Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht - Karls - University

Heidelberg

Presented by

M.Sc. Risa Suzuki

born in: Mie, Japan

Oral examination: 07.07.2025

Genetic Variation Shapes Visual-Motor Behavior: Insights from High-Throughput Optomotor Response Screening in Medaka Inbred Population

Referees: Prof. Dr. rer. nat. Joachim Wittbrodt
Prof. Dr. rer. nat. Lázaro Centanin

Abstract

Congenital eye diseases (CEDs) have been associated with monogenic mutations, yet the variability in clinical severity among individuals indicates a strong polygenic component.
The underlying gene-gene and gene-environment interactions causing this variability have so far been largely unexplored, owing to the inherent challenges of disentangling such interactions in human populations and the substantial sample sizes required to achieve statistical power in animal models. In this study, I made use of the Medaka Inbred Kiyosu-Karlsruhe (MIKK) panel, a collection of genetically and phenotypically diverse inbred strains of the Japanese rice fish medaka (*Oryzias latipes* and *Oryzias sakaizumi*) derived from a wild population. To dissect the genetic architecture underlying eye development and disease, I combined this resource with high-throughput behavioral phenotyping, using visuomotor behavior as a functional readout of visual system integrity.

To this end, I developed two high-throughput optomotor response (OMR) assays with an automated analysis pipeline to measure visual acuity, color and contrast sensitivity simultaneously, in 15 to 50 hatchlings at a time. Using this setup I demonstrated that visuomotor responses vary across albino mutant, inbred strains of medaka, and zebrafish, with the albino mutant exhibiting enhanced spatial, color and contrast sensitivity.
By optimizing light-induced retinal degeneration setup I successfully established, for the first time, a model of light-induced retinal degeneration in medaka hatchling, providing a novel model for investigating CEDs mechanisms. Moreover I showed that susceptibility to retinal damage varied across medaka strains and medaka in general is more resistant to retinal degeneration induced by the light than zebrafish.
A comparison of visual sensitivity and a spectrum of behavioral phenotypes (e.g. visual acuity, speed adjustment, and total distance traveled), showed great differences between the 76 tested inbred stains. Furthermore I performed functional assessments of retina in selected strains, including optokinetic response (OKR), electroretinography (ERG), and eye morphology analysis, which partially aligned with the OMR data and provided complementary evidence of visual impairments.
To uncover the genetic basis of these complex traits, I utilized strategic crosses between seven phenotypically contrasting MIKK strains and phenotyped the F2 generation. A genome-wide linkage analysis using F2 segregation has led to the identification of five quantitative trait loci (QTLs) associated with distinct visual and behavioral traits.

This work demonstrates that the automated OMR platform developed in this study can efficiently identify QTLs linked mechanisms involved in eye development, visual processing, and visuomotor integration and thereby setting the foundation to uncovering the network of causative genes contributing to the variability in congenital eye diseases.
It provides a robust framework for future studies aimed at dissecting gene-gene and gene-environment interactions, ultimately contributing to our understanding of the genetic and environmental factors driving variability in congenital eye diseases.

Zusammenfassung

Kongenitale Augenkrankheiten (CED) werden oft mit monogenen Mutationen in Verbindung gebracht, doch die Variabilität des klinischen Schweregrads bei einzelnen Personen deutet auf eine starke polygene Komponente hin.

Die zugrundeliegenden Gen-Gen- und Gen-Umwelt-Interaktionen, die diese Variabilität verursachen, sind bisher weitgehend unerforscht, da es schwierig ist, solche Interaktionen in menschlichen Populationen zu entschlüsseln da die Stichprobengröße die für statistisch aussagekräftige Verbindungen nur schwer zu erreichen sind. In dieser Studie verwende ich das Medaka Inbred Kiyosu-Karlsruhe (MIKK) Panel, eine Sammlung genetisch und phänotypisch unterschiedlicher Inzuchtstämme des japanischen Reisfisches Medaka (*Oryzias latipes* und *Oryzias sakaizumi*).

Um die genetische Architektur, die der Augenentwicklung und -erkrankung zugrunde liegt, zu entschlüsseln, kombinierte ich diese Ressource mit einer Hochdurchsatz-Verhaltensphänotypisierung, bei der das visuomotorische Verhalten als funktioneller Indikator dient. Zu diesem Zweck habe ich zwei Hochdurchsatz-Assays für die optomotorische Reaktion (OMR) mit automatisierten Analysepipelines entwickelt, um Sehschärfe, Farb- und Kontrastempfindlichkeit bei 15 bis 50 Larven gleichzeitig zu messen. Mit Hilfe dieses Systems konnte ich zeigen, dass die visuomotorischen Reaktionen bei Albino-Mutanten, Inzuchtstämmen von Medaka und Zebrafischen unterschiedlich sind.

Ich habe ein neuartiges Modell für lichtinduzierte Netzhautdegeneration bei Medaka entwickelt, um die Mechanismen von CEDs zu untersuchen. Damit konnte ich zeigen, dass die Anfälligkeit für Netzhautschäden bei Medakastämmen variiert und dass Medaka im Vergleich zu Zebrafischen eine größere Resistenz gegen Netzhautdegeneration aufweisen.

Der Vergleich der visuellen Empfindlichkeit und einiger weiterer relevanter Merkmale (z. B. Sehschärfe, Geschwindigkeitsanpassung und zurückgelegte Gesamtstrecke) zeigte große Unterschiede zwischen den 76 getesteten MIKK-Stämmen. Die funktionellen Untersuchungen der Netzhaut ausgewählter Stämme (optokinetische Reaktion (OKR), Elektroretinographie (ERG) und Augenmorphologie) lieferten ergänzende Hinweise auf Sehstörungen. Um die genetische Grundlage dieser komplexen Merkmale aufzudecken, habe ich strategische Kreuzungen zwischen sieben phänotypisch kontrastierenden

MIKK-Stämmen vorgenommen und die F2-Generation phänotypisiert. Eine genomweite Kopplungsanalyse unter Verwendung der F2-Segregation führte zur Identifizierung von fünf quantitativen Merkmalsloci (QTLs), die mit unterschiedlichen Seh- und Verhaltensmerkmalen assoziiert sind.

