnhA reporter activity in Tol2 and pIGLET lines

EnhA activity in Tol2 transgenesis background was analysed using *Tg(EnhA:βglobin:eGFP)* zebrafish transgenic line. In these embryos, strong GFP reporter expression was observed along the anterior and posterior lateral line (Fig. 14A-D).

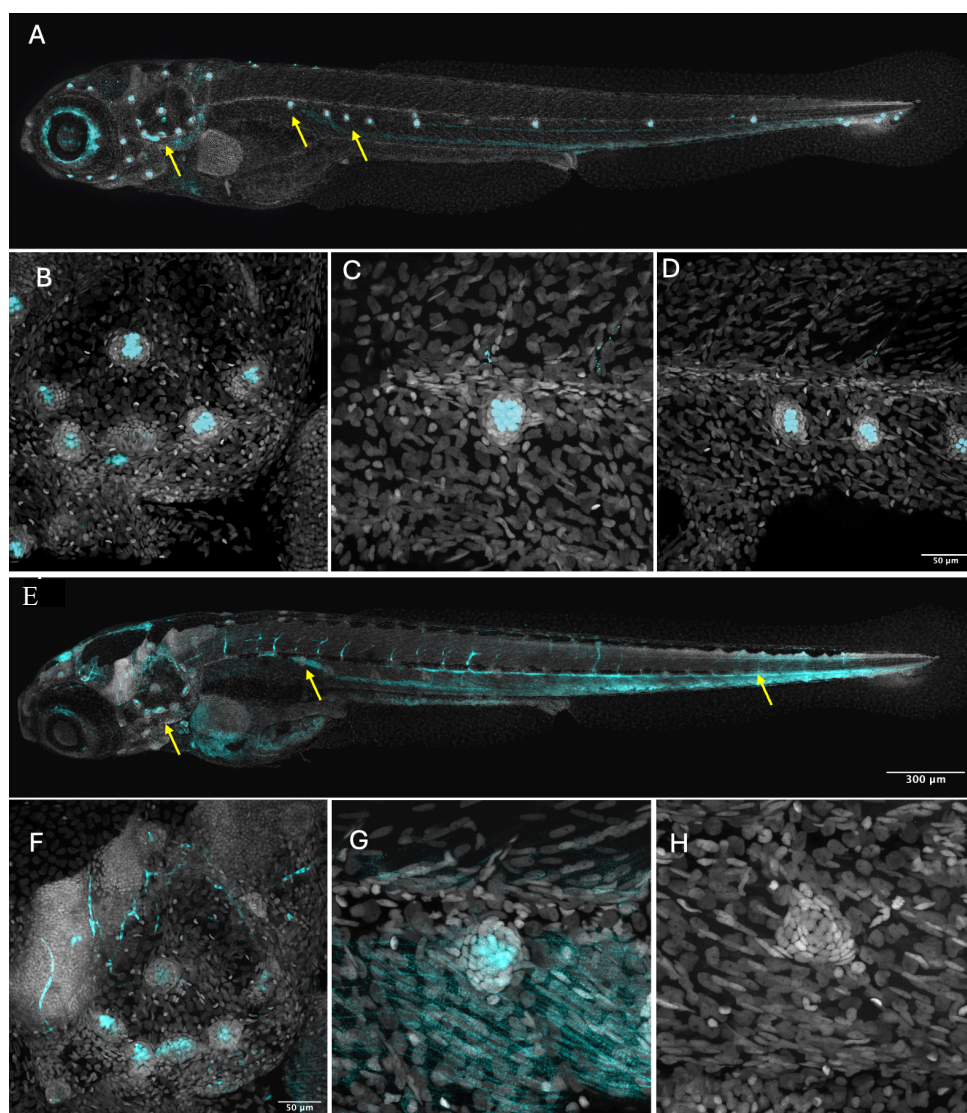


Figure 14: Comparison of EnhA-driven reporter activity in Tol2 and pIGLET transgenic zebrafish lines.

(A–D) Reporter expression in *Tg(EnhA:βglobin:eGFP)* Tol2 embryos. (A) Whole-embryo maximum projection showing strong GFP expression along the lateral line system (yellow arrows). (B) Higher magnification image of the anterior lateral line region showing multiple GFP-positive neuromasts. (C) Close-up of the L1 neuromast of the posterior lateral line displaying robust reporter activity. (D) Higher magnification image of ventral neuromasts v1.1, v1.2 and v1.3 of the

posterior lateral line. (E–H) Reporter expression in *Tg(EnhA:βglobin:eGFP)pIGLET24b* embryos. (E) Whole-embryo projection showing GFP-positive neuromasts primarily in the anterior lateral line (yellow arrows). (F) Higher magnification view of the anterior lateral line region. (G) Close-up of the L1 neuromast showing reporter activity. (H) Tail neuromast (T1) with no reporter activity. Scale bars: 300 μm (A, E) and 50 μm (B–D, F–H).

To determine whether this pattern of activity is maintained in transgenic lines created by pIGLET transgenesis system, 5dpf embryos from the *Tg(EnhA:βglobin:eGFP)pIGLET24b* transgenic line were imaged (Fig 14E–H). In contrast to the Tol2 line, the GFP intensity was much lower in the pIGLET line which made clear detection of the reporter activity difficult at 10X magnification. Higher magnification images revealed that reporter activity in the transgenic line was restricted to the anterior lateral line, with little signal in L1 neuromast and no signal in the subsequent neuromasts of the posterior lateral line.

These observations suggest that while EnhA seems to be capable of driving enhancer activity along whole of the lateral line across different Tol2 founders (see appendix), it is unable to drive reporter activity in the posterior lateral line when the enhancer-reporter construct is integrated at the attP site of pIGLET24b lines.

2. EnhE Mouse Ortholog Reporter Activity in Tol2 and pIGLET Lines

Reporter activity for the mouse ortholog of EnhE was similarly examined using *Tg(mmEnhE:βglobin:eGFP)* in Tol2 embryos. We observed strong GFP expression in the Tol2 background in some neuromasts of the anterior lateral line and posterior lateral line and the inter-neuromast cells of the anterior lateral line. (Fig. 15A–D). Additionally, we observed reporter activity in the olfactory pit. Higher magnification images confirmed activity in the posterior lateral line.

When we imaged the same enhancer construct inserted in *Tg(mmEnhE:βglobin:eGFP)pIGLET24b* lines (Fig 15E–H), the enhancer activity was again largely restricted to the anterior lateral line. Moreover, we could not detect any GFP signal in the olfactory pit unlike the Tol2 line. Higher magnification images revealed very faint GFP expression in L1 and no GFP expression in L2.

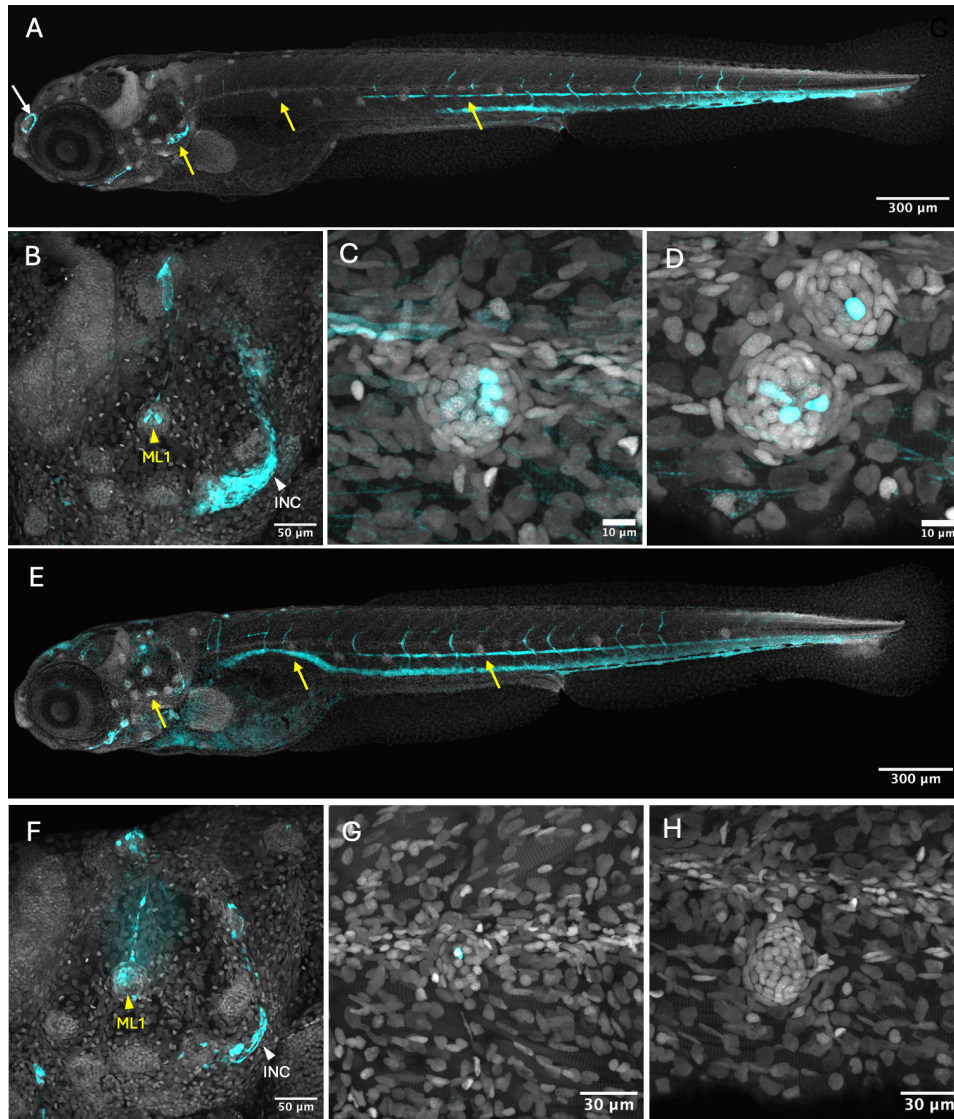


Figure 15: Comparison of EnhE-driven reporter activity in Tol2 and pIGLET transgenic zebrafish lines.

(A–D) Reporter expression in *Tg(EnhE:βglobin:eGFP)* Tol2 embryos. (A) Whole-embryo projection showing strong GFP expression along the lateral line system (yellow arrows). GFP expression is also seen on the olfactory pit (white arrow) (B) 40X image of the anterior lateral line region showing GFP-positive neuromasts. (C) 40X of the L1 neuromast of the posterior lateral line displaying reporter activity. (D) 40X of the adjacent L2 neuromast. (E–H) Reporter expression in *Tg(EnhE:betaglobin:eGFP)pIGLET24b* embryos. (E) Whole-embryo projection showing GFP-positive neuromasts primarily in the anterior lateral line (yellow arrows). (F) Higher magnification view of the anterior lateral line region. (G) 40X of the L1 neuromast of the posterior lateral line. (H) 40X of the adjacent L2 neuromast. Scale bars: 300 μm (A, E), 50 μm (B, F), and 10–30 μm (C–D, G–H).

Thus, like EnhA, the mouse ortholog of EnhE showed robust activity in the anterior and posterior lateral line in the Tol2 transgenesis background but did not drive robust reporter expression in pIGLET24b transgenic lines.

3. EnhH reporter activity in Tol2 and pIGLET Lines

We next examined the activity of EnhH in *Tg(EnhH:βglobin:eGFP)* in Tol2 transgenic embryos (Fig 16A-D). Similar to EnhA and EnhE tol2 transgenic lines, EnhH was able to drive reporter expression along the anterior and posterior lateral line. Higher magnification images of the L1 and T1 neuromasts of the posterior lateral line confirmed so as well.

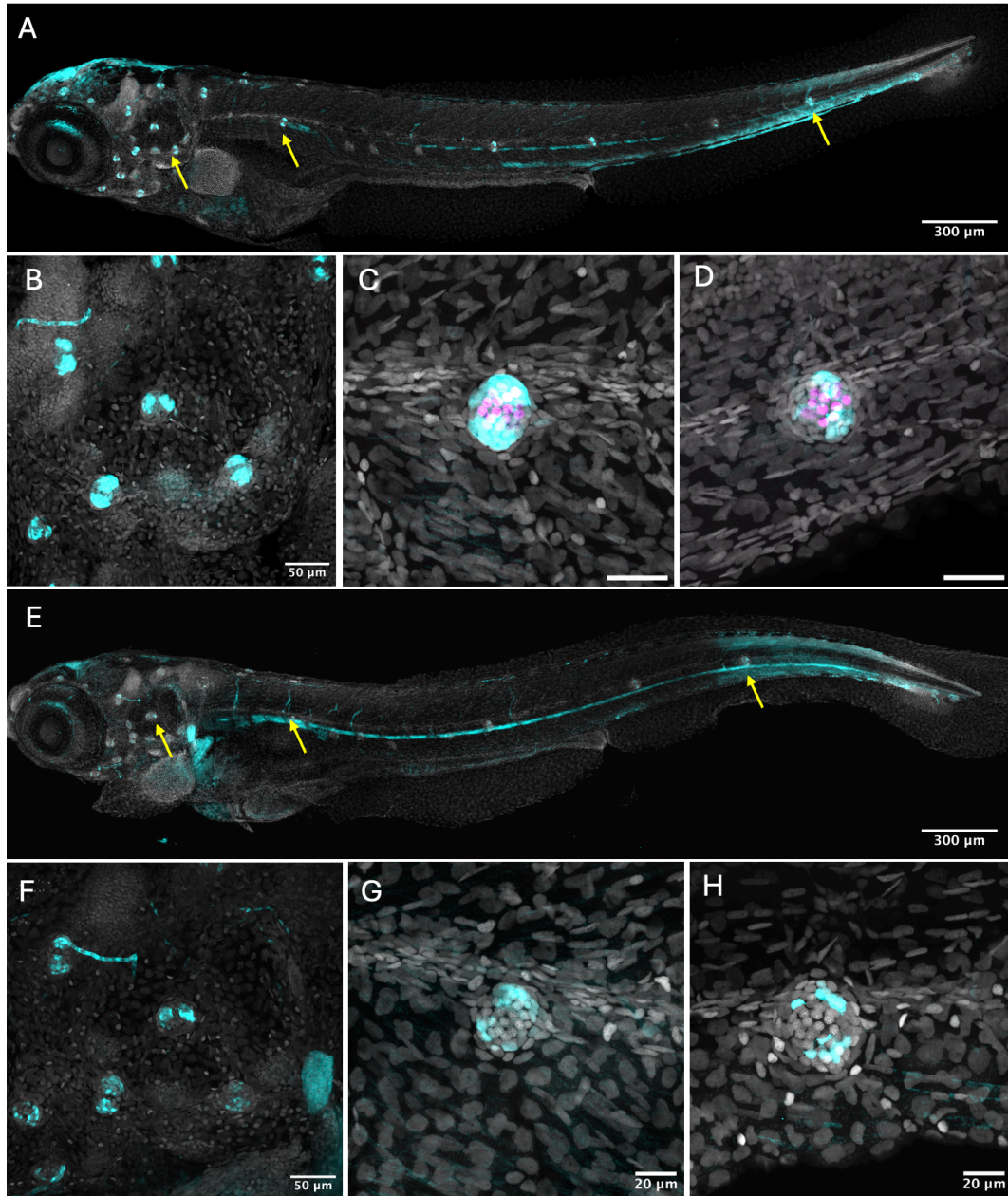


Figure 16: Comparison of EnhH-driven reporter activity in Tol2 and pIGLET transgenic zebrafish lines.

(A–D) Reporter expression in *Tg(EnhH:βglobin:eGFP)* Tol2 embryos. (A) Whole-embryo projection showing GFP expression along the lateral line system (yellow arrows). (B) Higher magnification image of the anterior lateral line region showing GFP-positive neuromasts. (C) Close-up image of a neuromast co-labeled with the hair cell marker *myo6b* (magenta), showing GFP-positive cells within the neuromast. (D) Tail neuromast (T1) displaying reporter activity. (E–H) Reporter expression in *Tg(EnhH:βglobin:eGFP)pIGLET14a* embryos. (E) Whole-embryo projection showing GFP-positive neuromasts along the lateral line (yellow arrows). (F) 40X image of neuromasts in the anterior lateral line. (G) 40X image of L1 neuromast in the posterior lateral line. (H) Tail neuromast (T1) showing GFP expression.

Scale bars: 300 μm (A, E), 50 μm (B, F), and 20 μm (C–D, G–H).

Next we examined EnhH activity in *Tg(EnhH:βglobin:eGFP)pIGLET14a* transgenic lines (Fig. 16E–H). These transgenic lines are different than the pIGLET lines of EnhA and EnhE shown in the above sections in that, the landing site occurs at a different genomic location than the pIGLET24b lines. Interestingly, in contrast to our observations seen in case of EnhA and EnhE pIGLET transgenic lines, the *Tg(EnhH:βglobin:eGFP)pIGLET14a* embryos displayed GFP expression throughout the lateral line. Higher magnification images also confirmed reporter activity in the neuromasts of both the anterior and posterior lateral line.

The spatial activity pattern of *Tg(EnhH:βglobin:eGFP)pIGLET14a* closely resembled the pattern of enhancer activity observed in the Tol2 transgenic line.