Diese Arbeit zeigt, dass die hier verwendete automatisierte OMR-Plattform effizient QTLs identifizieren kann, die mit Mechanismen verbunden sind, die an der Augenentwicklung, der visuellen Verarbeitung und der visuomotorischen Integration beteiligt sind, und damit die Grundlage für die Aufdeckung des Netzwerks der ursächlichen Gene schafft, die zur Variabilität bei angeborenen Augenerkrankungen beitragen. Die Studie bietet einen solide Grundlage für künftige Studien, die darauf abzielen, die Wechselwirkungen zwischen Genen sowie zwischen Genen und Umwelt zu untersuchen, und trägt damit letztlich zu unserem Verständnis der genetischen und umweltbedingten Faktoren bei, die die Variabilität bei angeborenen Augenerkrankungen bestimmen.

Abbreviations

AC	amacrine cell
AMD	age-related macular degeneration
ASD	anterior segment dysgenesis
BC	bipolar cell
CED	congenital eye disorder
CCW	counter-clockwise
CMZ	ciliary marginal zone
CRISPR	clustered regularly interspaced palindromic repeats
CW	clockwise
dpf	days post fertilization
dph	days post hatch
ERG	electroretinography
ERM	embryo rearing medium
FGF	fibroblast growth factor
GC	ganglion cell
GCL	ganglion cell layer
GWAS	genome-wide association study
HC	horizontal cell
INL	inner nuclear layer
IRD	inherited retinal disease
LD	linkage disequilibrium
LGN	lateral geniculate nucleus
LIRD	light induced retinal degeneration
MIKK panel	Medaka Inbred Kiyosu-Karlsruhe panel
MG	Müller glia
NFL	nerve fiber layer
OKR	optokinetic response
OMR	optomotor response
ONL	outer nuclear layer
POAG	primary open-angle glaucoma
PRC	photoreceptor cell

Shh	sonic hedgehog
SNP	single nucleotide polymorphism
TGF-β	transforming growth factor beta
QTL	quantitative trait loci
RGC	retinal ganglion cell
ROI	region of interest
ROS	reactive oxygen species
RPC	retinal progenitor cell
RP	retinitis pigmentosa
RPE	retinal pigment epithelium
TUNEL	terminal deoxynucleotidyl transferase dUTP nick end labeling

Contributions

The following people contributed to this work as outlined:

2.1 Development of a High-Throughput Assays

The linear-pool-style OMR arena and setup was built by **Gero Hofmann** and LED light was installed by **Niklas Grammling**. Stripe generating software was coded by **Jun Shern Chan**. **Jia Zheng Woo** performed pilot experiments for testing multiple stripe conditions which were later used in this study, under the supervision of **Tinatini Tavhelidse-Suck** and **Thomas Thumberger. Ian Brettell** gave great input in setting up the pipeline for fish tracking, image processing and analysis.

The prisms and the infinity-pool-style OMR setup was built by **Gero Hofmann**, the monitor was reassembled by **Niklas Grammling**. The stripe generating software was coded by **Damjan Kalšan**. Fish detection software pipeline was developed by **Damjan Kalšan** with input from **Risa Suzuki** and **Thomas Thumberger**. **Saul Pierotti** contributed to the initial analysis.

For light induced retinal degeneration setup, homemade dish and experimental box was built by **Gero Hofmann** and the homemade pump was built by **Niklas Grammling**.

2.3 MIKK panel screening for optomotor response

Egg collection was supported by the efforts of **Beate Wittbrodt**, **Marzena Majewski**, **Rebecca Lipp**, and **Tanja Kellner**.

2.4 Crossing scheme and further phenotyping

Jingjing Zang (UZH) conducted electroretinogram (ERG) on medaka and zebrafish.

2.6 Association Testing and QTL Mapping

Mireia Osuna Lopez (GeneCore EMBL) prepared the sequencing library for the F2 samples and carried out the sequencing at Genomics Core Facility at EMBL.

Esther Yoo (EMBL-EBI) performed imputation, generated the relatedness matrix, performed the association testing, and mapping.

Table of Contents

1

Introduction

Congenital eye disorders (CEDs) occur in approximately once in every 2,500 live births in Europe (Maillet et al., 2024). Although some of these conditions are linked to monogenic mutations with clear symptom complexes, they frequently display considerable phenotypic variability, even among individuals within the same family carrying the same pathogenic mutation (Harding & Moosajee, 2019).

This variability suggests the involvement of broader genetic interactions, such as modifier effects and epistasis, as well as potential environmental factors that contribute to the disorder's severity. Understanding this complexity is critical for unraveling the full genetic architecture of congenital eye disorders as well as the developmental processes underlying the visual system.

This thesis investigates the genetic basis of quantitative trait complexity with a focus on the visual system, using the teleost fish *Oryzias latipes* (medaka), a well-established vertebrate model. Specifically, this work employs a genetically diverse medaka inbred panel to associate distinct genotypes with phenotypic variation in visual traits. By leveraging natural variation across divergent genetic backgrounds along with high-throughput visual function assay, genetic crosses and large-scale sequencing enabled the identification of segregating loci linked to visual system phenotypes. This systems genetics approach establishes a powerful experimental framework for dissecting the polygenic basis of visual system development and congenital eye disorders, offering new insights into the genotype-to-phenotype relationship in a vertebrate model.