Together, these results suggest differences in the spatial activity of enhancers based on the transgenesis system and their genomic location. EnhA and EnhE displayed robust reporter expression throughout the lateral line in Tol2 transgenic lines whereas the reporter activity was restricted to the anterior lateral line in case of genomic integration at the pIGLET24b landing site. In contrast, EnhH displayed similar pattern of activity along the entire lateral line in Tol2 and pIGLET14a transgenic lines.

The reduced fluorescence intensity in pIGLET lines in general, when compared to the Tol2 lines could be attributed to differences in transgene copy number, as Tol2-mediated transgenesis can generate multiple in-tandem insertions at a single genomic location leading to an overall high reporter expression, as opposed to the single-copy insertion in pIGLET lines. The pIGLET24b and pIGLET14a specific differences in enhancer activity suggest that the genomic environment of the integration site may also influence the enhancer activity. Alternatively, the enhancer activity in the posterior lateral line could also vary depending on the strength of the promoter. Because these factors can produce variability, we relied on multiple independent established lines from Tol2 founders exhibiting similar reporter expression patterns, apart from the pIGLET transgenic lines to confirm that the observed enhancer activity was in fact, reproducible

25. Sequence Divergence between Orthologous Enhancers can influence their Spatial Regulatory Activity

To test whether the sequence divergence between orthologous zebrafish and mouse enhancers can influence their spatial regulatory, we compared the reporter expression pattern driven by the two enhancer orthologs in context of zebrafish development. The constructs were integrated into the zebrafish genome to establish separate zebrafish and mouse ortholog containing lines. By keeping the contextual parameter the same, we can examine the extent to which sequence changes between the orthologs can influence their spatial activity. Here, we focus on EnhA and EnhE as representative examples, although additional enhancers were analysed in this study.

1. Spatial Activity of the orthologs of EnhA

At 5 days post fertilization, both the orthologs of EnhA displayed largely overlapping activity patterns at the tissue-level, with the reporter expression observed in the zebrafish lateral line and inner ear structures (Fig. 17A–B). This suggests that both the orthologous sequences retain their ability to drive expression in mechanosensory structures.

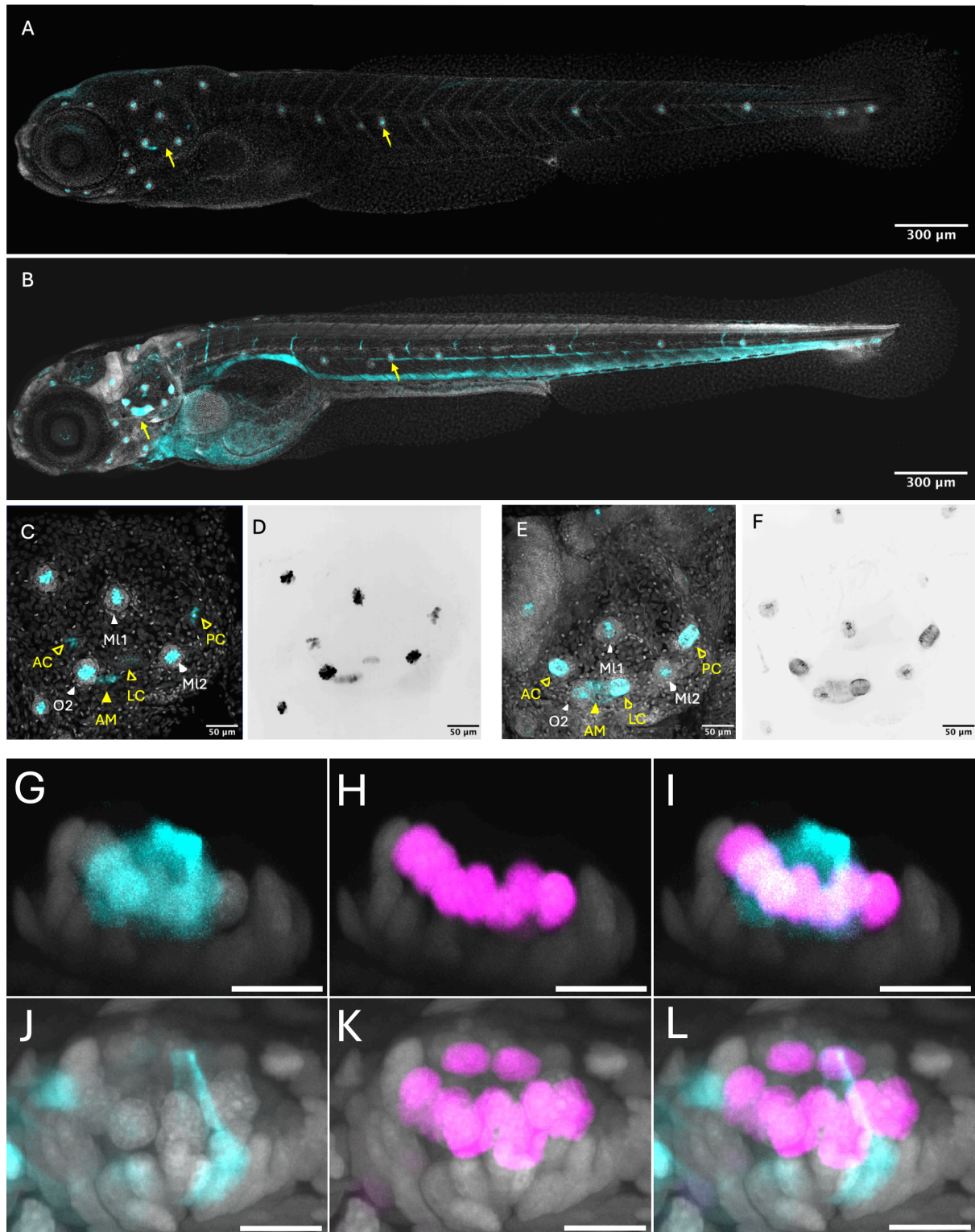


Figure 17: Spatial activity of the zebrafish EnhA enhancer and its mouse ortholog in zebrafish embryos at 5 dpf.

(A–B) Lateral views of 5 dpf zebrafish larvae showing reporter expression driven by the zebrafish EnhA enhancer (A) and the mouse ortholog (B). Both enhancers drive reporter expression in neuromasts of the lateral line system and in inner ear sensory structures.

(C–F) Higher magnification views of cranial neuromasts highlighting cell-type-specific differences in enhancer activity. The zebrafish EnhA enhancer shows reporter expression primarily in the central region of the neuromast, corresponding to hair cells (C,D), whereas the mouse ortholog displays reporter activity towards the peripheral region, consistent with mantle cell localization (E,F). Neuromast cell types are indicated: AC, accessory cells; LC, lateral cells; PC, peripheral cells; AM, anterior mantle cells; MI1 and MI2, mantle cell subtypes; O2, supporting cells.

(G–L) Confocal imaging of individual neuromasts showing reporter expression driven by the zebrafish enhancer (cyan) and the hair cell marker *myo6b* in magenta (G–I). The zebrafish enhancer shows activity predominantly in hair cells, whereas the mouse ortholog shows activity towards the peripheral region.

mouse ortholog shows reporter activity in a subset of central support cells, indicating differences in cellular specificity despite similar tissue-level expression patterns (J-L).

Scale bars: 300 μ m (A–B), 50 μ m (C–F).

However, a closer look at the neuromasts in the ear region revealed subtle differences in the reporter expression domains at the level of cell-type. In neuromasts of the zebrafish EnhA ortholog containing line, reporter activity was predominantly localised to the central region of the neuromasts which corresponds to the hair cells (Fig. 17C,D). In contrast, the mouse EnhA ortholog showed reporter activity preferentially towards the cells below the hair cells which corresponds to the central support cells (Fig. 17E,F). Further, imaging of the jaw-dissected fish to obtain a lateral view of the neuromast confirmed this distinction, as the activity of the zebrafish ortholog overlapped well with the hair cell nuclei marker *myo6b* in magenta, whereas the mouse ortholog drove activity partially in the cells surrounding the hair cells (Fig. 17G-L). These observations indicate that although the orthologous enhancers drive expression in same sensory structures, sequence divergence can lead to differences in terms of the specific cell-type populations, they drive activity in.

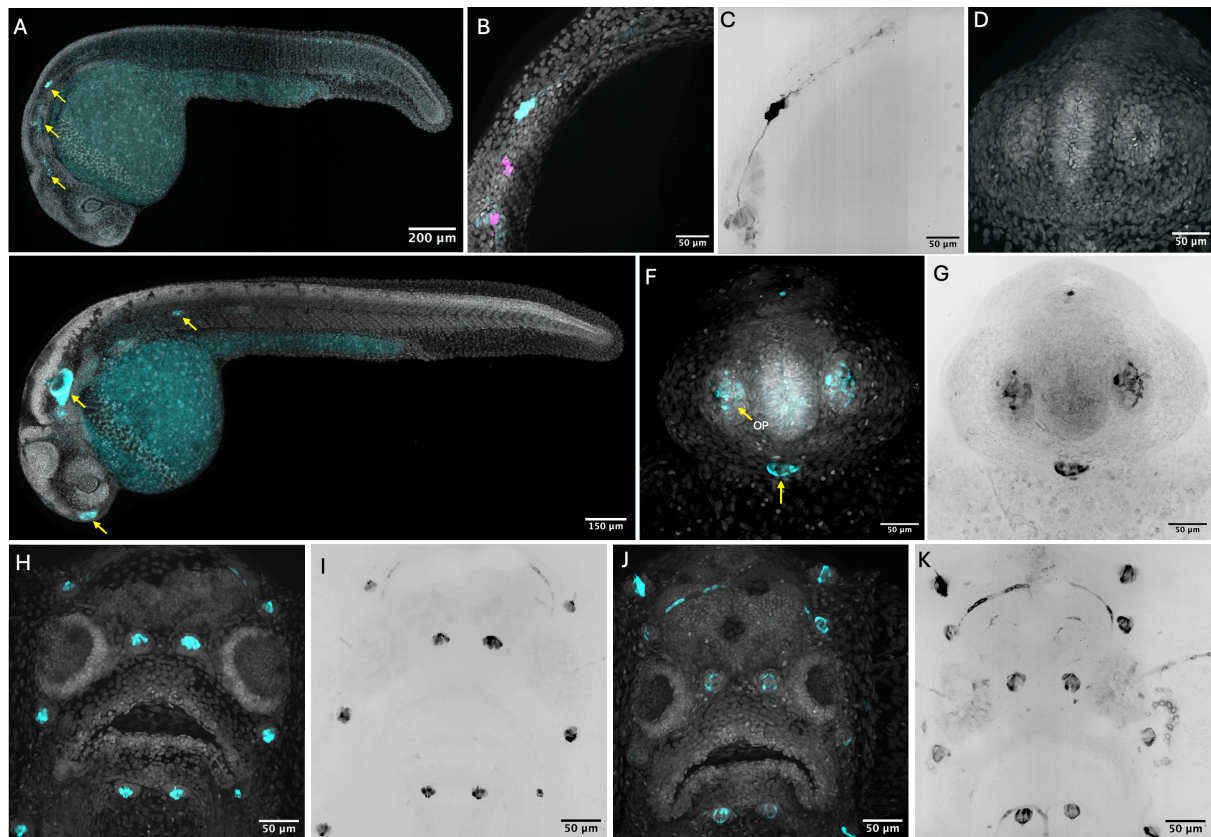


Figure 18: Developmental differences in the activity of zebrafish EnhA and the mouse ortholog.

(A–B) Reporter expression driven by the zebrafish EnhA enhancer. At 1 dpf, activity is observed in the posterior lateral line ganglion (C) Inverted GFP channel of image in B (D) Frontal views of embryos at 1 dpf with no detectable activity

(E) Reporter expression driven by the zebrafish EnhA ortholog at 1 dpf showing activity in the olfactory placode, otic placode and the lateral line primordium showing reporter expression driven by the mouse EnhA ortholog (F) in the olfactory placode (OP) and the developing mouth, indicated by arrows. (G) shows the image in D with an inverted LUT to improve visualization of the expression pattern. (H–I) Zebrafish EnhA enhancer at 5 dpf from the front (J–K) Mouse enhancer activity at 5 dpf

Scale bars: 150 μ m (C), 50 μ m (D–M).

We further wanted to examine whether any differences in activity would be observed at the tissue-level at earlier stages of development. Therefore we assessed enhancer activity at 1 dpf. Interestingly, at this stage, the mouse EnhA ortholog displayed reporter activity in both

the olfactory placode and a more ventral region (Fig. 18D–E). In contrast, the zebrafish EnhA ortholog showed no activity in the olfactory epithelium (Fig. 18C). By 5 dpf, however, the spatial activity patterns of both enhancers appeared largely similar at the tissue level (Fig. 18A–B).

2. Spatial Activity of the Orthologs of EnhE

We compared the activity the activity of the zebrafish EnhE enhancer and its mouse ortholog in transgenic zebrafish embryos. At the resolution of 10X, the activity of the enhancer orthologs in 5dpf embryos looked largely similar at the tissue level. Reporter expression for both the orthologs was observed in the lateral line system, including neuromasts and inter-neuromast cells in the eye and ear region (Fig.19 A-B).

However, taking a closer look at the head region revealed that the mouse ortholog showed additional reporter expression in the olfactory pit, whereas the zebrafish EnhE enhancer did not exhibit activity in this structure (Fig. 19C). Closer examination of the brain region in jaw-dissected embryos revealed that while the zebrafish ortholog of the enhancer is able to drive reporter expression robustly, the mouse ortholog drove the reporter expression sparsely in this region (Fig. 19D-E).

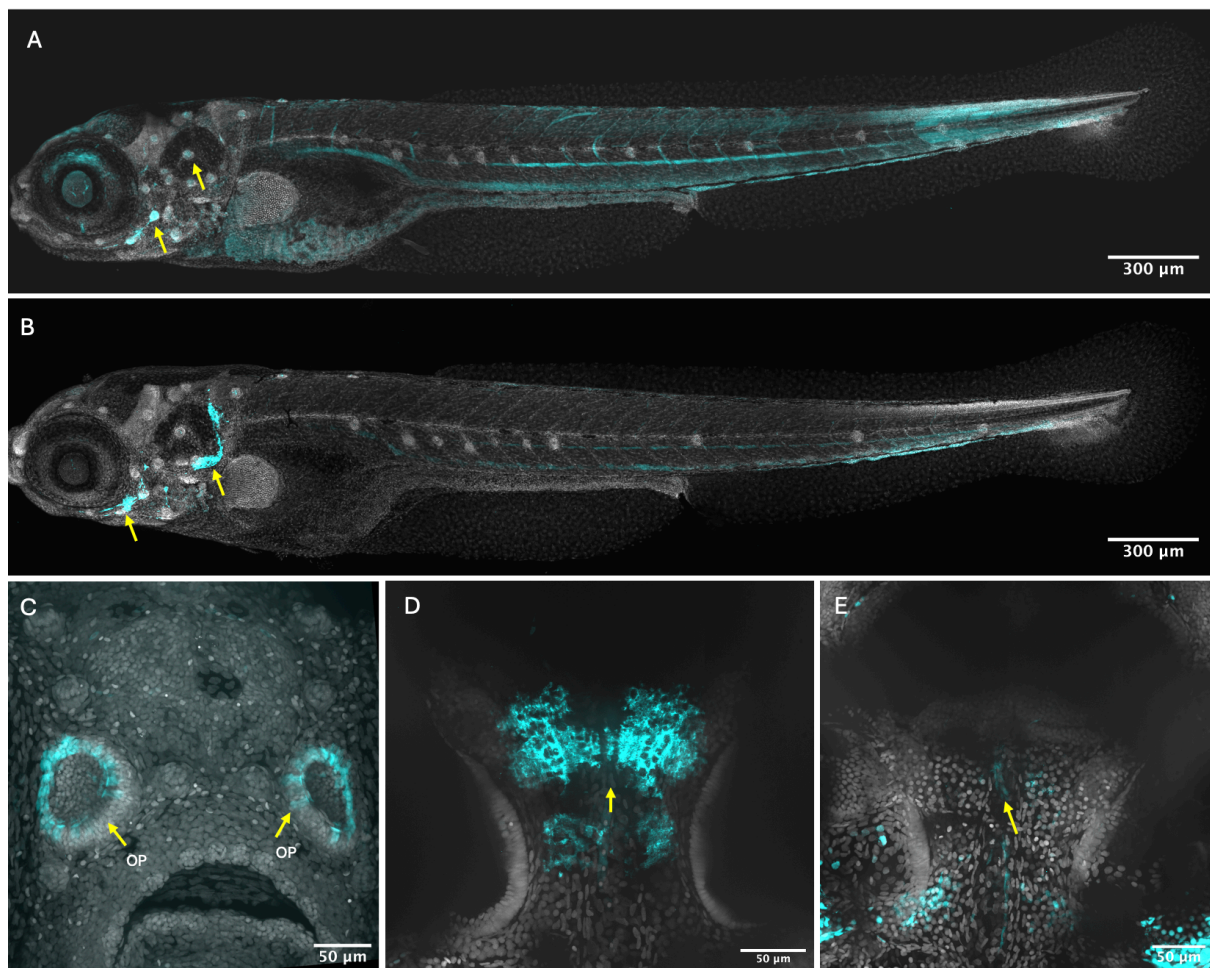


Figure 19: Comparison of the spatial activity of the zebrafish EnhE enhancer and its mouse ortholog in zebrafish embryos.

(A) Lateral view of a 5 dpf zebrafish larva showing reporter expression driven by the zebrafish EnhE enhancer. Reporter activity is observed in neuromasts and inter-neuromast cells of the lateral line system (B) Lateral view of a 5 dpf zebrafish larva showing reporter expression driven by the mouse ortholog of EnhE. Similar activity is observed in the lateral line system; however, additional reporter expression is detected in the olfactory pit (OP). (C) Higher magnification view of the lateral line region highlighting reporter expression in inter-neuromast cells and neuromasts. (D–E) Higher magnification

views of the brain region. The zebrafish EnhE enhancer drives robust reporter expression in the brain (D), whereas the mouse ortholog shows sparse activity in this region (E).

Scale bars: 300 μm (A–B), 50 μm (C–E).

Next, we examined whether the orthologs of EnhE exhibited differences in their activity at the level of cell-type in neuromasts. Imaging of the ear region with focus on the ML1 neuromast in the zebrafish ortholog containing transgenic line revealed reporter activity not only in the **central hair cells** but also in the surrounding **support-cells and mantle cells** (Fig. 20B–C). This wider distribution is particularly evident in the contrast-inverted image (Fig. 20C), where the reporter expression can be seen in the central hair cells and the peripheral cell of the neuromast.

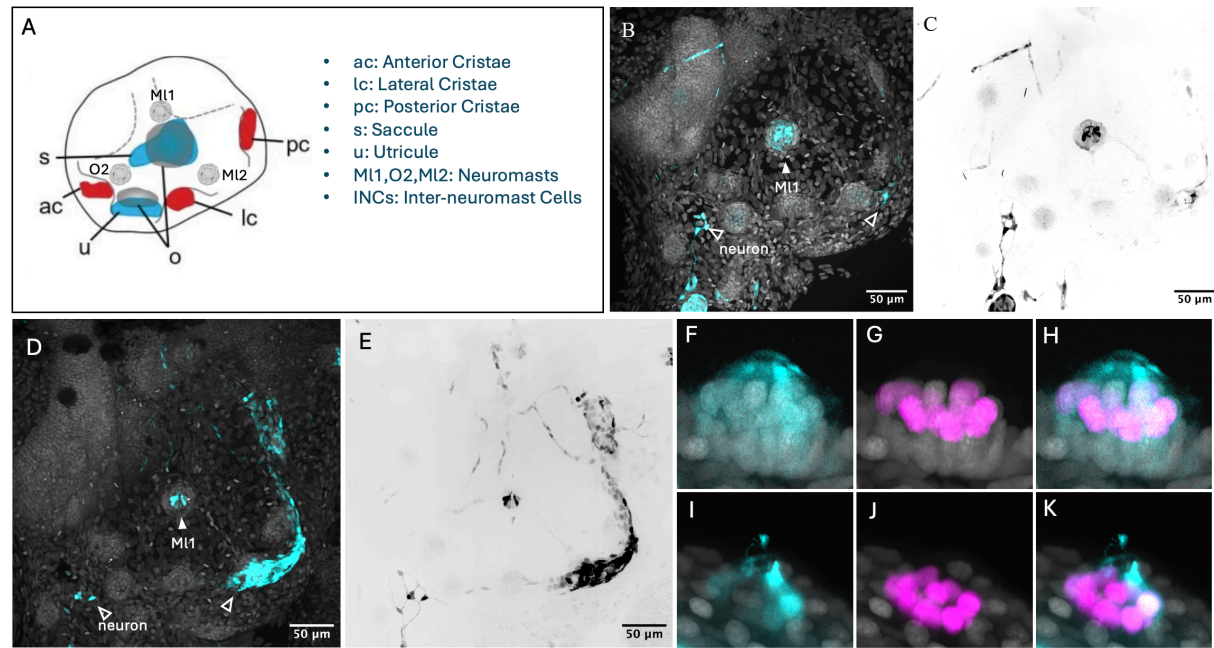


Figure 20: Cell-type specific differences in neuromast activity driven by the zebrafish EnhE enhancer and its mouse ortholog.

(A) Schematic representation of zebrafish inner ear and lateral line sensory structures indicating anatomical features and neuromast organization. Abbreviations: ac, anterior cristae; lc, lateral cristae; pc, posterior cristae; s, saccule; u, utricle; M1, O2, M2, neuromasts; INC, inter-neuromast cells. (B–C) Neuromast expressing the zebrafish EnhE ortholog. Reporter activity (cyan) is detected across multiple cell layers of the neuromast, including central hair cells as well as surrounding supporting and mantle cells, resulting in broader expression within the rosette structure. Panel C shows the same image with an inverted LUT to enhance visualization of the reporter signal. (D–E) Neuromast expressing the mouse EnhE ortholog. Reporter expression is largely restricted to the central hair cell population and is absent from peripheral supporting and mantle cells. Panel E shows the same image with an inverted LUT. (F–H) Lateral view of a neuromast expressing the zebrafish EnhE enhancer showing reporter activity extending across multiple cell layers. (I–K) Lateral view of a neuromast expressing the mouse EnhE ortholog, where reporter expression is restricted to the central hair cells, confirmed by co-labeling with hair-cell markers (magenta).

Scale bars: 50 μm .

In contrast, the mouse ortholog of EnhE exhibits a more restricted activity domain, where the reporter expression can be seen to be more restricted towards the centre, i.e., the hair cells, and was absent in the surrounding support-cells and mantle cells (Fig. 20D–E). Notably, the reporter expression was seen only in a subset of hair cells of the ML1 neuromast.

Imaging the ML1 neuromast laterally in jaw-dissected 5dpf embryos further supported these observations (Fig. 20F–K).

Taken together, the comparative analysis of the zebrafish and mouse orthologs of EnhA and EnhE, when the organismal context remains the same reveals that the sequence divergence between the orthologs can lead to changes in their regulatory output at both the cell and

tissue-level. In both the cases, the orthologs retain their ability to drive the reporter function in the lateral line and the cranial sensory structures. However, closer examination revealed differences in cellular specificity, spatial distribution, and developmental timing of enhancer activity. These observations suggest that sequence differences between enhancer orthologs can subtly alter the regulatory-output, while a significant overlap still remains in their activity, given the context organismal context remains the same.

26. Activity of EnhC in zebrafish and mouse embryos

To determine whether the validated enhancers show conserved or divergent spatial activity in zebrafish and mouse embryos, we analysed the spatial activity of the zebrafish orthologs of enhancers in zebrafish embryos and mouse ortholog in mouse embryos from a mouse enhancer-reporter transgenic line.

1. Activity of the zebrafish ortholog of EnhC activity in zebrafish embryos

Reporter activity driven by EnhC was examined in 2dpf zebrafish embryos as at this stage, EnhC activity is the highest. GFP expression was observed prominently in the otic vesicle, which is the developing inner ear structure in an embryo (Fig. 21A–B). Higher magnification images further confirmed that the activity seems to be in the developing macula of the inner ear (Fig. 21B–D). **No detectable enhancer activity was observed in the lateral line even at later stages of development (Fig S.),** despite it being recognised as a lateral line enhancer candidate from the ATAC-seq dataset.

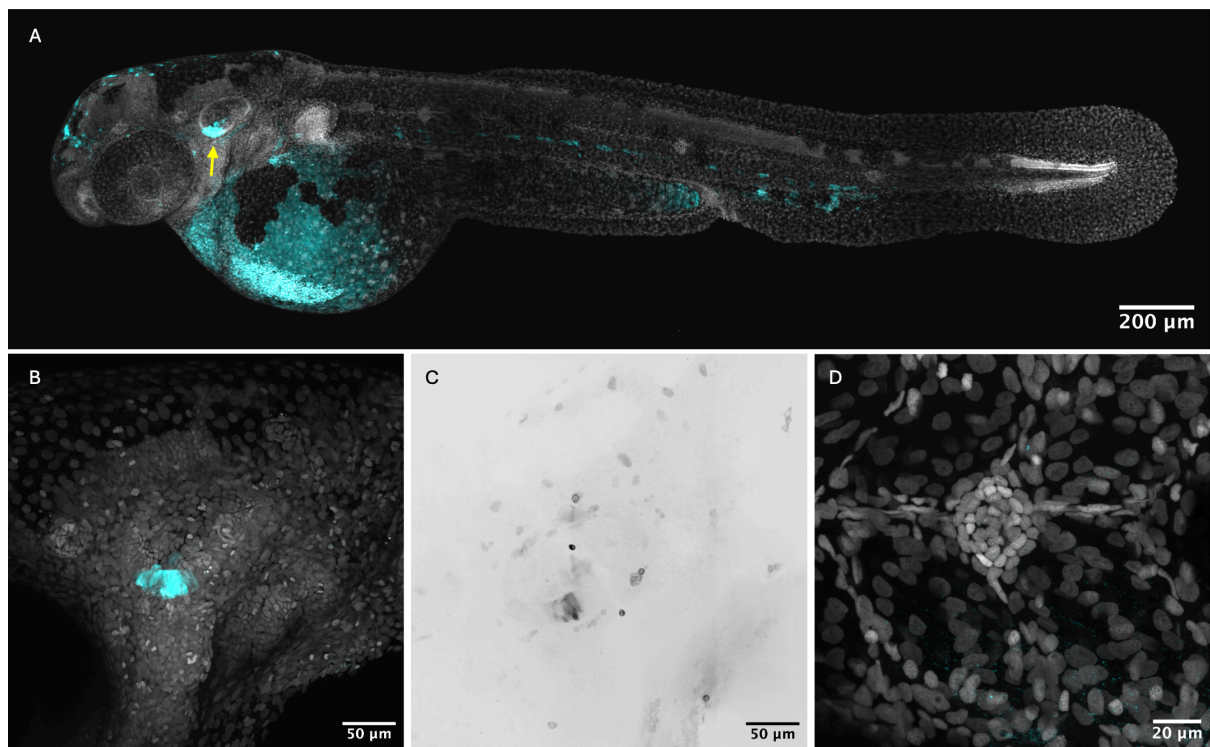


Figure 21: Activity of the zebrafish ortholog of EnhC in zebrafish embryos.

(A) Whole-mount image of a 2 dpf zebrafish embryo carrying a reporter driven by the zebrafish ortholog of EnhC. Fluorescent reporter expression (cyan) is detected in the otic vesicle (arrow), with no detectable activity along the lateral line system. Scale bar: 200 µm. (B) Higher magnification view of the otic vesicle region showing localized reporter expression in the sensory epithelium. Nuclei are counterstained in grey. Scale bar: 50 µm. (C) Inverted contrast view of the same region highlighting cellular morphology and reporter localization. (D) Posterior lateral line neuromast L1 showing no GFP expression indicating lack of enhancer activity. Scale bar: 20 µm.

Although, the reporter expression seen in the otic vesicle is consistent with the fact that they developmental programs and functional relationships are shared between hair cell so the inner

ear structures and the lateral line, suggesting that this enhancer may be associated with regulatory programs active in sensory epithelial structures.

2. Activity of the mouse ortholog of EnhC in transgenic mouse embryos

To examine whether there are novel spatial activity domains of the mouse ortholog of EnhC in mouse embryos, the mouse ortholog of EnhC was inserted into a reporter construct and analysed in mouse embryos at approximately E10.5. Sagittal sections of the mouse embryos were imaged to detect reporter expression driven by the enhancer.

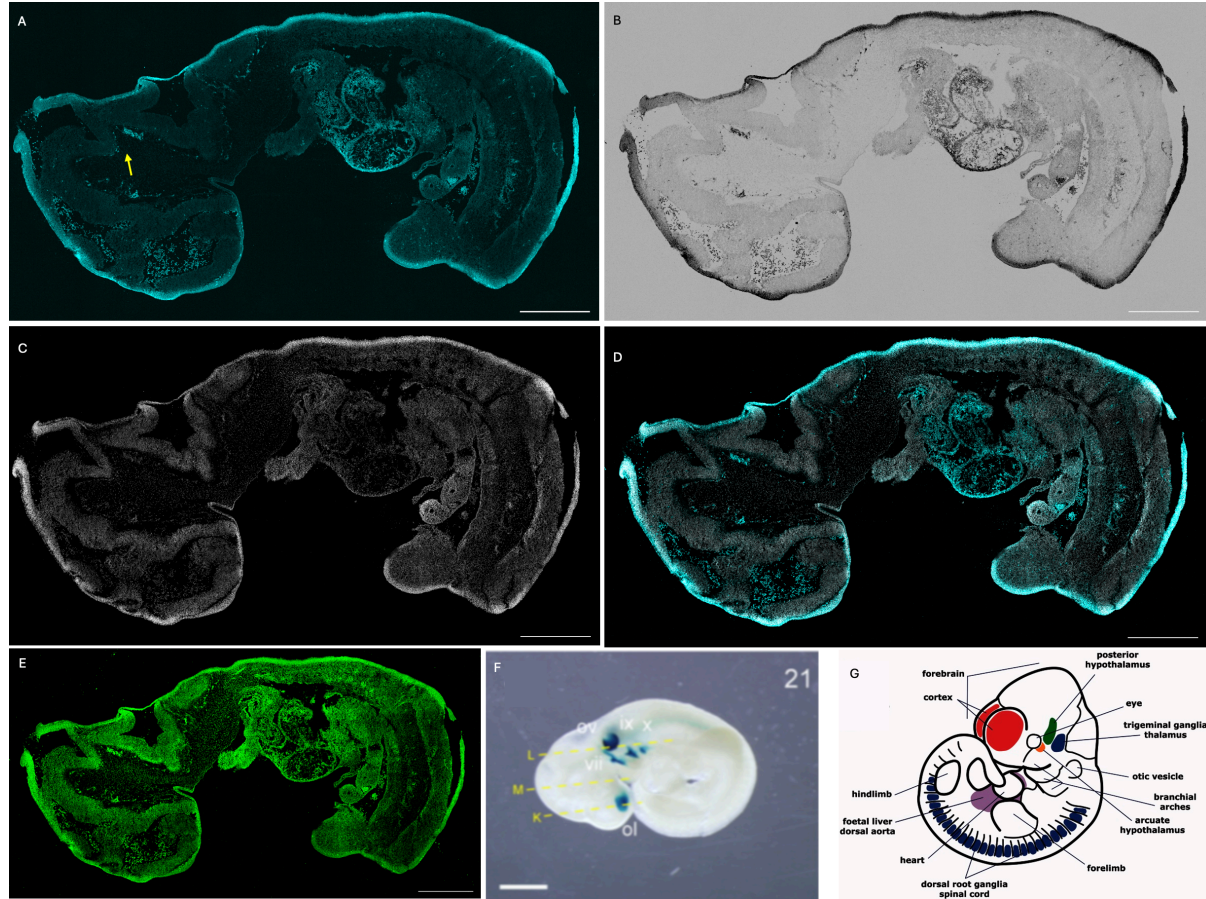


Figure 22: Analysis of the mouse ortholog of EnhC in a transgenic mouse embryo.

(A) Sagittal section of a mouse embryo showing endogenous GFP fluorescence (488 nm) driven by the mouse ortholog of EnhC. A localized signal in the head region (arrow) may correspond to the otic vesicle. (B) Inverted contrast image of the same section highlighting tissue morphology. (C) Nuclear staining with DAPI. (D) Merged image of GFP and DAPI signals. (E) Immunofluorescence staining using an anti-GFP antibody (568 nm). Despite antibody amplification, high background fluorescence is observed across the section. (F) Previously reported activity of this enhancer in mouse embryos showing expression in olfactory placode (ol), otic vesicle (ov), cranial sensory ganglia (VII–X), epibranchial placodes, and Rathke's pouch, adapted from Sato et al. 2012). (G) Schematic diagram of an E11 mouse embryo illustrating major anatomical structures, adapted from De Souza et al. 2013

Imaging the endogenous GFP signal directly revealed low-level fluorescence throughout the section (Fig. 22A), resulting in a high background signal. This obscured the clear identification of spatial activity domains of EnhC.

Despite the background, a localised bright signal was seen in the head region of the embryo. Based on the position of this signal in the sagittal section, it is difficult to tell the anatomical structure this might belong to. Due to the high level of background fluorescence, it was difficult to confidently assign it as a bona fide signal due to enhancer activity.

To improve the detection of GFP, we carried out immunofluorescence staining using an anti-GFP antibody (568 nm) (Fig. 22E). However, this did not improve the signal to noise ration

significantly across the section. Therefore technical limitations pertaining to histology and imaging currently hinder the clear identification of enhancer-driven GFP expression and further troubleshooting will be required to optimize this assay.

Interestingly, a previous study carried out by Sato et al. 2012 demonstrated that EnhC drives the reporter expression in olfactory placode, otic vesicle, cranial sensory ganglia, epibranchial placodes, and Rathke's pouch during mouse embryonic development (Fig. 22F).

Taken together, these observations suggest that the activity of EnhC is conserved across vertebrates in that, it is associated with sensory organ development, but its regulatory output may differ between the aquatic and terrestrial species. In zebrafish embryos it is primarily active in the otic vesicle whereas in tetrapods, this enhancer has expanded its activity domains to various multiple cranial placodes and sensory ganglia.

Although technical limitations prevented us from the clear identification of the enhancer activity in mouse embryos, previously published data supports the interpretation that EnhC has expanded and diversified its activity in terrestrial vertebrates. These findings are consistent with our broader hypothesis that conserved regulatory sequences can be repurposed to acquire novel functions, which might contribute to the diversification of sensory structures during vertebrate evolution.