1.1 Vertebrate Visual System

Vision is a critical function for survival in many animal species, allowing them to detect and interpret environmental stimuli and in turn respond with behaviors such as escaping from predators, capturing prey, or coordinating positions with conspecifics. Vertebrates move in

reaction to visual cues through a series of neurological processing steps that mainly involve the retina, different brain regions for interpretation and computation, and finally the spinal cord and motor pathways for a coordinated response.

In vertebrates, the pathway starts with the light being detected by the photoreceptor cells (PRCs) in the outer retina. Then, PRCs generate electric signals, which are transmitted to the bipolar cells (BCs) and ganglion cells (GCs) in the retina. The signals are further transmitted through the axons of the GCs as the optic nerves to the brain.

In mammals, there are two visual pathways. In the geniculate pathway, axons of retinal ganglion cells (RGCs) project to lateral geniculate nucleus (LGN) in the thalamus (Schneider, 1969; Stepniewska et al., 2000). Axons of those neurons of the LGN go to the primary visual cortex then to the higher visual cortices. This pathway is crucial for computation of high-resolution visual information such as form, color, and motion vital for higher order cognitive functions.

In the extrageniculate pathway, RGC axons project to the superior colliculus in the midbrain. From there to lateral posterior nucleus-pulvinar complex in thalamus then to higher order visual cortices (Forrester et al., 2020). The extrageniculate system facilitates rapid reflexive responses to visual stimuli such as potential threats like predators (Q. V. Le et al., 2013). Once a decision to move is made, motor commands are initiated in the motor cortex and transmitted via descending pathways to the spinal cord, activating motor neurons that control muscle contractions. The cerebellum and basal ganglia in the striatum play modulatory roles, refining movement coordination and execution (Auer et al., 2024).

In non-mammalian vertebrates, such as fish and amphibians, the visual processing architecture differs. RGC axons project directly to the optic tectum which integrates visual information and coordinates appropriate motor response. The optic tectum, which is functionally analogous to the mammalian superior colliculus, serves as the center for visual processing (Nevin et al., 2010). From the optic tectum, signals are transmitted to various brain regions, including the pretectum, tegmentum, cerebellum, and reticular formation (Baier & Scott, 2024; Heap et al., 2017). In zebrafish, it has been reported that RGC axons also project to the pretectum in the diencephalon directly, these neurons in turn projects to cerebellum, tegmentum, reticular formation, and hypothalamus that regulate locomotor patterns (Matsuda & Kubo, 2021).

1.2 Eye Structure

In the visual pathway, the eye serves as the entry point, where light is first received and converted into neural signals. The eye is highly conserved in its development, structure, and cellular composition across vertebrates. The anterior segment of the eye comprises the cornea, lens, iris, ciliary body, and the tissues of the iridocorneal angle. Those are the structures that collectively regulate and facilitate the entry and focusing of light onto the retina (Ahsanuddin & Wu, 2023; Richardson et al., 2017). The posterior segment of the eye consists of the vitreous body, the neural retina and the retinal pigment epithelium (RPE), and the posterior choroid and sclera. Cumulatively these structures work together to detect light and protect the retina. The neural retina contains photoreceptor cells (PRCs) that initiate visual signaling, while the RPE absorbs excess light to minimize scattering, supports PRC health and provides protection against photooxidative damage from natural light (Strauss, 1995; Yang et al., 2021).

The retina consists of seven main cell types in three nuclear layers: rod and cone PRCs, three cell types of interneurons, retinal ganglion cells (RGCs), and Müller glia (MG) cells. PRC nuclei are located in the outer nuclear layer, the most apical layer. The nuclei of all three interneurons (horizontal cells: HC, bipolar cells: BC, and amacrine cells: AC) and MG cells are contained in the inner nuclear layer. RGCs and displaced ACs are located in the ganglion cell layer (Centanin & Wittbrodt, 2014). The outer retina is principally responsible for phototransduction, where PRCs (rods and cones) convert light into electrochemical signals, while subsequent layers made up of BCs, HCs, and ACs integrate and modulate these signals (Centanin & Wittbrodt, 2014). The RPE and choroid, although not neuronal, provides critical metabolic support and participates in the recycling of visual pigments; such interactions are essential for maintaining the health and function of PRCs. MG cells are the main glial cell type in the retina and contribute to the maintenance of retinal homeostasis and visual function (Devoldere et al., 2019).

The retina plays a central role in determining which visual signals are transmitted to the optic nerve. Opsins, that are light-sensitive proteins, along with chromophores expressed in PRCs define which wavelengths of light trigger neural responses. The density and distribution of photoreceptors fundamentally determine visual acuity, in conjunction with optical and refractive quality, which is shaped by physical elements such as the lens, cornea, and

photoreceptor mosaic (Querubin et al., 2009). While certain types of RGCs are known for their role in regulating circadian rhythms, emerging evidence also suggests that they contribute to enhancing contrast sensitivity (Chien et al., 2023).

1.3 Eye Development

Eye formation begins early in embryogenesis and follows the same basic developmental and organisational principles across vertebrates (Diacou et al., 2022; Sinn & Wittbrodt, 2013). It is initiated by the specification of a single eye field in the anterior neural plate (Wilson & Houart, 2004), where a coordinated network of transcription factors (Zuber et al., 2003) establishes retinal progenitor cell identity. This primordium splits into two lateral domains, each of which evaginates to form an optic vesicle (Harding & Moosajee, 2019). The surface ectoderm adjacent to the optic vesicle thickens to form the lens placodes which subsequently delaminates to form lens vesicle or solid lens rudiment (Harding & Moosajee, 2019; Richardson et al., 2017). Optic vesicle forms the bilayerd optic cup (Heermann et al., 2015). The inner layer of the cup differentiates into the neural retina, while the outer layer becomes the RPE, both of which constitute the posterior segment of the eye (Sinn & Wittbrodt, 2013). The RPCs of the neural retina differentiate into the above outlined differentiated neuronal cell types, driven by, among others, Fibroblast Growth Factor signaling, while the RPCs of the presumptive RPE differentiate driven by Wnt/beta-catenin signaling. The anterior segment of the eye originates from the surface ectoderm and surrounding periocular mesenchyme (Richardson et al., 2017).