27. Profiling of Enhancer Activity Across Multiple Founders

Name	Injected	Founder_ID	Imaging
enhA			
enhB			
enhC			
enhD			
enhE			
enhF			
enhG			
enhH			
enhI			
enhJ			
enhK			
enhL			

Given above is a table showing the 12 candidate enhancers of the lateral line that can drive reporter function. The following subsections explore the activity patterns for these various enhancers across founders of zebrafish and mouse orthologs, wherever available.

1. EnhA

EnhA activity in Zebrafish (Tol2 Transgenesis)

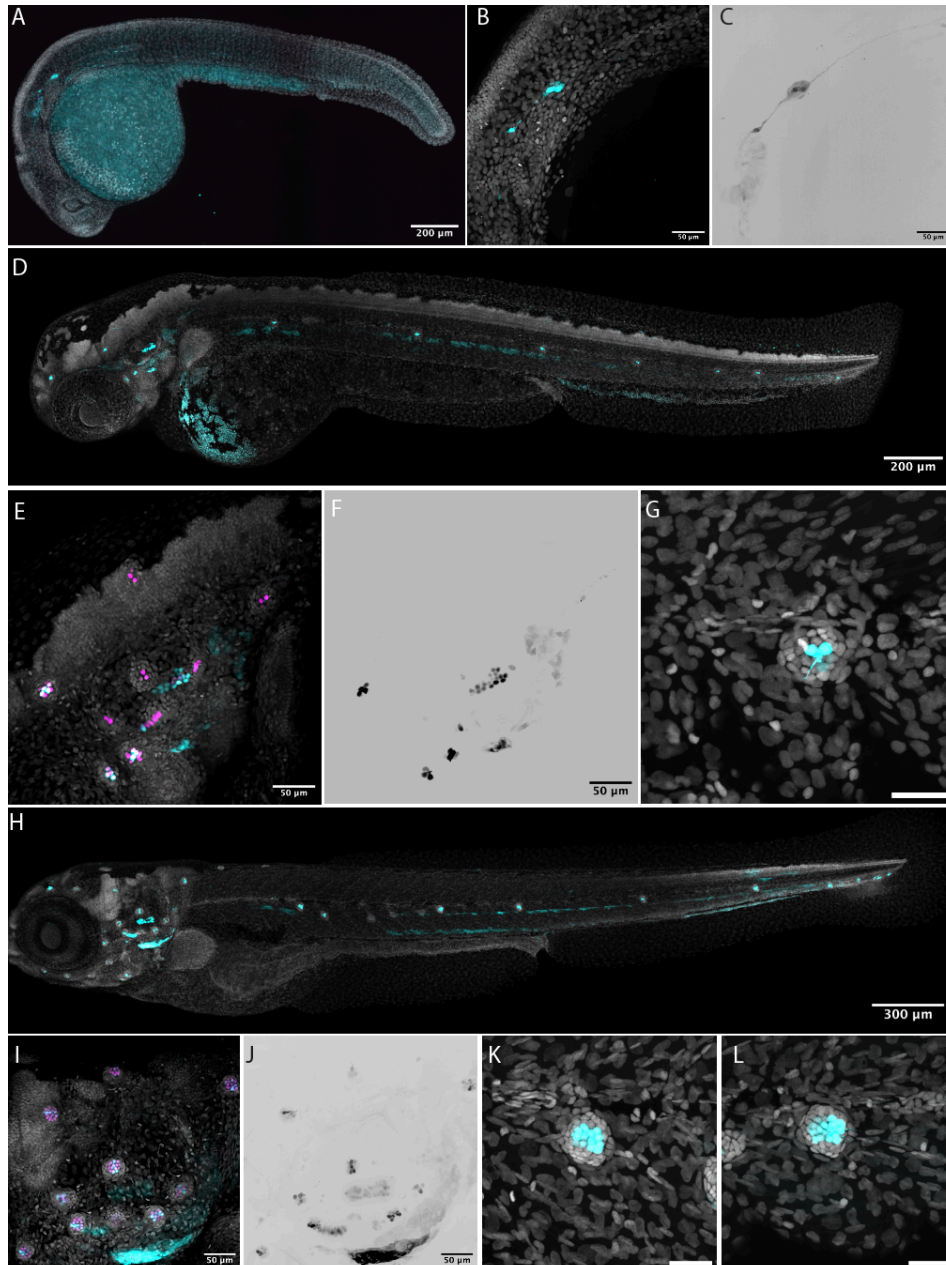


Figure 23: EnhA Tol2 founder 2 established line exhibits stage-dependent activity in inner ear structures and lateral line neuromasts.

(A) Lateral view of a 1 dpf embryo showing enhancer activity. (B–C) High-magnification (40X) views of the otic vesicle, with detectable activity in the posterior lateral line ganglion. (D) Whole-embryo view at 2 dpf showing expanded enhancer activity. (E–F) Magnified views of the ear region at 2 dpf. (G) High-magnification image of the L1 neuromast at 2 dpf. (H) Whole-embryo view at 5 dpf (10X), showing robust enhancer activity along the lateral line and in cranial regions. (I–J) Magnified views of the inner ear structures, including the cristae and maculae. (K–L) High-magnification images of L1 and L2 neuromasts, respectively.

The EnhA activity of the zebrafish ortholog was assessed in two founders using Tol2 transgenesis. We found enhancer activity to be reproducible across the two founders at 5dpf. Reporter activity was observed prominently in inner ear structures (cristae and maculae) as well as in the anterior and posterior lateral line neuromasts. Yet, minor differences were observed in the established line derived from founder 2, where we could see additional EnhA activity at the ear border, a feature that we did not detect in the established line derived from founder 1. Apart from this variation, the enhancer activity across the two founders was fairly

conserved. The activity was observed in the hair cells of the inner ear structures and the neuromasts.

EnhA Activity in Zebrafish (pIGLET Transgenesis)

We also assessed EnhA activity using the pIGLET transgenesis system in zebrafish (single founder). The comparison between the Tol2 and pIGLET transgenesis systems, including the differences in the strength of enhancer activity have been discussed in detail, in section 24.1.

EnhA in Mouse Ortholog (pIGLET Transgenesis)

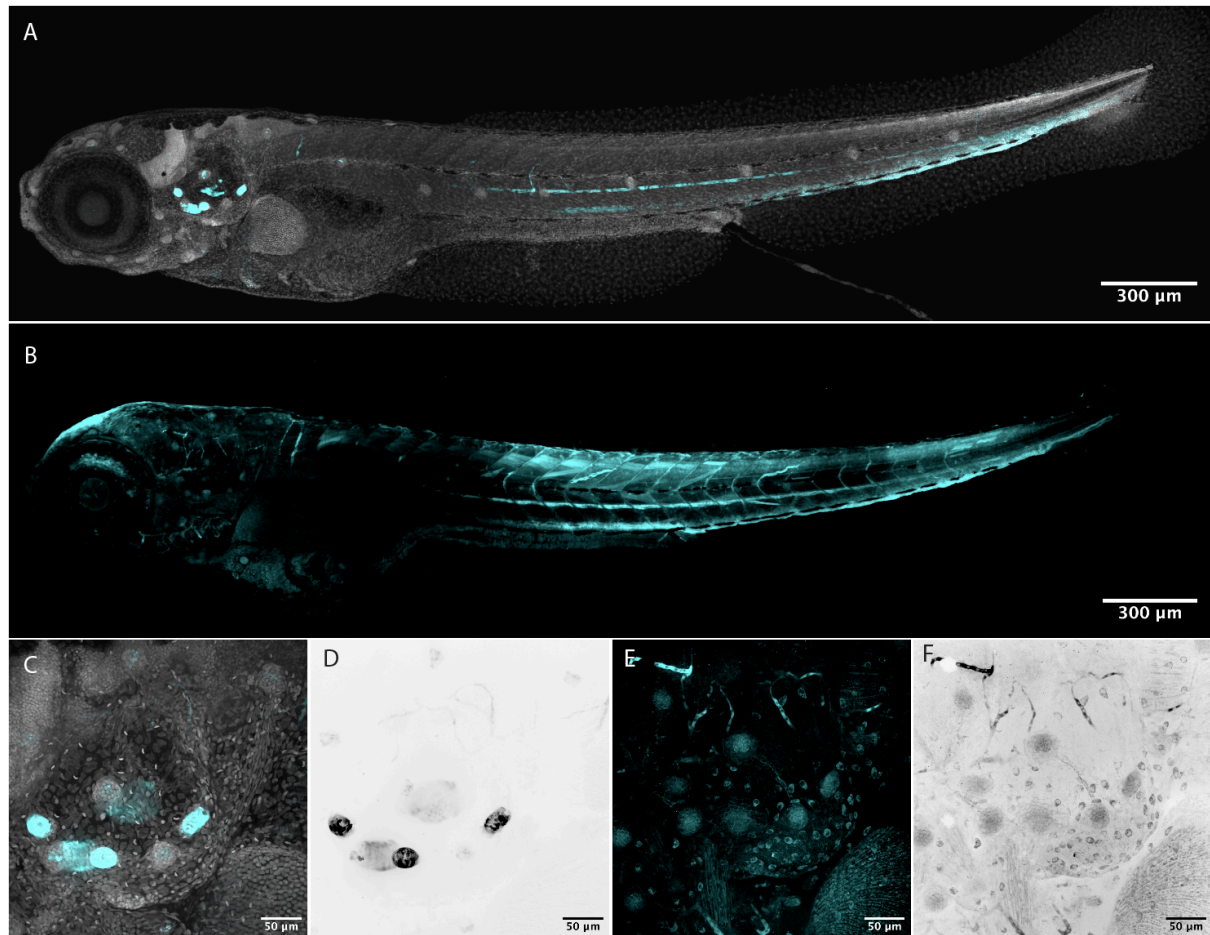


Figure 24: Mouse ortholog of EnhA (pIGLET24b) shows variable enhancer activity across independent founders at 5 dpf.

(A–B) Whole-embryo lateral views (10×) of two independent founders: Founder 1 (A) and Founder 2 (B), showing differences in overall enhancer activity levels. (C–D) High-magnification views of the inner ear in Founder 1 progeny, showing robust expression in the cristae and maculae. (E–F) Corresponding views of the inner ear in Founder 2 progeny, displaying markedly reduced enhancer activity.

Two independent founders using the pIGLET transgenesis system were analysed—

Tg(mmEnhA:βglobin:eGFP)pIGLET24b_01 and

Tg(mmEnhA:βglobin:eGFP)pIGLET24b_02. We observed notable differences across these two founders. Founder1 derived established line exhibited strong enhancer activity with a robust reporter expression in the neuromasts of the anterior and posterior lateral line, and the cristae and the maculae of the inner ear. But the Founder 2 derived established line displayed markedly reduced activity overall. While enhancer activity was still detectable in the inner ear structures and the neuromasts, the signal intensity was substantially lower. The enhancer activity was detectable in L1 neuromasts but not in L2 and L3, indicating reduced or absence

of activity. These differences suggest variability across the founders even in a controlled transgenesis system.

Developmental dynamics of Mouse EnhA (Founder 1) in pIGLET24b

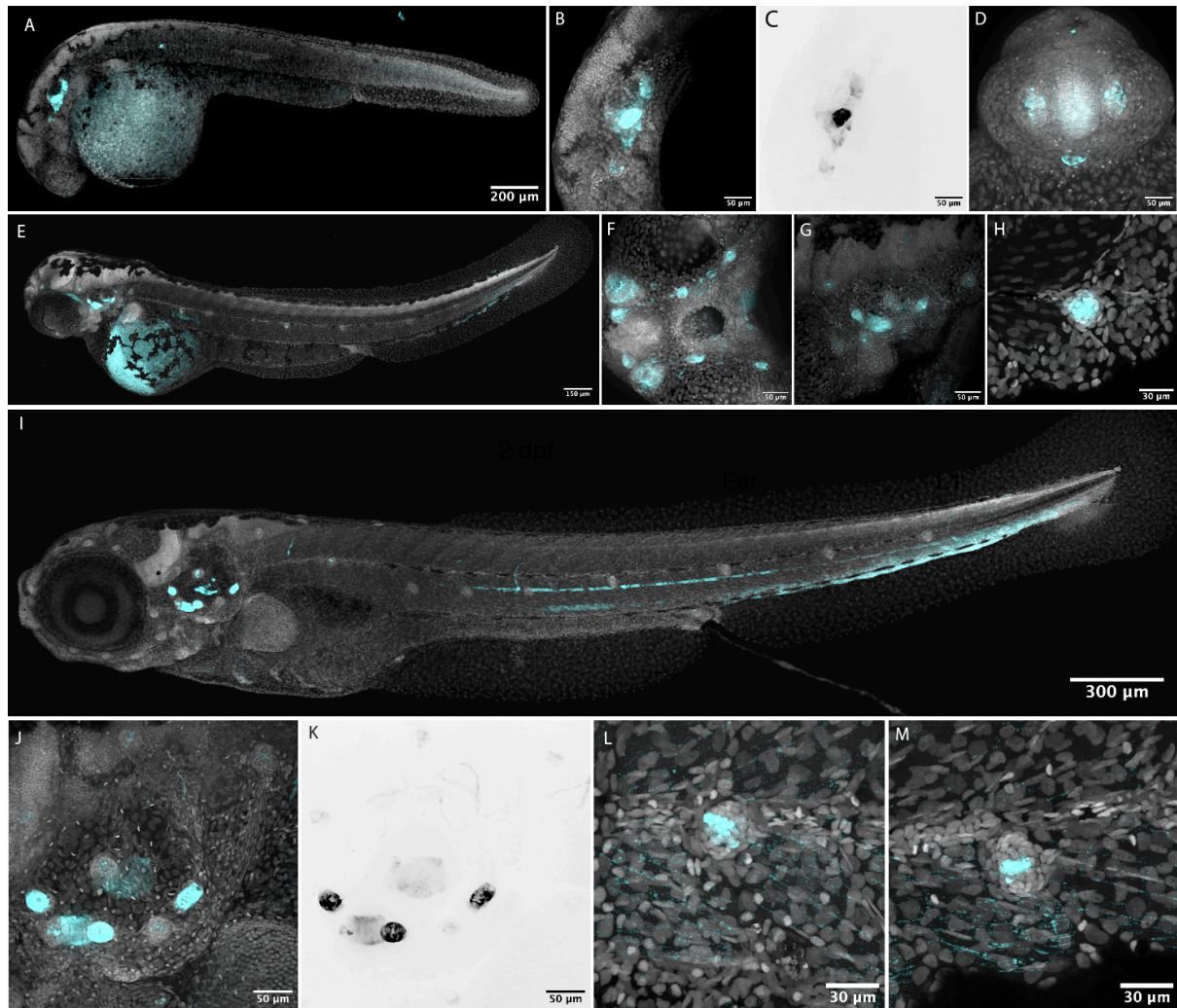


Figure 25: Mouse ortholog of EnhA (pIGLET24b, founder 1) exhibits dynamic enhancer activity across developmental stages

(A) Whole-embryo lateral view at 1 dpf, showing enhancer activity in cranial regions. (B) High-magnification view of the otic vesicle at 1 dpf. (C) Inverted GFP signal highlighting enhancer activity. (D) High-magnification view of the snout region, showing anterior expression domains. (E) Whole-embryo lateral view at 2 dpf (10 $\times$), showing enhancer activity. (F) Ventral view highlighting activity in the developing mouth and olfactory epithelium. (G) High-magnification view of the otic vesicle. (H) High-magnification image of the L1 neuromast. (I) Whole-embryo lateral view at 5 dpf (10 $\times$), showing robust enhancer activity in inner ear structures and along the lateral line. (J) High-magnification view of the inner ear, with expression in the cristae and maculae. (K) Inverted GFP signal of the ear region. (L–M) High-magnification images of L1 and L2 neuromasts, respectively, showing enhancer activity within neuromast-associated cells.

At 1 dpf, enhancer activity could be detected in the otic vesicle, deposited neuromast, the olfactory epithelium and a more ventral region which could potentially correspond to the developing oral region. At 2 dpf, the enhancer activity became more defined in the cristae and maculae, as well as the anterior and posterior lateral line neuromasts. Expression in the olfactory epithelium persisted. At 5dpf the enhancer exhibited strong activity in the cristae, maculae and the lateral line neuromasts. The activity was specifically seen in the support cell population of the inner ear structures and most neuromasts.

Comparison of the Activity across Zebrafish and Mouse orthologs of EnhA

A detailed comparison of enhancer activity between zebrafish and mouse orthologs is provided in Section 25.1.

Briefly, the zebrafish ortholog of EnhA exhibits reproducible enhancer activity across different founders in the Tol2 lines. In contrast, the mouse ortholog exhibits greater variability across founders in terms of the strength and spatial activity.