In sum, the eye arises from multiple distinct embryological tissues and develop through a tightly regulated process involving coordinated upregulation and downregulation of genes and transcription factors, such as OTX2, SOX2, SIX3, PAX6, RAX (Rx2 and Rx3 in medaka) (Harding & Moosajee, 2019; Sinn & Wittbrodt, 2013). The division of the eye field into optic primordia is orchestrated by secreted signaling molecules from the transforming growth factor beta (TGF-β), fibroblast growth factor (FGF), and Sonic hedgehog (Shh) pathways (Carl & Wittbrodt, 1999; Sampath et al., 1998; Sinn & Wittbrodt, 2013). The subsequent evagination of the optic vesicles is regulated by the members of the Retinal homeobox (Rx, RAX) transcription factor family (Loosli et al., 2003; Mathers et al., 1997; Rojas-Muñoz et al., 2005; Winkler et al., 2000).

1.4 Congenital Eye Disorders (CEDs)

Various defects during the developmental process of the eye can result in different types of congenital eye disorders (CEDs) in humans. CEDs such as anophthalmia, microphthalmia, coloboma, anterior segment dysgenesis, and inherited retinal degenerative disorders occur at varying frequencies and recent studies indicate a polygenic origin of its phenotypic diversity (Ford & Petersen-Jones, 2025; N.-Q. Le et al., 2024).

Anophthalmia refers to complete absence of ocular tissue, whereas microphthalmia is the condition of a small eye with all to some ocular tissues missing (Verma & Fitzpatrick, 2007). Recent studies suggest that the occurrence of these malformations could be as high as 30 per 100000 individuals (Verma & Fitzpatrick, 2007). Over 90 genes including developmental regulators such as SOX2, OTX2, RAX, VSX2, and PAX6, have been described to be associated (Harding & Moosajee, 2019). Heterozygous loss-of-function mutations in SOX2 are one of the most frequent causes of bilateral anophthalmia or extreme microphthalmia, accounting for an estimated 15-40% of such cases (Harding & Moosajee, 2019). PAX2 mutations on the other hand cause renal coloboma syndrome, a condition with a symptom complex encompassing both renal and ocular malformations (Bower et al., 1993).

However, each individual gene typically explains only a small fraction of cases, reflecting extensive genetic heterogeneity (for example, no conclusive disease-causing variants were identified in 8 out of the 15 index patients (53%) (Haug et al., 2021). Strikingly, even when a known pathogenic mutation is present, phenotypes can vary widely among individuals, ranging from unilateral microphthalmia to bilateral anophthalmia in families with the same mutation (Harding & Moosajee, 2019).

Overall, despite the many genes discovered for anophthalmia and microphthalmia, a molecular diagnosis cannot be made in up to 20-80% of patients (depending on severity and laterality), implying that numerous causative genes and genetic interactions remain to be identified (Haug et al., 2021). This unexplained portion likely involves a combination of individually rare mutations and polygenic risk factors that current gene tests have yet to capture (Haug et al., 2021).

Anterior Segment Dysgenesis (ASD) encompasses developmental disorders of the anterior part of the eye, including cornea, iris, lens, and trabecular meshwork. The overall prevalence of congenital cataract was in the range from 0.63 to 9.74 per 10000 (Sheeladevi et al., 2016). Overall 255 genes (such as FOXC1 and PAX6) are known to be associated, placing the total the diagnostic success rate is about 50% (Reis et al., 2024). High phenotypic variability

within and between families are reported, indicating modifiers and involvement of rare or more common variants in genes in the same pathway.

Inherited retinal diseases (IRDs) are genetic disorders affecting retinal development or maintenance, leading to vision loss, most frequently due to the progressive degeneration of PRCs (Hanany et al., 2020). Unlike the above mentioned diseases, many IRDs follow Mendelian inheritance, with over 270 disease genes identified to date for various forms of IRDs (Chen et al., 2021). The prevalence of retinitis pigmentosa was 1:4,000 with recessive causative mutation being responsible for 50-60% of cases and ~65% of cases causing mutations being identified (Hanany et al., 2020). Interestingly digenic inheritance has been reported, for retinitis pigmentosa caused by concurrent heterozygous mutations in PRPH2 (peripherin-2) and ROM1, where each mutation alone is insufficient and only double heterozygotes cause retinitis pigmentosa (Kajiwara et al., 1994; T. R. Lewis et al., 2023). Thus, even in the domain of hereditary retinal disorders, there is a growing recognition of polygenic influences and epistatic (gene by gene) interactions.

Investigating the polygenic contributions to CEDs in humans poses several challenges, particularly due to phenomena such as variable expressivity (the severity and features of the disease can differ widely among individuals with the same mutation) and incomplete penetrance (not everyone who inherits CED-related mutations will manifest the disease) (Harding & Moosajee, 2019). Further complexity arises from genetic heterogeneity, where mutations in different genes can lead to similar phenotypes, and pleiotropy, where a single gene can influence multiple traits (Mackay & Anholt, 2024). Epistasis, in which the effect of one gene is modified by the presence of others, adds yet another layer of complexity to genetic analyses, making it difficult to trace clear gene to gene relationships in CEDs (Mackay & Anholt, 2024). These factors complicate the prediction of disease onset and severity, even among individuals carrying the same pathogenic variant.