2. EnhB

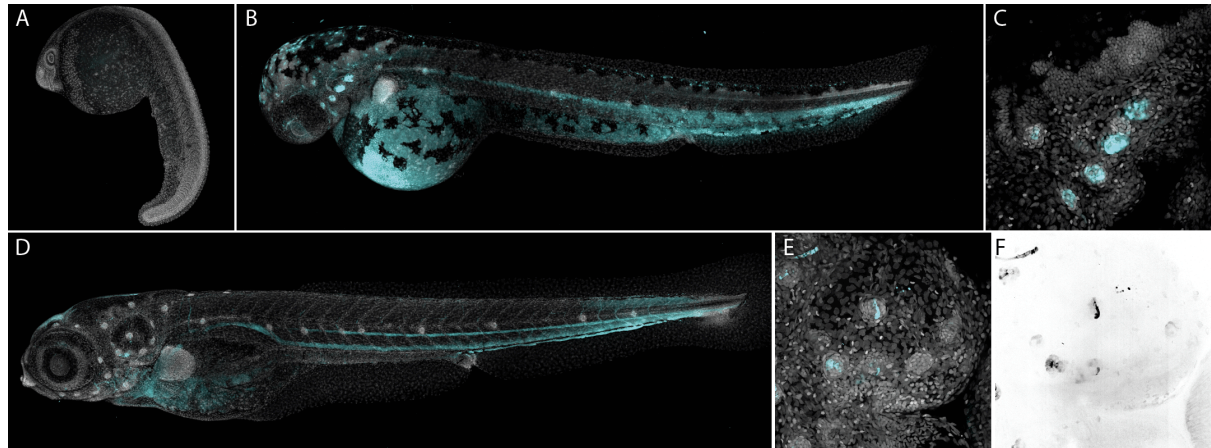


Figure 26: EnhB (Tol2, founder 1) stage-dependent enhancer activity in zebrafish.

(A) Lateral view of a 1 dpf embryo showing no detectable enhancer activity. (B) Whole-embryo view at 2 dpf, with onset of enhancer activity. (C) High-magnification (40×) view of the otic vesicle at 2 dpf, showing localized expression. (D) Whole-embryo view at 5 dpf. (E–F) High-magnification views of the inner ear at 5 dpf, showing expression in the cristae and ML1 and O2 neuromasts

We assessed the activity of EnhB using Tol2 transgenesis in zebrafish in a single founder. At 1 dpf, no detectable activity was observed. Enhancer activity became apparent at 2 dpf in the otic vesicle and some early deposited neuromasts as well. At 5 dpf, enhancer activity was detected in the inner ear structures, cristae and maculae. Within the lateral line, the activity seemed to be stronger in the neuromasts of the anterior lateral line and comparatively weaker in the posterior lateral line. At the cell-type level, the pattern of expression suggests that EnhB is active predominantly in the mantle cells of the neuromast.

3. EnhC

Developmental Progression of EnhC Activity in Zebrafish Founder 1 (Tol2 Transgenesis)

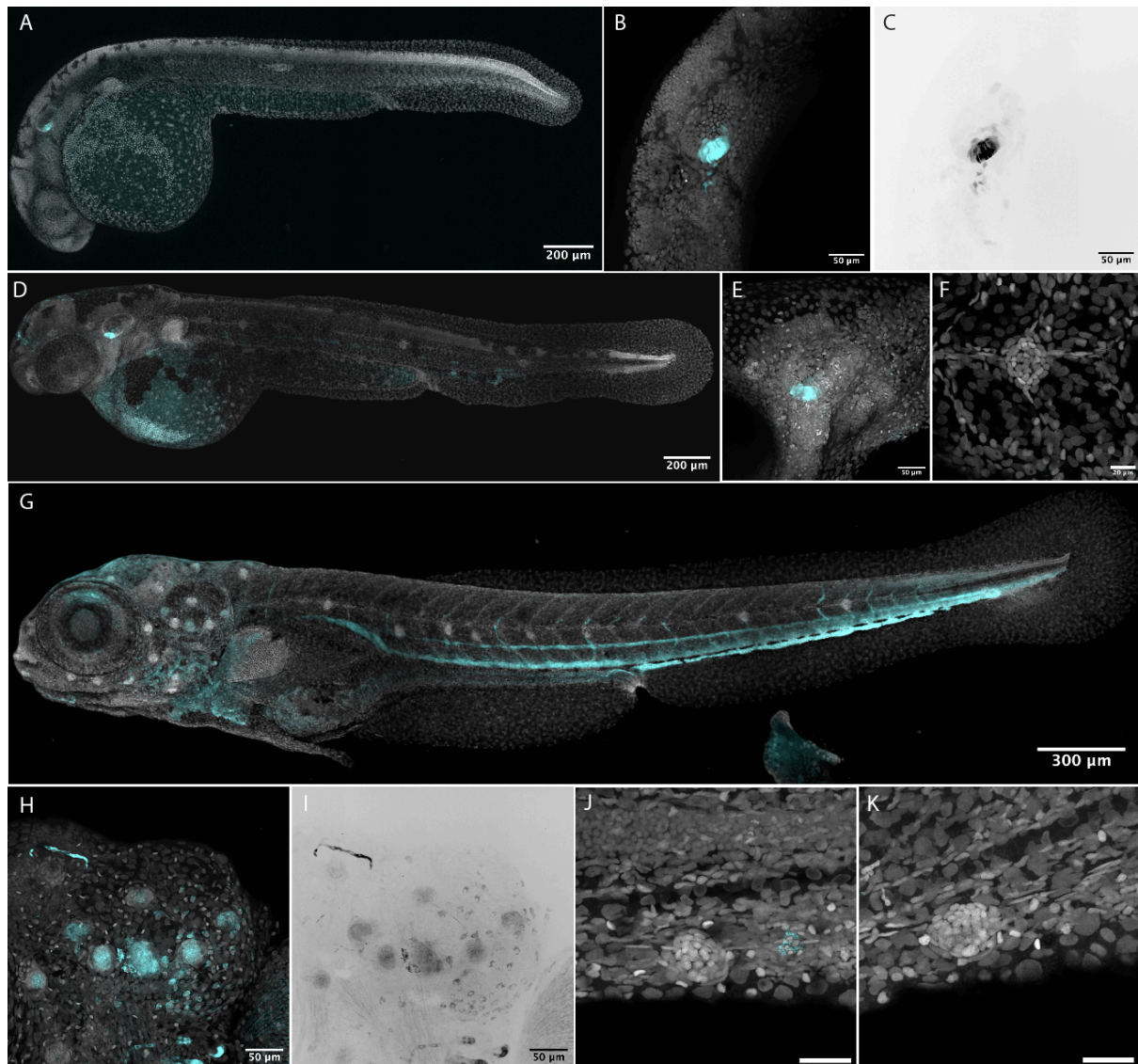


Figure 27: EnhC (Tol2, founder 1) exhibits restricted activity in inner ear structures and anterior lateral line neuromasts across development.

(A) Whole-embryo lateral view at 1 dpf (10×) showing early enhancer activity. (B–C) High-magnification views of the otic vesicle at 1 dpf. (D) Whole-embryo view at 2 dpf, with activity maintained in the otic region. (E) High-magnification view of the otic vesicle at 2 dpf. (F) High-magnification image of the L1 neuromast at 2 dpf. (G) Whole-embryo lateral view at 5 dpf (10×), showing enhancer activity along the anterior lateral line. (H–I) High-magnification views of the inner ear, with expression in the cristae. (J–K) High-magnification images of L1 and L2 neuromasts, respectively.

At 1dpf, EnhC activity was detected in the otic vesicle in a localised region. At 2pf, the enhancer activity remained localised within the restricted region of the otic vesicle. At this stage, the inner ear structures are not yet fully differentiated, making it difficult to assign activity to specific domains such as cristae or maculae. At 5dpf, the enhancer activity became more spatially defined, with reporter expression in the cristae and anterior lateral line neuromasts. It was difficult to assign the cell-type specific activity due lack of sharp activity domains in the activity of enhancer.

EnhC Activity in Zebrafish (Tol2 Transgenesis)

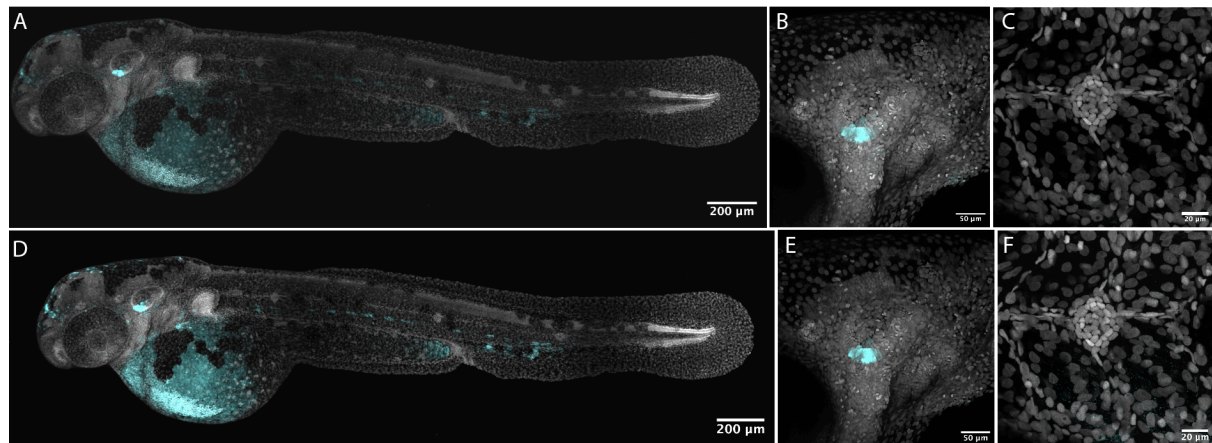


Figure 28: EnhC (Tol2) shows reproducible enhancer activity across independent zebrafish founders at 2 dpf.

(A–C) Zebrafish EnhC (Tol2) founder 1 at 2 dpf. (A) Whole-embryo lateral view showing enhancer activity. (B) High-magnification view of the otic vesicle, with localized expression. (C) High-magnification view of a neuromast.

(D–F) Zebrafish EnhC (Tol2) founder 2 at 2 dpf. (D) Whole-embryo lateral view showing a similar pattern of enhancer activity. (E) High-magnification view of the otic vesicle. (F) High-magnification view of a neuromast.

We assessed the activity of zebrafish ortholog of EnhC using two independent founders established using Tol2 transgenesis system. At 5dpf, both founders exhibited largely similar pattern of activity, which gives us confidence in their activity being reproducible. EnhC activity was consistently detected in the cristae of the inner ear and the anterior lateral line neuromasts, with no detectable activity found in the posterior lateral line neuromasts.

EnhC activity in Mouse Ortholog (pIGLET Transgenesis)

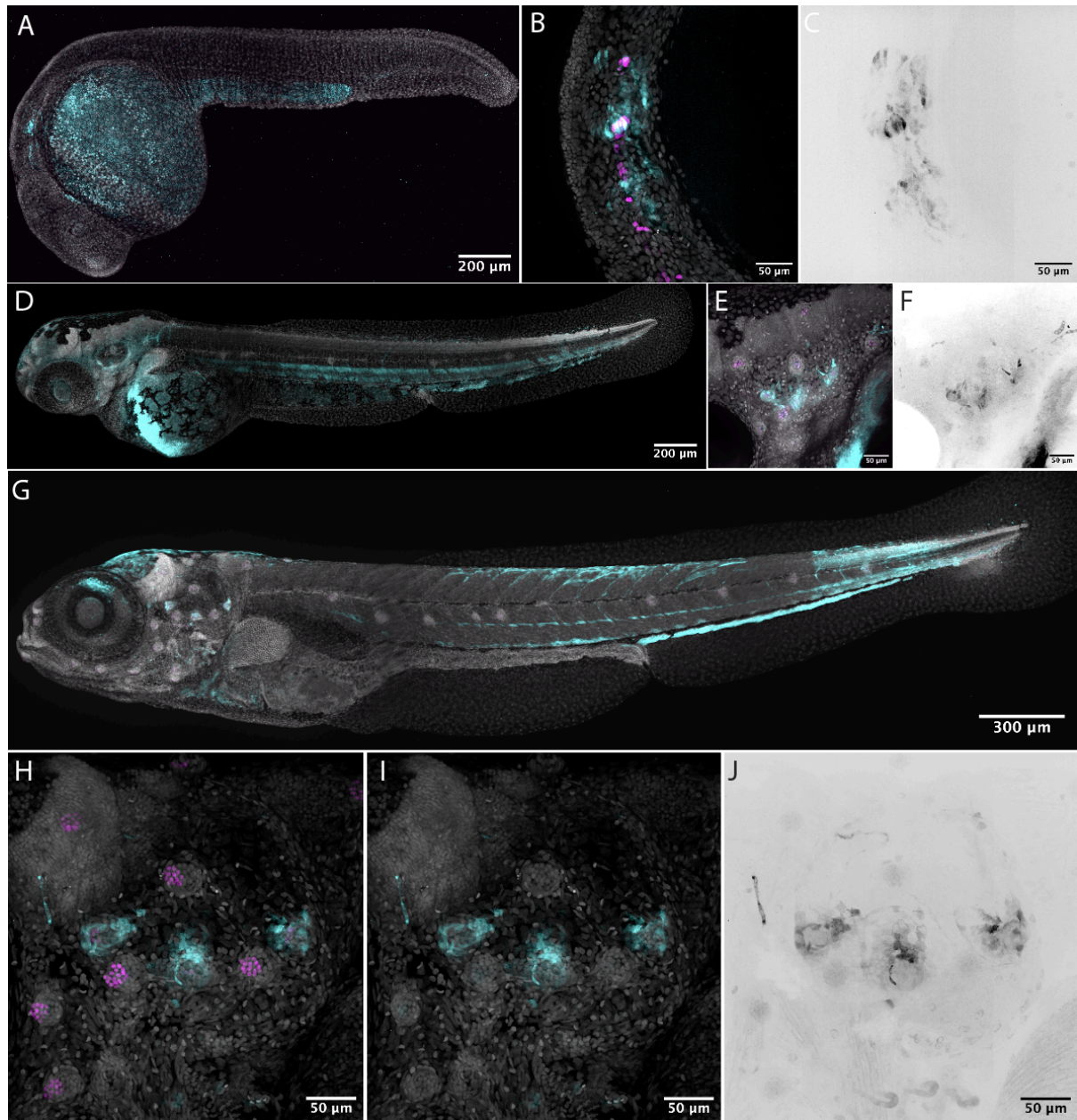


Figure 29: Mouse ortholog of EnhC (pIGLET24b) founder 1 progeny activity across developmental stages.

(A) Whole-embryo lateral view at 1 dpf (10×) showing enhancer activity. (B–C) High-magnification views of the otic vesicle at 1 dpf. (D) Whole-embryo view at 2 dpf, with enhancer activity maintained in the otic region. (E–F) High-magnification views of the otic vesicle at 2 dpf.

(G) Whole-embryo lateral view at 5 dpf, showing prominent enhancer activity in cranial regions. (H–J) High-magnification views of the inner ear at 5 dpf: (H) merged image showing hair cell marker (magenta), GFP, and DAPI, (I) GFP channel and DAPI, and (J) inverted GFP signal, highlighting enhancer activity in the cristae.

The activity of EnhC was tested using a single pIGLET24b founder. At 1 dpf, the activity was detected in the otic vesicle, with activity observed in rather broader regions of the otic vesicle, as well as in a ventral domain beneath the otic region, which could not clearly be assigned an anatomical structure. At 2 dpf, the enhancer activity could be detected in the developing cristae of the inner ear. At 5 dpf, the reporter expression was observed in the cristae. However, no robust activity was detected in the lateral line neuromasts, with only faint signal, if any.

Comparison of the Activity across Zebrafish and Mouse orthologs of EnhC

Comparison of the activity of both the orthologs of EnhC revealed both overlapping and divergent features. The activity in the inner ear, pertaining to cristae was consistent in both the orthologs at 5dpf. However, notable differences were observed in the spatial activity at 1dpf and the lateral line at 5dpf. The zebrafish enhancer shows clear activity in the anterior lateral line neuromasts, whereas the mouse ortholog exhibits little to no detectable activity in neuromasts. Additionally, the mouse version at 1dpf displays broader expression domains in the otic vesicle, including regions more ventral to the otic vesicle, that were not observed in the zebrafish ortholog of the enhancer.

4. EnhE

EnhE Activity in Zebrafish (pIGLET Transgenesis)

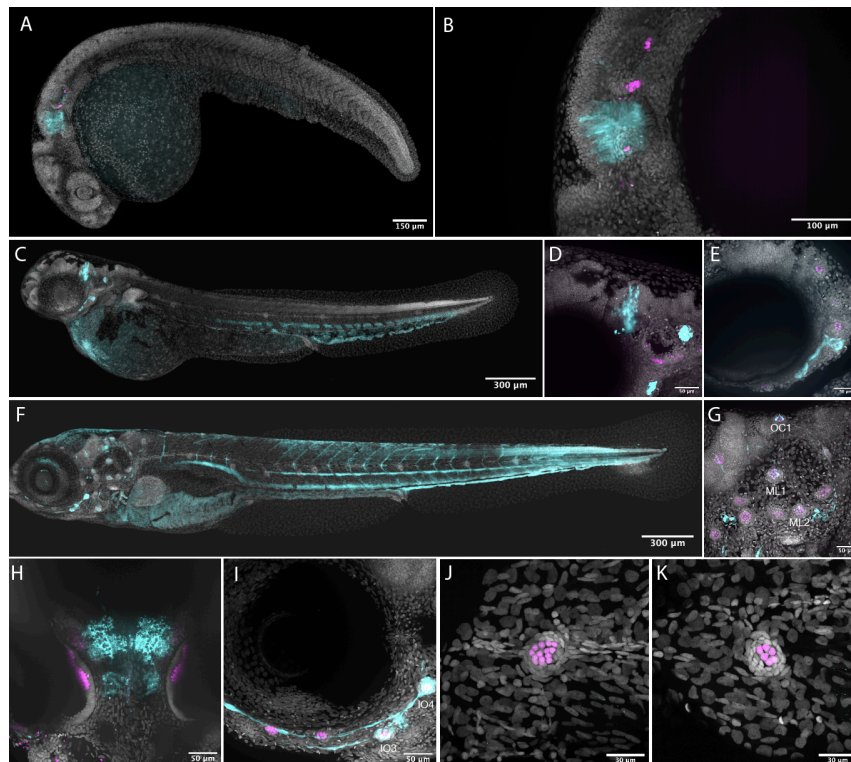


Figure 30: EnhE (zebrafish ortholog, pIGLET24b founder 1) exhibits neural and lateral line-associated activity across developmental stages.

(A) Whole-embryo lateral view at 1 dpf showing GFP (cyan), DAPI (grey), and hair cell marker (magenta). (B) High-magnification (40×) view of the otic region at 1 dpf. (C) Whole-embryo view at 2 dpf (10×), showing expanded enhancer activity. (D) High-magnification view of the otic vesicle at 2 dpf. (E) View of the eye region at 2 dpf, showing anterior expression. (F) Whole-embryo lateral view at 5 dpf (10×), showing robust activity along the hindbrain and lateral line. (G) High-magnification view of the ear region, including anterior lateral line neuromasts (OC1, ML1, ML2). (H) Dissected brain showing strong enhancer activity in the hindbrain. (I) High-magnification view of the eye region at 5 dpf (IO2–IO3 region). (J–K) High-magnification images of L1 and L2 neuromasts, respectively, with hair cells marked in magenta and minimal or no overlap with EnhE-driven GFP expression, indicating activity in non-hair cell populations.

Enhancer activity for the zebrafish ortholog of EnhE was assessed in a single founder of pIGLET24b founder across multiple developmental stages. At 1 dpf, enhancer activity was seen in the developing hindbrain. At 2dpf, the activity persisted in hindbrain and was additionally detected in the early deposited neuromasts. Enhancer activity was also observed in the region, posterior to eye, which could correspond to the anterior lateral line primordium. At 5dpf, the enhancer activity became more refined spatially, with activity detected in components of the lateral line and the nervous system. Expression was observed in the anterior lateral line neuromasts, particularly including those in the ear region (OC1, ML1,

ML2) and beneath the eye (IO3, IO4). In addition, we observed activity in regions ventral to the eye, which likely correspond to the neuronal projections due to the presence of cell-body like structures in the vicinity. Dissected samples further confirmed strong enhancer activity in the hindbrain, indicating that neural expression is a prominent feature of EnhE activity.

EnhE Activity in Mouse Ortholog (Tol2 Transgenesis)

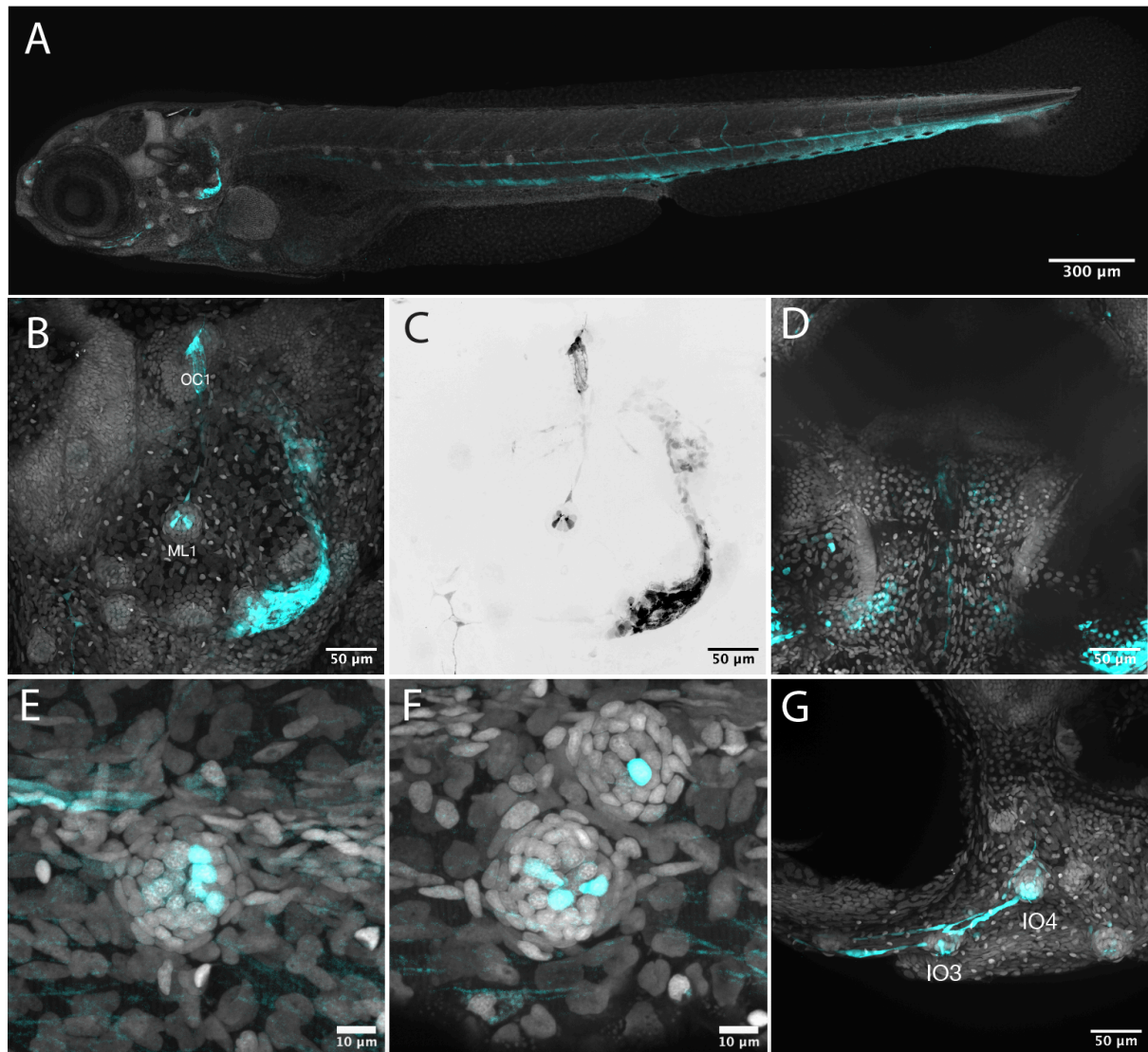


Figure 31: Mouse ortholog of EnhE (Tol2, founder 1) shows robust activity in lateral line neuromasts, ear-associated regions, and neural tissues at 5 dpf.

(A) Whole-embryo lateral view (10×) showing enhancer activity along the lateral line, ear and the olfactory pits. (B–C) High-magnification views of the ear region, including anterior lateral line neuromasts (OC1, ML1) and surrounding tissues. (D) View of the brain region following jaw dissection, revealing enhancer activity in neural tissues. (E–F) High-magnification images of L1 and L2 neuromasts, respectively, showing enhancer activity within hair cells. (G) High-magnification view of the eye region, with activity in the IO3 and IO4 neuromasts.

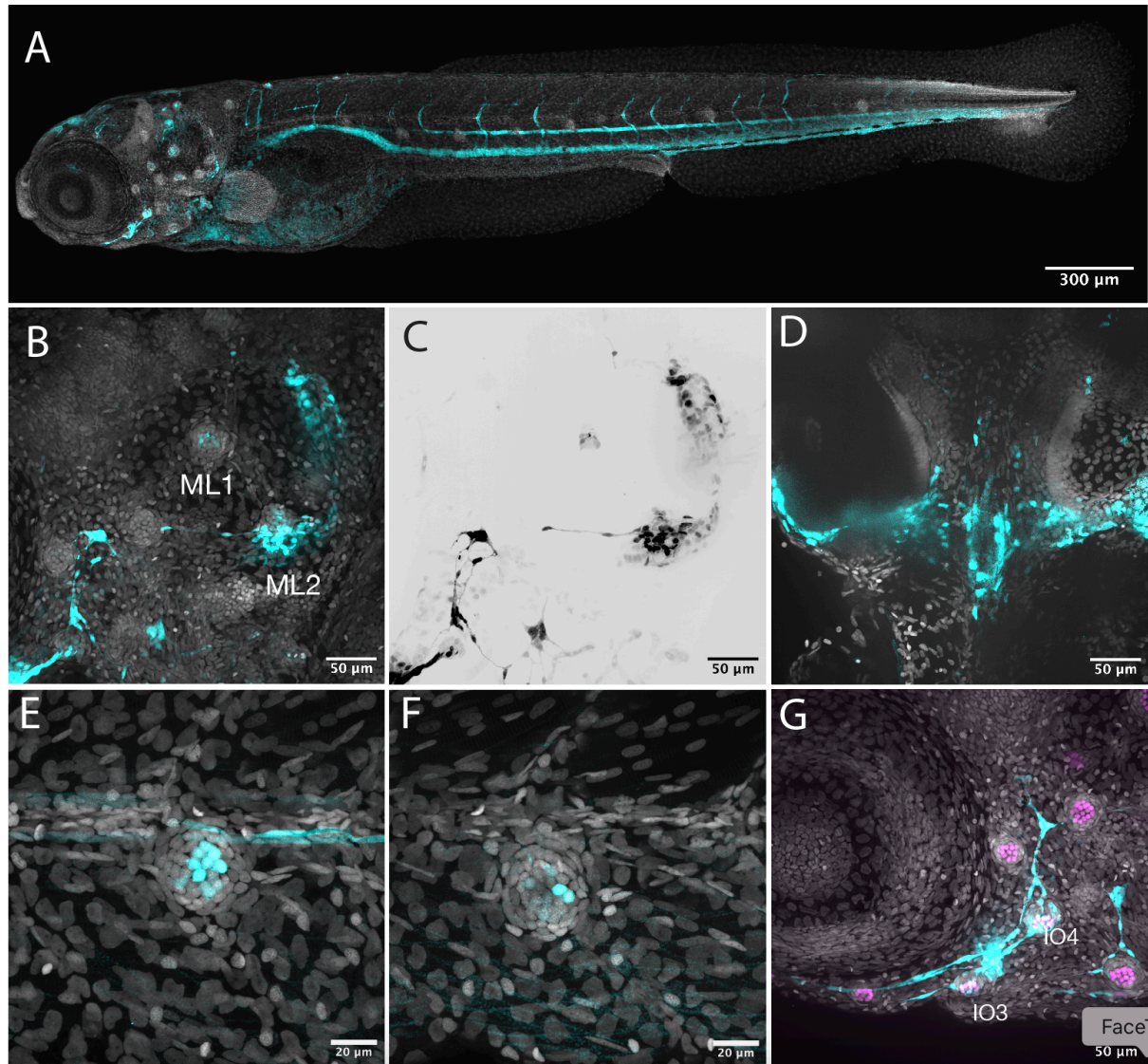


Figure 32: Mouse ortholog of EnhE (Tol2, founder 2) enhancer activity at 5dpf

(A) Whole-embryo lateral view (10×) showing enhancer activity along the lateral line and in cranial regions. (B–C) High-magnification views of the ear region, including anterior lateral line neuromasts (ML1, ML2) and surrounding tissues. (D) View of the brain region following jaw dissection, showing enhancer activity in neural tissues. (E–F) High-magnification images of L1 and L2 neuromasts, respectively, showing enhancer activity within hair cells. (G) High-magnification view of the eye region, with activity in the IO3 and IO4 neuromasts. In all panels, GFP (cyan) represents enhancer activity, magenta marks myo6b-positive hair cells, and DAPI (grey) labels nuclei.

Enhancer activity of the mouse ortholog of EnhE was assessed in 2 independent founders and their activities were compared at 5dpf. Both the founders displayed largely consistent expression patterns, with a key difference in the enhancer activity seen in the olfactory pit in case of the progeny from founder1 while no detectable enhancer activity was observed in this region in case of founder2. The overlapping expression between the two founders included reporter expression in the anterior lateral line neuromasts (ML1, OC1 and ML2 in the head region and IO3 and IO4 beneath the eye), cells-associated with the neuronal projections in the ear and near the eye, and low-level activity in the hind brain which became more apparent upon dissection.

EnhE Activity in Mouse Ortholog (pIGLET Transgenesis)

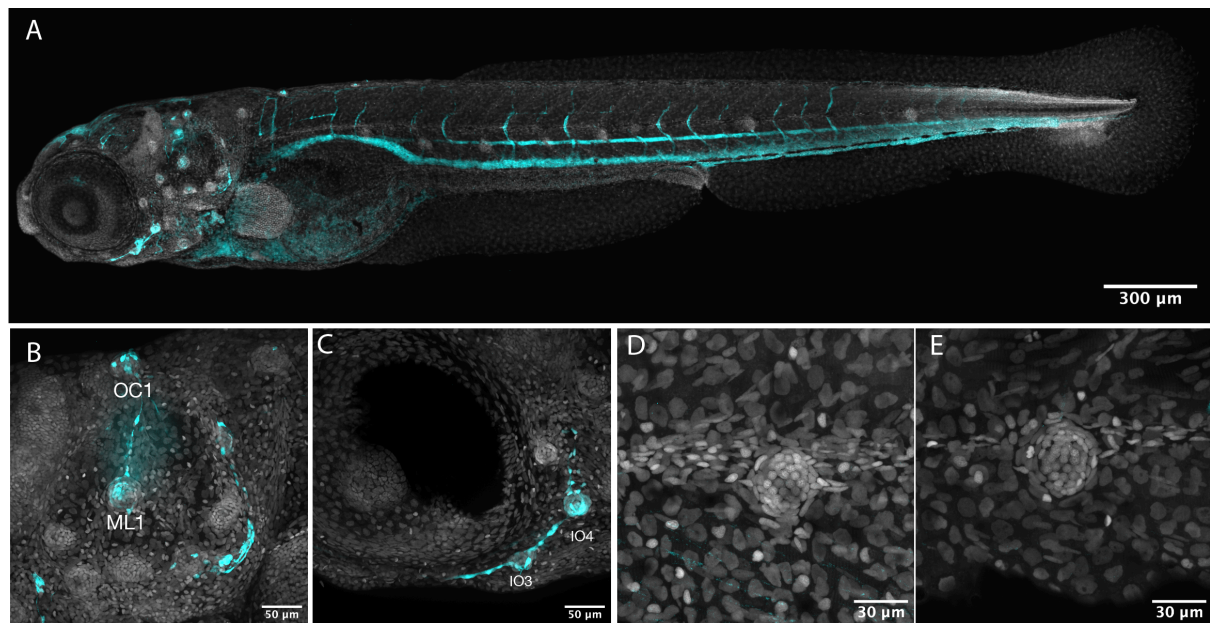


Figure 33: Mouse ortholog of EnhE (pIGLET24b, founder 1) enhancer activity at 5 dpf.

(A) Whole-embryo lateral view (10×) showing enhancer activity along the lateral line and in cranial regions. (B) High-magnification view of the ear region, including anterior lateral line neuromasts (OC1, ML1). (C) High-magnification view of the eye region, with activity in the IO3 and IO4 neuromasts. (D–E) High-magnification images of L1 and L2 neuromasts, respectively, showing enhancer activity within neuromast-associated cells.

Enhancer activity for the mouse ortholog was also assessed using a single pIGLET24b founder at 5dpf. The reporter expression pattern was largely consistent with the Tol2 founders including activity in the ear region, brain, and neuromasts beneath the eye (IO3, IO4). However, no activity was observed in the olfactory pit, aligning more closely with Tol2 founder 2. In contrast to the Tol2 lines, activity in the posterior lateral line neuromasts was minimal, with only sporadic expression in one or two cells, consistent with the reduced enhancer activity that is observed in the posterior lateral line in general.

Comparison of the Activity across Zebrafish and Mouse orthologs of EnhE

A detailed comparison is provided in Section 25.2.

Comparison of Transgenesis Systems (Tol2 vs pIGLET)

A detailed comparison between Tol2 and pIGLET systems is provided in Section 24.2.