On top of these complex genetic factors, environmental influences also contribute to the complexity of CEDs, interacting with genetic factors to affect disease onset, severity, and variability. A range of environmental risk factors has been implicated in the development of CEDs. These include maternal infections (Salomè et al., 2023), alcohol consumption during pregnancy (Tandon & Mulvihill, 2009), and exposure to teratogenic substances (Dubucs et al., 2024). Additional maternal factors, such as elevated pre-pregnancy body mass index (BMI) (Lee et al., 2021), smoking during pregnancy (Santiago-Colón et al., 2020), and the

use of certain medications (Dubucs et al., 2024), including anti-inflammatory drugs and antibiotics have also been associated with an increased risk of ocular anomalies. Together, these environmental influences can disrupt critical stages of eye development, either independently or in combination with underlying genetic susceptibilities.

1.5 Strategies to Study Genetic Architectures of Complex Phenotype

To understand genetic mechanisms of complex phenotypes such as CEDs, it is necessary to associate different genetic variants with various effect sizes to a trait of interest. Quantitative trait loci (QTL) mapping is a statistical method used to identify genomic regions that contribute to variation in quantitative traits, which are continuously measurable phenotypes such as plants height (Wu et al., 2022), body weight (Delpero et al., 2024) or social behavior in mouse (Knoll et al., 2018). The core principle of QTL mapping is to correlate phenotypic variation with molecular markers distributed across the genome, enabling researchers to infer the number, location, and effect sizes of loci involved in complex traits.

In the classical design, QTL mapping begins with the selection of genetically and phenotypically divergent inbred strains for crossing (Mackay & Anholt, 2022). The resulting F1 generation consists of genetically uniform, heterozygous individuals with phenotypes typically intermediate between the parental strains. Intercrossing F1 individuals produces a recombinant F2 population, where each individual inherits a unique combination of haplotype blocks due to meiotic recombination. These blocks serve as the basis for mapping statistical associations between genetic markers and trait variation, allowing the identification of QTLs. This controlled approach maximizes genetic contrast and minimizes environmental noise, making it particularly effective for dissecting the genetic architecture of polygenic traits. There is a growing trend toward utilizing more complex crosses with an expanded number of founder strains (Milner et al., 2016; Scott et al., 2020).

Genome-wide association studies (GWAS) explore statistical associations between phenotypes and genome-wide SNP data across many individuals. GWAS leverages linkage disequilibrium (LD), the non-random association of alleles between nearby loci, to identify genomic regions linked to traits (Uffelmann et al., 2021). Although recombination events gradually erode LD over generations, such events are infrequent between closely spaced single nucleotide polymorphisms (SNPs). Consequently, LD tends to be stronger between

nearby markers and declines with increasing physical distance. Unlike linkage mapping with F2 populations, GWAS benefits from historical recombination accumulated over generations, increasing mapping resolution without the need to generate crosses. GWAS in human patients on primary open-angle glaucoma (POAG) uncovered at least 127 independent risk loci (Singh et al., 2023). However, GWAS require large sample sizes, which poses a challenge for studying rare conditions such as congenital eye diseases.

To identify causal genetic elements beyond naturally occurring variants, forward genetics approaches using mutagenesis have long been utilized in various model organisms (E. B. Lewis & Bacher, 1968). Chemical mutagenesis with ENU (N-ethyl-N-nitrosourea) has been utilized in various models such as mice (Acevedo-Arozena et al., 2008), zebrafish (Haffter et al., 1996), and medaka (Furutani-Seiki et al., 2004; Taniguchi et al., 2006). Additional mutagens used in forward genetic screens include EMS (ethyl methanesulfonate**),** widely used in *Drosophila* (St Johnston, 2002) and *Arabidopsis* (Koornneef & Meinke, 2010), which induces primarily G/C to A/T transitions (W. Yan et al., 2021). Radiation-based mutagenesis such as X-rays or gamma rays, which typically cause larger chromosomal deletions and rearrangements (Shima & Shimada, 1988). These tools have been instrumental in discovering key developmental regulators and gene networks (Nüsslein-Volhard & Wieschaus, 1980) (Acevedo-Arozena et al., 2008).

A limitation of chemically or radiation-induced mutagenesis is that it often results in multiple concurrent mutations within a single organism (Acevedo-Arozena et al., 2008), and many of which may not reflect natural variation. These challenges are particularly pronounced when investigating polygenic traits, where the interplay between multiple genes contributes to the phenotype. The presence of numerous background mutations can introduce artifacts, making it difficult to discern the true genetic architecture underlying complex traits, as it is crucial to consider the cumulative and interactive effects of multiple genes within their natural contexts (Mackay, 2004).

In contrast, reverse genetics approaches begin with a known gene and aim to uncover its function by targeted disruption or modification. Clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 system, uses a guide RNA to direct the Cas9 nuclease to a specific DNA sequence, creating a double-strand break that results in targeted gene disruption or mutation during repair or knock-in sequences via non-homologous end joining or homology-directed repair (Agarwal et al., 2025; Chowdhury et al., 2022; Doering et al.,

2023; Gücüm et al., 2021; Gutierrez-Triana et al., 2018). Sequences can also be knocked-in this way by providing so-called donor DNA sequences that are then integrated during HDR (Gutierrez-Triana et al., 2018). Engineered variations of the Cas9 enzyme can even fulfill related but different genetic editing actions. Base editors directly convert one DNA base to another (Cornean et al., 2022; Pakari et al., 2023). Prime editors can insert or swap short DNA sequences without requiring donor templates (Caso & Davies, 2022). Dead Cas9 (dCas9) fused to effectors can regulate gene expression or modify epigenetic marks such as DNA methylation at specific loci (Espino-Saldaña et al., 2020).