5. EnhF

Developmental Dynamics of Enhancer Activity in Zebrafish

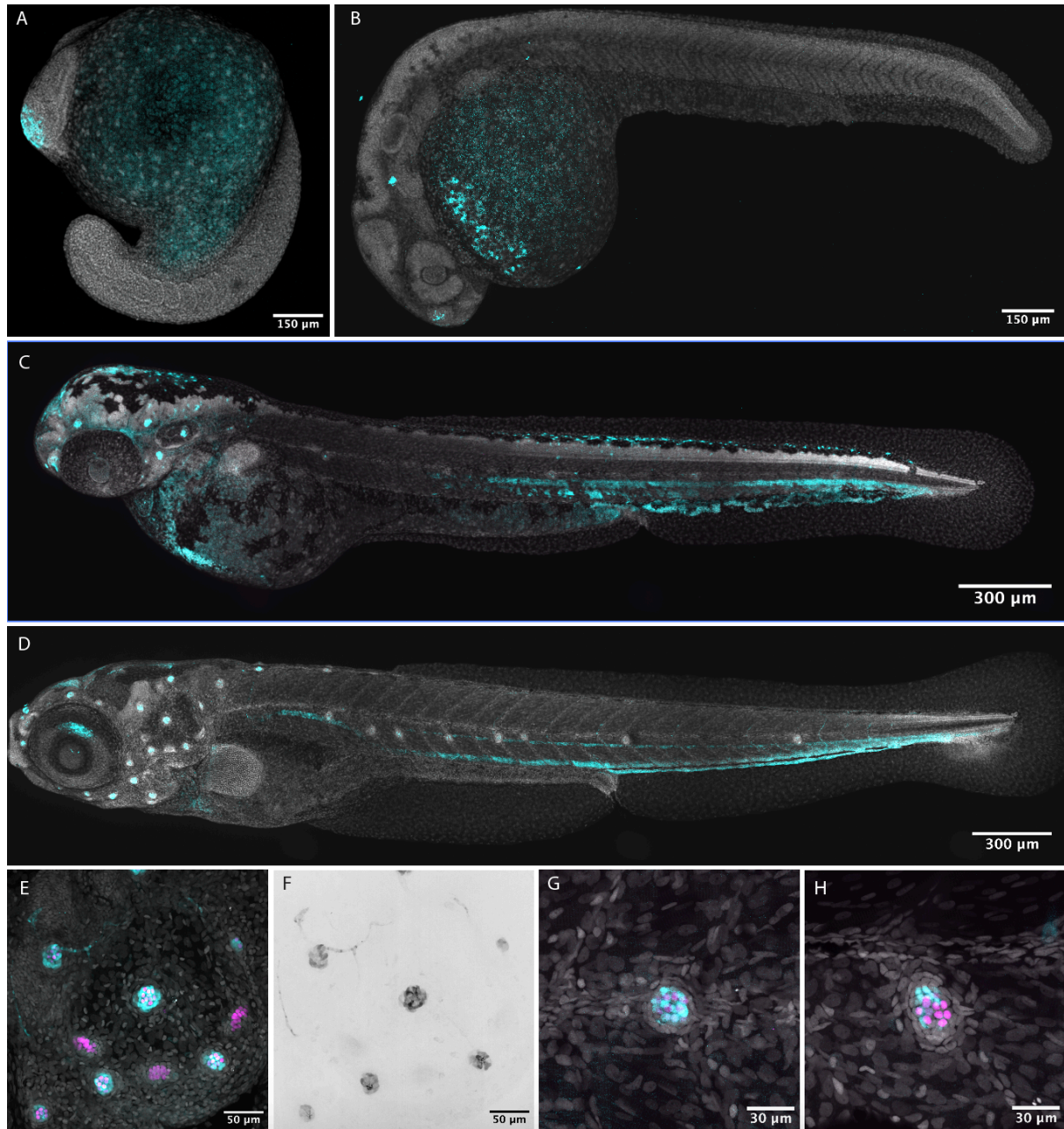


Figure 34: EnhF (Tol2, founder 1 progeny) stage-dependent activity

(A) Lateral view at ~16 somite stage, showing early enhancer activity in the olfactory epithelium. (B) Lateral view at ~32 somite stage, with activity in the olfactory epithelium, hatching gland cells, and an early deposited neuromast. (C) Whole-embryo view at 2 dpf, showing expanded activity along the lateral line system. (D) Whole-embryo view at 5 dpf, with activity predominantly in the anterior lateral line neuromasts. (E) High-magnification view of the ear region, showing GFP (cyan), myo6b-positive hair cells (magenta), and DAPI (grey). (F) Inverted GFP signal highlighting enhancer activity. (G–H) High-magnification images of L1 and L2 neuromasts, respectively, showing enhancer activity within neuromast-associated cells.

Enhancer activity for EnhF was assessed using Tol2-mediated transgenesis for the progeny from founder1 across multiple developmental stages. At ~16 somite stage (early 1 dpf) enhancer activity was observed in the olfactory epithelium. At 32 somite stage (late 1 df) the activity persisted in the olfactory epithelium and expanded to include the hatching glands and an early deposited neuromast, suggesting the onset of activity associated with lateral line. At 2 dpf, the enhancer is prominently active in the developing lateral line which persists

throughout to 5dpf, where enhancer activity is primarily seen in the anterior and posterior lateral line.

6. EnhG

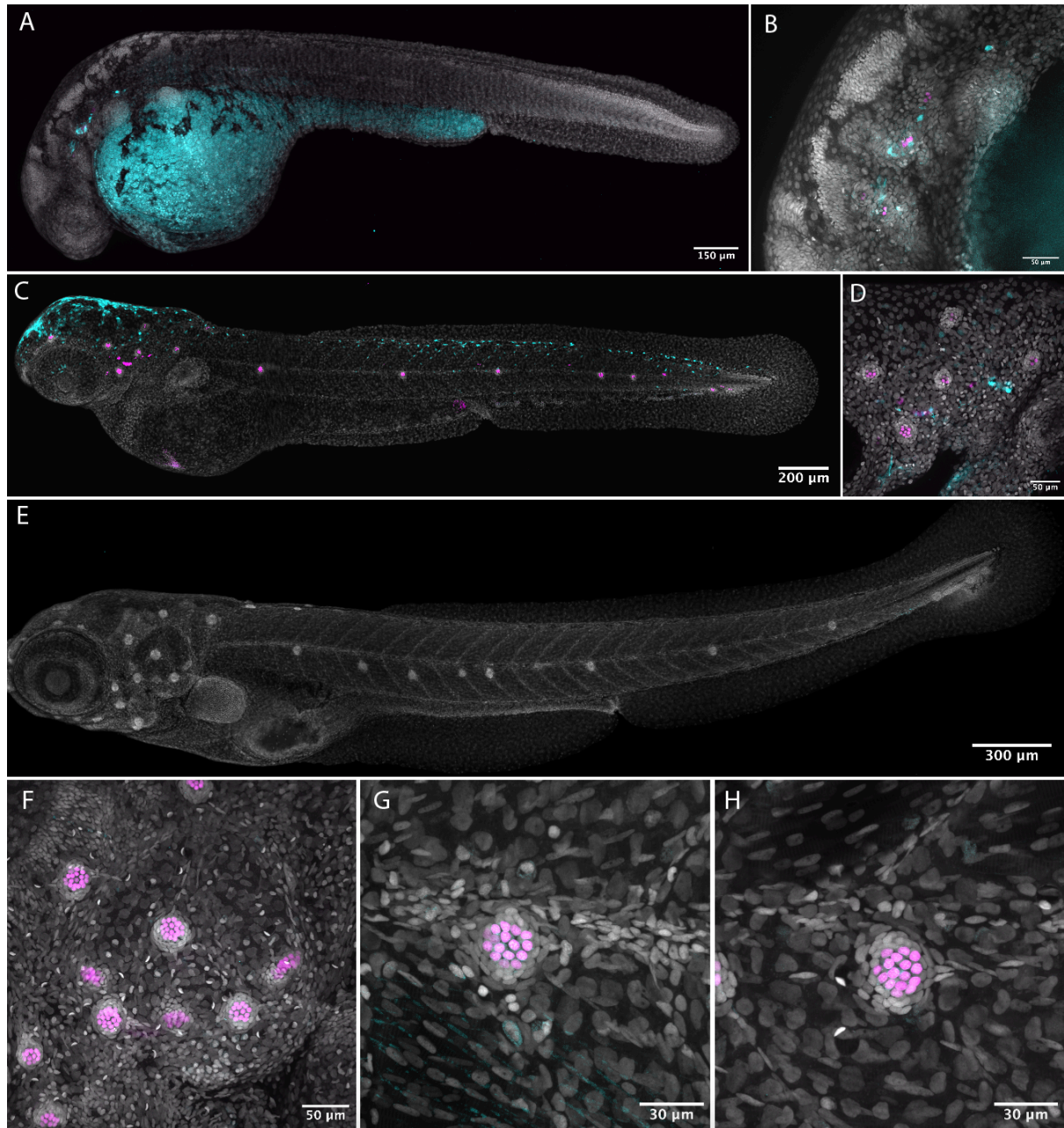


Figure 35: EnhG (Tol2, founder 1 progeny) exhibits weak and transient enhancer activity restricted to early developmental stages.

(A) Whole-embryo lateral view at 1 dpf, showing low-level enhancer activity. (B) High-magnification (40×) view of the ear region, with activity detected in the otic vesicle and a cell cluster anterior to the ear, possibly corresponding to an early deposited neuromast. (C) Whole-embryo view at 2 dpf (10×), showing reduced enhancer activity. (D) High-magnification view of the otic vesicle at 2 dpf, with sparse signal in a few cells. (E) Whole-embryo view at 5 dpf, showing no detectable enhancer activity. (F) High-magnification view of the ear region, (G–H) L1 and L2 neuromasts, respectively, all lacking detectable GFP signal. Magenta marks myo6b-positive hair cells

Enhancer activity for EnhG was assessed using Tol2-mediated transgenesis in zebrafish. Analysis was performed on an established line derived from a single founder (Founder 1) across multiple developmental stages. At 1 dpf, enhancer activity was detected sporadically in the otic vesicle, indicating weak and variable early activation in inner ear-associated regions. At 2dpf the already limited activity observed at earlier stages was observed in very

few cells within the otic vesicle. At 5dpf, no detectable enhancer activity was observed at this suggesting a decline of the activity of this enhancer over time.

7. EnhH

All the acquired data pertaining to EnhH can be found in section 24.3.

8. EnhJ

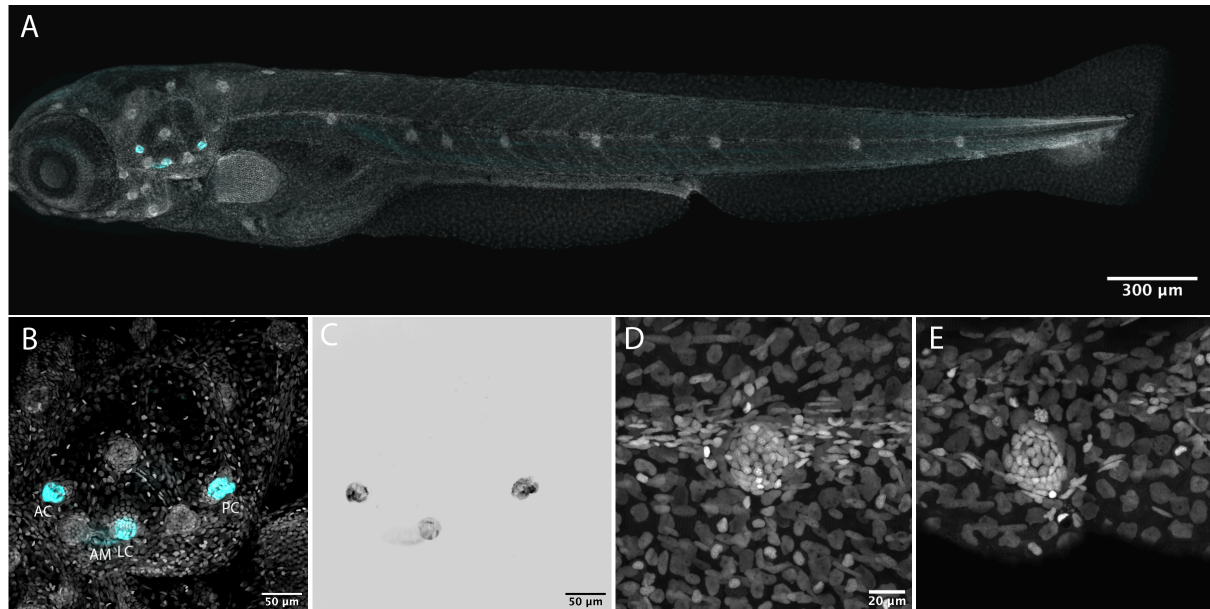


Figure 36: EnhJ (zebrafish ortholog, pIGLET14a, founder 1) exhibits inner ear-specific enhancer activity at 5 dpf.

(A) Whole-embryo lateral view (10×) at 5 dpf, showing restricted enhancer activity in cranial regions. (B) High-magnification view of the inner ear, with expression localized to the cristae (AC, PC, LC) and maculae (AM). (C) Inverted GFP signal, highlighting enhancer activity in the ear. (D–E) High-magnification images of L1 and L2 neuromasts, respectively, showing no detectable enhancer activity.

Enhancer activity for EnhJ for a single founder line (pIGLET14a, Founder 1) at 5 dpf is shown. At this stage, EnhJ exhibited clear and specific activity in inner ear structures, with expression localized to the cristae and maculae. Notably, no enhancer activity was detected in the lateral line system, including both anterior and posterior neuromasts.

Chapter 4 Discussion

Understanding the role of cis-regulatory elements in the evolution of gene expression and thereby, the phenotypic diversification is a central question in evolutionary-developmental biology. Enhancers play a critical role in regulating the spatiotemporal expression of genes; therefore, changes in enhancer activity have been proposed as an important mechanism underlying phenotypic diversification across species (Carroll 2008). Since enhancer functions in a modular fashion, evolutionary changes in enhancer sequences can alter gene expression programs without disrupting protein-coding sequences, thereby becoming flexible substrates for evolution by sequence-level mutations (Prud'homme et al. 2007). However, there are very limited studies exploring enhancer repurposing mechanistically in context of vertebrate evolution. In this study, we investigated whether the sequences identified as potential enhancers of the zebrafish lateral line by chromatin accessibility assays exhibit functional activity *in vivo* by carrying out enhancer-reporter assays and whether orthologous sequences from different species display different regulatory outputs in the same organismal context. By combining genomic analysis with transgenic enhancer-reporter assays in zebrafish, we sought to examine whether sequence differences between orthologous enhancers can lead to changes in their spatial or cell-type-specific activity.

To begin answering this question, we first identified candidate enhancer elements using previously generated ATAC-seq data, which captured chromatin regions that were open in the zebrafish lateral line. Open chromatin regions are often used as a proxy for regulatory elements as regions such as promoters, enhancers and other regulatory elements are often bound by transcription factors (Buenrostro et al. 2013). However, not all accessible regions correspond to enhancers necessarily which made an *in-vivo* validation of enhancer activity necessary. To functionally test the enhancer candidates, we employed enhancer-reporter assays by Tol2-mediated transgenesis in zebrafish, enabling efficient transgenesis and stable genomic integration of the constructs in zebrafish embryos (Kawakami 2007). This approach enabled the visualization of spatial enhancer activity in developing embryos and allowed us to assess whether ATAC-identified regions indeed function as enhancers.

Using this strategy, we found 12 enhancer candidates which could drive the reporter expression in sensory structures, particularly in the lateral line and the inner ear of zebrafish embryos. Interestingly, enhancer activity was also observed in inner ear structures which contain similar sensory hair cells detecting vestibular fluid motion. The presence of enhancer activity in both lateral line and inner ear is consistent with their shared developmental programs and functional characteristics of these mechanosensory organ systems. The function of hair cells in both organs relied on similar molecular pathways and therefore they might also have shared elements carrying out regulatory function. These findings support the idea that same enhancers, although known to be tissue-specific, can function in different tissues if they have similar functions.

While Tol2-mediated transgenesis allowed us to functionally validate our candidate enhancer elements, this system relies on random genomic integration, which can induce variability due to position effects. This is specifically undesirable for a comparison between the different orthologs of the same enhancer. Differences in integration site can influence reporter function due to the local chromatin environment or influence from other regulatory elements, inducing variability between transgenic lines with the same enhancer-reporter construct (Auer and Del

Bene 2014). To address this issue, we utilised the pIGLET transgenesis system, which allows targeted insertion of exogenous constructs. This provided us with a common landing platform for both the orthologs of our candidate enhancer elements, allowing for a better comparison of their spatial activity.

However, in some cases, the pIGLET-integrated enhancers could not drive reporter expression in the posterior lateral line. One possible explanation is that some enhancers might not be strong enough to drive reporter expression from a single genomic insertion. Alternatively, they might require inputs from additional regulatory elements to drive gene expression in their native contexts, and such elements are not present in the enhancer-reporter construct. In contrast, Tol2 consistently gave us a bright GFP signal, which often results from multiple copies of the transgene being inserted in tandem at a single genomic locus, which can lead to higher levels of GFP expression. While this feature can enhance the detectability of enhancer activity, it may also exaggerate the apparent strength of the enhancer. Because of these complementary advantages and limitations, we did not rely exclusively on either of these systems. To mitigate the potential variability arising due to position effects in Tol2 transgenesis, we analysed multiple independent founders for each enhancer construct and assessed whether they consistently produced comparable reporter expression patterns for the same enhancer. Observing similar reporter expression across different founders of the same enhancer increased confidence that the reporter expression is due to genuine enhancer activity rather than effects due to insertion sites.

Not all enhancers showed activity in the lateral line, while some showed activity only in the anterior lateral line. For example, EnhJ is active only in the inner ear structures, cristae and maculae, while EnhC shows very little activity in the anterior lateral line and no activity in the posterior lateral line. This observation suggests that although these enhancer candidates were identified from chromatin accessibility assay associated with the lateral line, an open region does not necessarily mean that it plays an active regulatory role in the tissue. Alternatively, it could also mean that while the enhancer in its native context might carry out some function in the lateral line, it is unable to drive the reporter function. These results also imply that chromatin accessibility datasets might capture elements in broadly similar sensory gene regulatory networks, even if their activity in-vivo is restricted to a particular tissue/organ type.

An important question that we wanted to address was to determine the extent to which sequence alone can encode the regulatory repurposing of orthologous enhancers. To investigate this question, we tested orthologous sequences from the mouse in zebrafish embryos. Since both the orthologs were subjected to the same developmental context of zebrafish, this experiment allowed us to isolate the contribution of sequence differences from the environmental context. Our results indicate that while the orthologous enhancers exhibit broadly similar patterns in zebrafish embryos, we could see differences in the reporter expression at different stages and in different tissues, and more subtle changes across the different neuromast cell-types. These observations suggest that sequence differences between orthologous enhancers can influence the specificity of regulatory output at multiple levels. Such differences may arise due to alterations in transcription factor binding motifs or other sequence-specific features which influence the enhancer grammar (Long et al. 2016).

These findings are also consistent with the idea that sequence changes during evolution can modulate enhancer activity without necessarily abolishing enhancer function entirely. Instead, such changes may lead to more subtle modifications leading to a regulatory tuning,

which could result in evolutionary innovation while still preserving essential developmental functions (Rebeiz et al. 2011).

It is important to note that this study focuses on characterizing the spatial activity of the candidate enhancers rather than establishing functional relationships of these enhancers with their regulatory targets. Yet, we examined the genes lying in the same topologically-associated domain as our candidate enhancers as the likely regulatory targets by looking at the overlap between domain of gene expression and enhancer activity. It is also important to note that the enhancer does not drive reporter in its endogenous genomic context. Chromatin environment, long-range interactions, and interactions with other regulatory elements can all contribute to enhancer function in vivo (Carter et al. 2002). Therefore, while our enhancer-reporter assays provide valuable information about the regulatory potential of the candidate enhancers, they represent a simplified system for studying enhancer activity.

We further examined the spatial activity of EnhC in transgenic mouse embryos to assess whether the conserved elements have been co-opted to carry out activity in novel functional domains. Although the successful insertion of enhancer-reporter construct was confirmed by genotyping, detection of endogenous GFP proved to be challenging due to high background fluorescence in the tissue sections. While we did see a localised signal in the head region of the embryo, we could not confidently assign the tissues of enhancer activity due to the overall signal-to-noise ratio. We attempted antibody staining against the endogenous GFP, which failed to reduce the background fluorescence as well, indicating that further optimization of staining and imaging conditions might be necessary.

Despite these technical limitations, we have insights into the potential regulatory function of this enhancer from a previous study (Sato et al. 2012). They demonstrated that this conserved enhancer element regulates *six1* expression and reporter activity in several cranial sensory structures in mouse embryos, including the otic vesicle. These tissues arise from cranial placodes and require the expression of *six1* gene for their proper development and function. Interestingly, our assessment of overlapping enhancer activity domains with gene expression domains facilitated by zebrafish single-cell datasets suggests that EnhC regulates the expression of the orthologs of *six1* gene as well. These results suggest that EnhC might have undergone a functional expansion or diversification during vertebrate evolution. Such changes could be reflected due to sequence changes between the zebrafish and mouse orthologs of this enhancer, which enable interactions with different transcription factors. Alternatively, it could also happen due to the distinct organismal and developmental contexts.

This study set out to answer the broad question of whether conserved enhancer candidates have been co-opted for novel function during vertebrate evolution, by studying this idea in context of the water-to-land transition of vertebrates. The tissue-level differences and more subtle cell-type differences in the spatial activity of zebrafish and mouse orthologs of the same enhancer within the same developmental context support this idea. Further, the expansion of spatial activity in the case of EnhC in the developmental context of mouse, when compared to the context of zebrafish, supports the idea that with sequence divergence and context changes during evolution, enhancers may have been repurposed during vertebrate evolution to generate variation in gene expression patterns and give rise to novel phenotypic features.

Future studies could be targeted towards establishing functional relationships between the conserved enhancers and their potential regulatory targets in zebrafish and mouse. Comparing the effects of perturbation across species could provide insights into evolutionary

conservation or divergence of the function of enhancers. In some cases, orthologous enhancers may regulate similar genes across the two species but drive spatial expression differently. In other cases, the repurposing of enhancer function led to different regulatory targets altogether. Establishing such relationships would give insights into whether repurposing of enhancers has primarily altered spatial domains of enhancer activity or whether the underlying regulatory network has been rewired as well.

Another promising direction would be to study the sequence features between the orthologous enhancers and identify differences in transcription factor binding motifs. Motif discovery and transcription factor binding prediction tools could be applied to both zebrafish and mouse enhancer sequences to identify such changes in motifs.

Together, these studies may push our descriptive understanding of enhancer co-option towards a more mechanistic understanding of how enhancer repurposing contributes to the diversification of gene expression patterns across vertebrate species.

5. References

- Auer, Thomas O., and Filippo Del Bene. 2014. “CRISPR/Cas9 and TALEN-Mediated Knock-in Approaches in Zebrafish.” *Methods* (San Diego, Calif.) 69 (2): 142–50. <https://doi.org/10.1016/j.ymeth.2014.03.027>.
- Buenrostro, Jason D., Paul G. Giresi, Lisa C. Zaba, Howard Y. Chang, and William J. Greenleaf. 2013. “Transposition of Native Chromatin for Fast and Sensitive Epigenomic Profiling of Open Chromatin, DNA-Binding Proteins and Nucleosome Position.” *Nature Methods* 10 (12): 1213–18. <https://doi.org/10.1038/nmeth.2688>.
- Cande, Jessica Doran, Vivek S. Chopra, and Michael Levine. 2009. “Evolving Enhancer-Promoter Interactions within the Tinman Complex of the Flour Beetle, *Tribolium Castaneum*.” *Development* 136 (18): 3153–60. <https://doi.org/10.1242/dev.038034>.
- Carroll, Sean B. 2008. “Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution.” *Cell* 134 (1): 25–36. <https://doi.org/10.1016/j.cell.2008.06.030>.
- Carter, David, Lyubomira Chakalova, Cameron S. Osborne, Yan-feng Dai, and Peter Fraser. 2002. “Long-Range Chromatin Regulatory Interactions in Vivo.” *Nature Genetics* 32 (4): 623–26. <https://doi.org/10.1038/ng1051>.
- Cvekl, A., and J. Piatigorsky. 1996. “Lens Development and Crystallin Gene Expression: Many Roles for Pax-6.” *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 18 (8): 621–30. <https://doi.org/10.1002/bies.950180805>.
- De Souza, F. S. J., L. F. Franchini, and M. Rubinstein. 2013. “Exaptation of Transposable Elements into Novel Cis-Regulatory Elements: Is the Evidence Always Strong?” *Molecular Biology and Evolution* 30 (6): 1239–51. <https://doi.org/10.1093/molbev/mst045>.
- Dixon, Jesse R., Siddarth Selvaraj, Feng Yue, et al. 2012. “Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions.” *Nature* 485 (7398): 376–80. <https://doi.org/10.1038/nature11082>.
- Farrell, Jeffrey A., Yiqun Wang, Samantha J. Riesenfeld, Karthik Shekhar, Aviv Regev, and Alexander F. Schier. 2018. “Single-Cell Reconstruction of Developmental Trajectories during Zebrafish Embryogenesis.” *Science* (New York, N.Y.) 360 (6392): eaar3131. <https://doi.org/10.1126/science.aar3131>.
- Galupa, Rafael, Gilberto Alvarez-Canales, Noa Otilie Borst, et al. 2023. “Enhancer Architecture and Chromatin Accessibility Constrain Phenotypic Space during Drosophila Development.” *Developmental Cell* 58 (1): 51-62.e4. <https://doi.org/10.1016/j.devcel.2022.12.003>.
- Gaunt, Stephen J. 2022. “Seeking Sense in the Hox Gene Cluster.” *Journal of Developmental Biology* 10 (4): 48. <https://doi.org/10.3390/jdb10040048>.

- Hill, Mark S., Pétra Vande Zande, and Patricia J. Wittkopp. 2021. “Molecular and Evolutionary Processes Generating Variation in Gene Expression.” *Nature Reviews. Genetics* 22 (4): 203–15. <https://doi.org/10.1038/s41576-020-00304-w>.
- Hintermann, Aurélie, Christopher C. Bolt, M. Brent Hawkins, et al. 2025. “Co-Option of an Ancestral Cloacal Regulatory Landscape during Digit Evolution.” *Nature* 648 (8092): 109–16. <https://doi.org/10.1038/s41586-025-09548-0>.
- Howe, Kerstin, Matthew D. Clark, Carlos F. Torroja, et al. 2013. “The Zebrafish Reference Genome Sequence and Its Relationship to the Human Genome.” *Nature* 496 (7446): 498–503. <https://doi.org/10.1038/nature12111>.
- Kawakami, Koichi. 2007. “Tol2: A Versatile Gene Transfer Vector in Vertebrates.” *Genome Biology* 8 Suppl 1 (Suppl 1): S7. <https://doi.org/10.1186/gb-2007-8-s1-s7>.
- Kimmel, C. B., W. W. Ballard, S. R. Kimmel, B. Ullmann, and T. F. Schilling. 1995. “Stages of Embryonic Development of the Zebrafish.” *Developmental Dynamics: An Official Publication of the American Association of Anatomists* 203 (3): 253–310. <https://doi.org/10.1002/aja.1002030302>.
- Lalonde, Robert L., Harrison H. Wells, Cassie L. Kemmler, et al. 2024. “pIGLET: Safe Harbor Landing Sites for Reproducible and Efficient Transgenesis in Zebrafish.” *Science Advances* 10 (23): eadn6603. <https://doi.org/10.1126/sciadv.adn6603>.
- Levine, Mike. 2010. “Transcriptional Enhancers in Animal Development and Evolution.” *Current Biology: CB* 20 (17): R754–763. <https://doi.org/10.1016/j.cub.2010.06.070>.
- Long, Hannah K., Sara L. Prescott, and Joanna Wysocka. 2016. “Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution.” *Cell* 167 (5): 1170–87. <https://doi.org/10.1016/j.cell.2016.09.018>.
- Lopez-Delisle, Lucille, Leily Rabbani, Joachim Wolff, et al. 2021. “pyGenomeTracks: Reproducible Plots for Multivariate Genomic Datasets.” *Bioinformatics* 37 (3): 422–23. <https://doi.org/10.1093/bioinformatics/btaa692>.
- Lumen Learning. n.d. “Eukaryotic Epigenetic Gene Regulation.” In *OpenStax Biology*.
- Lupiáñez, Darío G., Katerina Kraft, Verena Heinrich, et al. 2015. “Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions.” *Cell* 161 (5): 1012–25. <https://doi.org/10.1016/j.cell.2015.04.004>.
- McDonald, Jeanne M. C., and Robert D. Reed. 2024. “Beyond Modular Enhancers: New Questions in Cis-Regulatory Evolution.” *Trends in Ecology & Evolution* 39 (11): 1035–46. <https://doi.org/10.1016/j.tree.2024.07.005>.
- Megginson, Leon C. 1963. *Petroleum Management*. (Louisiana State University).
- Mouse Genome Sequencing Consortium, Robert H. Waterston, Kerstin Lindblad-Toh, et al. 2002. “Initial Sequencing and Comparative Analysis of the Mouse Genome.” *Nature* 420 (6915): 520–62. <https://doi.org/10.1038/nature01262>.
- Nicolson, T., A. Rüsç, R. W. Friedrich, M. Granato, J. P. Ruppertsberg, and C. Nüsslein-Volhard. 1998. “Genetic Analysis of Vertebrate Sensory Hair Cell Mechanosensation:

- The Zebrafish Circler Mutants.” *Neuron* 20 (2): 271–83.
[https://doi.org/10.1016/s0896-6273\(00\)80455-9](https://doi.org/10.1016/s0896-6273(00)80455-9).
- Nicolson, Teresa. 2005. “The Genetics of Hearing and Balance in Zebrafish.” *Annual Review of Genetics* 39: 9–22. <https://doi.org/10.1146/annurev.genet.39.073003.105049>.
- Ohmori, Yoshihiro. 2026. “Conserved Statistical Organization of Cis-Regulatory Landscapes Defines a Fluctuation-Structured Regulatory Phase.” Preprint, Genomics, February 6. <https://doi.org/10.64898/2026.02.06.704399>.
- Ohno, S. (1970). *Evolution by gene duplication*. Springer-Verlag.
- Patthey, Cedric, Gerhard Schlosser, and Sebastian M. Shimeld. 2014. “The Evolutionary History of Vertebrate Cranial Placodes--I: Cell Type Evolution.” *Developmental Biology* 389 (1): 82–97. <https://doi.org/10.1016/j.ydbio.2014.01.017>.
- Petit, Florence, Karen E. Sears, and Nadav Ahituv. 2017. “Limb Development: A Paradigm of Gene Regulation.” *Nature Reviews. Genetics* 18 (4): 245–58. <https://doi.org/10.1038/nrg.2016.167>.
- Postlethwait, J. H., Y. L. Yan, M. A. Gates, et al. 1998. “Vertebrate Genome Evolution and the Zebrafish Gene Map.” *Nature Genetics* 18 (4): 345–49. <https://doi.org/10.1038/ng0498-345>.
- Prud’homme, Benjamin, Nicolas Gompel, and Sean B. Carroll. 2007. “Emerging Principles of Regulatory Evolution.” *Proceedings of the National Academy of Sciences of the United States of America* 104 Suppl 1 (Suppl 1): 8605–12. <https://doi.org/10.1073/pnas.0700488104>.
- Rebeiz, Mark, Nick Jikomes, Victoria A. Kassner, and Sean B. Carroll. 2011. “Evolutionary Origin of a Novel Gene Expression Pattern through Co-Option of the Latent Activities of Existing Regulatory Sequences.” *Proceedings of the National Academy of Sciences of the United States of America* 108 (25): 10036–43. <https://doi.org/10.1073/pnas.1105937108>.
- Sato, Shigeru, Keiko Ikeda, Go Shioi, Kazuki Nakao, Hiroshi Yajima, and Kiyoshi Kawakami. 2012. “Regulation of Six1 Expression by Evolutionarily Conserved Enhancers in Tetrapods.” *Developmental Biology* 368 (1): 95–108. <https://doi.org/10.1016/j.ydbio.2012.05.023>.
- Shubin, Neil, Cliff Tabin, and Sean Carroll. 2009. “Deep Homology and the Origins of Evolutionary Novelty.” *Nature* 457 (7231): 818–23. <https://doi.org/10.1038/nature07891>.
- Signor, Sarah A., and Sergey V. Nuzhdin. 2018. “The Evolution of Gene Expression in Cis and Trans.” *Trends in Genetics: TIG* 34 (7): 532–44. <https://doi.org/10.1016/j.tig.2018.03.007>.
- Spence, Rowena, Gabriele Gerlach, Christian Lawrence, and Carl Smith. 2008. “The Behaviour and Ecology of the Zebrafish, *Danio Rerio*.” *Biological Reviews of the Cambridge Philosophical Society* 83 (1): 13–34. <https://doi.org/10.1111/j.1469-185X.2007.00030.x>.

- Spitz, François, and Eileen E. M. Furlong. 2012. “Transcription Factors: From Enhancer Binding to Developmental Control.” *Nature Reviews. Genetics* 13 (9): 613–26. <https://doi.org/10.1038/nrg3207>.
- Sur, Abhinav, Yiqun Wang, Paulina Capar, Gennady Margolin, Morgan Kathleen Prochaska, and Jeffrey A. Farrell. 2023. “Single-Cell Analysis of Shared Signatures and Transcriptional Diversity during Zebrafish Development.” *Developmental Cell* 58 (24): 3028–3047.e12. <https://doi.org/10.1016/j.devcel.2023.11.001>.
- Thomas, Eric D., Ivan A. Cruz, Dale W. Hailey, and David W. Raible. 2015. “There and Back Again: Development and Regeneration of the Zebrafish Lateral Line System.” *Wiley Interdisciplinary Reviews. Developmental Biology* 4 (1): 1–16. <https://doi.org/10.1002/wdev.160>.
- Thorvaldsdottir, H., J. T. Robinson, and J. P. Mesirov. 2013. “Integrative Genomics Viewer (IGV): High-Performance Genomics Data Visualization and Exploration.” *Briefings in Bioinformatics* 14 (2): 178–92. <https://doi.org/10.1093/bib/bbs017>.
- Venkataraman, Vishruth, Noel H. McGrory, Theresa J. Christiansen, Joaquin Navajas Acedo, Michael I. Coates, and Victoria E. Prince. 2025. “Development of the Zebrafish Anterior Lateral Line System Is Influenced by Underlying Cranial Neural Crest.” *Developmental Biology* 525 (September): 102–21. <https://doi.org/10.1016/j.ydbio.2025.05.025>.
- Villar, Diego, Camille Berthelot, Sarah Aldridge, et al. 2015. “Enhancer Evolution across 20 Mammalian Species.” *Cell* 160 (3): 554–66. <https://doi.org/10.1016/j.cell.2015.01.006>.
- Westerfield, M. 2000. *The Zebrafish Book*. 4th ed. University of Oregon Press, Eugene.
- Wike, Candice L., Yixuan Guo, Mengyao Tan, et al. 2021. “Chromatin Architecture Transitions from Zebrafish Sperm through Early Embryogenesis.” *Genome Research* 31 (6): 981–94. <https://doi.org/10.1101/gr.269860.120>.
- Xin, Yaqu, Yann Le Poul, Liucong Ling, et al. 2020. “Enhancer Evolutionary Co-Option through Shared Chromatin Accessibility Input.” *Proceedings of the National Academy of Sciences of the United States of America* 117 (34): 20636–44. <https://doi.org/10.1073/pnas.2004003117>.