To dissect the genetic basis of complex traits, a combination of approaches is essential (Gierten et al., 2025; Welz et al., 2025). On the other hand QTL mapping and GWAS provide powerful methods for identifying candidate loci associated with phenotypic variation, particularly when applied to populations with natural genetic diversity (Gierten et al., 2025; Welz et al., 2025). However, these approaches require high-throughput phenotyping and benefit significantly from the use of inbred panels, which offer reproducibility and genetic stability. Inbred panels, collections of genetically distinct yet homozygous strains, allow for the systematic exploration of phenotypic diversity linked to underlying genetic variation (Gierten et al., 2025; Welz et al., 2025). Reverse genetics complements these approaches by enabling direct functional validation of candidate genes and testing of gene-gene or gene-environment interactions (Welz et al., 2025).

Together, trait mapping, inbred panels, and targeted genome editing form a comprehensive toolkit for understanding the genetic architecture of complex traits. Such panels have been successfully established in a variety of model organisms, including Arabidopsis thaliana (Atwell et al., 2010), Drosophila (Mackay et al., 2012), and mouse (Ashbrook et al., 2021; Dumont et al., 2024), highlighting their broad utility across species.

1.6 Medaka

1.6.1 Vertebrate Model with Long History of Genetic Research

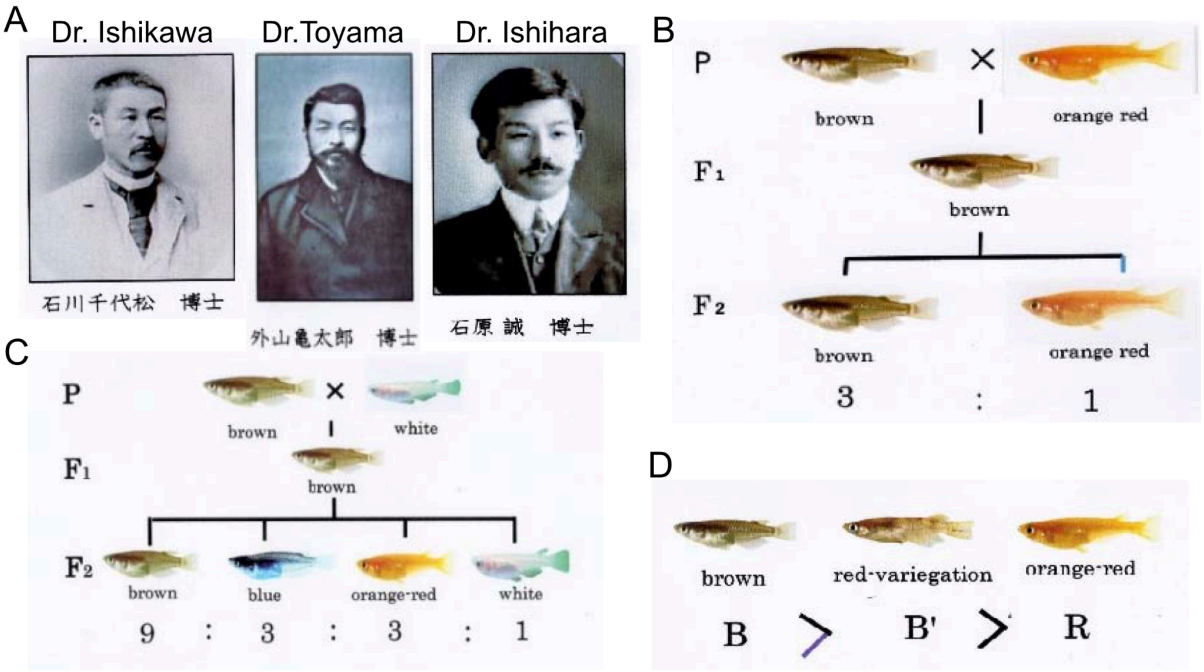


Figure 1.1 History of Genetic Studies with Medaka.
A, Pioneer researchers who used medaka as a model organism. From the left, Dr. Chiyomatsu Ishikawa, Dr. Kametaro Toyama, and Dr. Makoto Ishihara. (B-D) Body color in Medaka shows Mendelian inheritances reported by Dr. Ishikawa (B), by Dr. Toyama (C), and Ishihara (D). Adapted from the figures originally generated by Tetsuro Takeuchi (*Life of My Teacher Tatsuo AIDA*, 2020)

Medaka (*Oryzias latipes*) is a well-established vertebrate model used in diverse fields such as developmental biology, evolution, ecology, neuroscience, and biomedical research. The earliest recorded medaka research dates back to 1912, when Dr. Chiyomatsu Ishikawa conducted genetic experiments involving a cross between brown-bodied and orange-red-bodied medaka (Ishikawa, 1912). All F1 offspring displayed the brown phenotype, while the F2 generation exhibited a 3:1 segregation of brown to orange-red individuals (Figure 1.1) (Ishikawa, 1912). A few years later, Dr. Kametaro Toyama demonstrated that brown body color is dominant over orange-red, and orange-red is dominant over white. His studies revealed a 3:1 segregation in the F2 generation, and mating brown and white individuals produced all brown F1 offspring, with the F2 segregating into brown:blue:orange-red:white in a 9:3:3:1 ratio. Notably, all white individuals were reported to be female (Toyama, 1916). In the same year, Dr. Makoto Ishihara reported that brown is dominant over red variation, which in turn is dominant over orange-red (Ishihara, 1916).

Medaka naturally inhabit a wide geographical range across China, Korea, Taiwan, and Japan. Their habitats span a variety of environments, including freshwater (e.g., canals, ponds, lakes, and rice fields) and saltwater regions (Iwamatsu, 2018). Genetic studies have revealed significant divergence between Northern and Southern Japanese medaka populations. These populations not only exhibit large genetic distances but also display distinct schooling behaviors when coexisting in the same river systems. As a result, the Northern medaka has been classified as *Oryzias sakaizumii* (Asai et al., 2011), while the Southern counterpart retains the name *Oryzias latipes*.

Medaka is particularly well-suited for large-scale genetic screenings due to several advantageous traits: a compact genome (~700 Mb), small adult body size (~4 cm), high fecundity, daily spawning of externally developing transparent eggs, short developmental and maturation periods (2-3 months), ease of maintenance, broad tolerance to temperature and salinity variations, low breeding costs, and the availability of multiple inbred and mutant strains. (Iwamatsu, 2004; Kasahara et al., 2007; Kirchmaier et al., 2015; Wittbrodt et al., 2002).

Moreover, medaka exhibit a remarkably high tolerance to inbreeding, which enabled the development of the first near-isogenic vertebrate population panel (Fitzgerald et al., 2022; Leger et al., 2022; Spivakov et al., 2014).

1.6.2 Medaka Inbred Panel Kiyosu Karlsruhe (MIKK)

Several classical medaka strains, originating from different latitudes have undergone over 100 generations of inbreeding resulting in genetically stable lines (Hyodo-Taguchi & Egami, 1985; Kirchmaier et al., 2015; Spivakov et al., 2014). Among these are the southern-derived strains HdrR, HO5, and Cab, and the northern-derived strain Kaga, representing distinct geographic origins within the medaka population. Building on this foundation, the first near-isogenic population resource in a vertebrate model was established from wild medaka collected in Kiyosu, near Toyohashi City in Aichi Prefecture, Japan. This resource, known as the MIKK panel, comprises over 80 inbred lines and serves as a powerful tool for population genomics (Fitzgerald et al., 2022; Leger et al., 2022).

Across the MIKK panel, a total of 3,001,493 genetic variants, including SNPs and insertions or deletions, compared to the HdrR reference genome were identified (Fitzgerald et al., 2022). 70% (2,248,228) of these variants are located in intergenic or synonymous regions . Functional annotation revealed 644,509 non-synonymous variants, 36,444 splice site variants,

and 82,312 predicted loss-of-function mutations, including premature stop codons and frameshifts (Fitzgerald et al., 2022). Following nine generations of inbreeding, more than 70% of the MIKK lines reached homozygosity across over 70% of their genomes. However, chromosome 1 consistently exhibited lower levels of homozygosity across all lines, reflecting the presence of the sex-determination region, which remains heterozygous in males (Fitzgerald et al., 2022). These patterns of genetic variation and homozygosity have important implications for GWAS and QTL mapping with LD across the MIKK panel revealed genomic blocks with a minimum size of 12.5 kb, providing a framework for high-resolution trait mapping (Fitzgerald et al., 2022).

One of the key advantages of medaka as a model organism is its compatibility with genome editing technologies. Modern tools such as CRISPR-Cas and base editors enable precise functional interrogation of candidate genes and causal variants (Cornean et al., 2022; Pakari et al., 2022). Furthermore, more than 70% of human genes have orthologs in medaka, and nearly all major human organ systems have functional counterparts in teleosts (Kasahara et al., 2007; Sakamaki et al., 2007), greatly enhancing the translational relevance of discoveries. Consequently, the MIKK panel, combined with the ease of genome editing in medaka, offers a unique opportunity to advance QTL mapping and uncover the genetic architecture of complex visual traits, including congenital eye diseases (CEDs) and visual system development.

1.7 Fish Model to Study Eye Development

Fish have been extensively utilized as model organisms to investigate eye development, disease, function, and regeneration due to their genetic tractability, optical transparency, and regenerative capacity (Loosli et al., 2004; Lust & Wittbrodt, 2018; Sinn et al., 2014). Along with transient gene knockdown methods like morpholino, the advent of genome editing tools, such as CRISPR/Cas9, has enabled precise manipulation of gene expression and generation of targeted mutants and transgenic reporter lines to model human eye diseases and eye development (Gücüm et al., 2021; Winkler et al., 2000). For example, pax6 mutants in zebrafish and medaka have been used to mimic congenital eye disorders (Mikula Mrstakova & Kozmik, 2024; Takamiya et al., 2015). Random mutagenesis screens in zebrafish have helped identify genes critical for visual system development and function (Brockerhoff et al., 1995; S. C. F. Neuhauss, 2010). Research using cavefish has demonstrated the essential role of the lens in supporting retinal cell survival, highlighting evolutionary adaptations in eye

structure and function (Strickler et al., 2007). Moreover, fish provide a unique advantage for regeneration studies: while zebrafish possess a remarkable ability to regenerate all retinal cell types, medaka exhibit a more restricted regenerative response, making them valuable for comparative analyses of regenerative mechanisms (Lust & Wittbrodt, 2018)

1.8 Assays to Evaluate Visual Functions and Mimic Retinal Degeneration

1.8.1 Optomotor Response (OMR) Assay

The optomotor response (OMR) is a behavioral reflex in which an animal moves its head or body in response to a moving visual stimulus. It has been widely used to assess visual function in mice and various fish models (Nonarath et al., 2025; Shi et al., 2018). In a typical OMR assay, the subject is placed in an arena surrounded by moving stripes (often black and white) (Shi et al., 2018). When the striped pattern is rotated, the animal turns in the direction of motion, effectively stabilizing its visual scene (Schlegel & Neuhauss, 2020). OMR is a robust innate behavior indicating functional vision and visuomotor circuitry and elicited purely by visual cues. Neural circuits underlying OMR involve retinal processing of motion signals and downstream activation of motor pathways to adjust swim direction, via the optic tectum or pretectal neurons that detect global optic flow in fish (Matsuda & Kubo, 2021). In zebrafish, it has become increasingly common to present the visual stimulus from below by placing the arena on top of a screen or monitor as this orientation has been shown to elicit a stronger optomotor response (Matsuda & Kubo, 2021). Zebrafish (Danio rerio) larvae develop OMR behaviour as early as vision matures from 5 days post fertilization (dpf), and it has been exploited in large-scale genetic screens. For example, Neuhauss et al. (1999) screened hundreds of zebrafish mutants using an OMR-based visual behavior test, identifying numerous loci important for visual responses (S. C. Neuhauss et al., 1999).

Medaka (*Oryzias latipes*) have also been studied with OMR. Carvalho et al. (2002) developed an optomotor testing system for medaka spanning larval to adult stages. Using a rotating striped drum, they recorded both the swimming (OMR) and eye movement with optokinetic responses (OKR) of medaka to moving patterns, and quantitatively measured visual performance improves significantly during the development after hatch. In medaka, OMR has typically been elicited using rotating stripes surrounding the fish, and the response has been most reliably observed from 5 days post-hatch (dph) onward (Carvalho et al., 2002; Furukawa & Ijiri, 2002; Imada et al., 2010). Although early-stage larvae can detect motion,

(when I started this study), consistent OMR behavior had been reported to usually take at least 10 days after hatch to develop (Carvalho et al., 2002).

One limitation of the OMR assay is its throughput. Traditional setups, while effective, are low-throughput and time-consuming. Neuhauss et al. (1999) placed up to 100 zebrafish larvae above a flat monitor displaying moving stripes, allowing rapid identification of visually impaired individuals based on their failure to track motion (S. C. Neuhauss et al., 1999). However, the designs are optimized for detecting severe vision impairments and do not allow for the assessment of individual visual performance, nor do they eliminate confounding factors such as interactions between larvae.

Other attempts to increase throughput include the use of 24-well plate formats (Brastrom et al., 2019), though challenges remain in terms of the robustness of behavioral readouts and the implementation of high-throughput analysis pipelines downstream of data acquisition. The challenges with behavioural readouts include difficulty in distinguishing true OMR behavior from random movement due to the small arena size. Furthermore, receiving the stimulus applied from below the fish usually hinders automation of fish detection, which in turns further limits the assay's throughput.

While suitable for identifying complete blindness with low to medium throughput, such state-of-the-art setups may not reliably capture subtle visual phenotypes or individual differences, particularly when not designed for group comparisons.

1.8.2 Optokinetic Response (OKR) Assay

The optokinetic response (OKR) is a reflexive eye movement that occurs when an animal's entire visual field moves. In an OKR assay, the subject is typically held stationary (often immobilized), and a moving pattern of stripes is presented around it. The animal's eyes will automatically rotate to follow the drifting pattern (slow phase), then flick back in the opposite direction (fast phase) once the eyes reach their movement limit. This reflex stabilizes the image on the retina during whole-field motion and is mediated by neural circuits connecting the retina to oculomotor nuclei, often via the accessory optic system or pretectum (Schlegel & Neuhauss, 2020).

OKR is a widely used assay in zebrafish larvae, where it provides a sensitive readout of visual ability. Zebrafish develop a robust OKR by about 4 dpf, coincident with the maturation of the retinal circuitry (Schlegel & Neuhauss, 2020). For example, Brockerhoff's laboratory established an OKR-based screening protocol to identify zebrafish visual mutants. Each

larva's OKR can be measured in about one minute, allowing many individuals to be tested in succession (Rodwell et al., 2023). Medaka fish also exhibit an OKR, though this has been less intensively exploited than in zebrafish. In the medaka visual function study by Carvalho et al., both optomotor and optokinetic responses were recorded from larval and adult medaka using the same rotating stripe apparatus. The number of eye oscillations (Crouzier et al., 2022) or the eye velocity (Mueller et al., 2011) can quantify the OKR strength. However, because OKR inherently requires individual handling of each fish, including mounting them one by one in methylcellulose and adjusting their orientation. Although each step is seemingly minor, it demands ongoing manual intervention. Consequently, someone must be present throughout the processes, making OKR lower-throughput compared to OMR as well as more prone to systematic errors introduced by the experimenter with each handling of the fish.

1.8.3 Electroretinogram (ERG)

Electroretinography (ERG) is a classical physiological assay that measures the electrical response of the retina to light stimulation. An ERG is the recording obtained, which is typically a trace of voltage over time, when a flash of light or a visual stimulus is presented to the eye. The ERG reflects summed field potentials from retinal cells in vivo (Chrispell et al., 2015). In an experiment, an electrode is placed on the cornea of the eye and a reference electrode is placed on the body or in the surrounding medium. When a flash of light is delivered, it evokes a characteristic biphasic waveform: in vertebrates, a negative-going a-wave originating from photoreceptors (rods/cones) followed by a positive b-wave generated largely by ON-bipolar cells. The amplitude and timing of these ERG components provide insight into retinal function, for instance, reduced a-wave implies photoreceptor dysfunction, while an isolated a-wave with no b-wave suggests inner retinal (bipolar cell) defects. Under anesthesia, larval zebrafish at 5-7 dpf can yield clear ERG readings, and notably at this age only cone photoreceptors are functional (rods develop later) (Chrispell et al., 2015).

1.8.4 Light-Induced Retinal Degeneration (LIRD) Models

Light-induced retinal degeneration (LIRD) refers to experimental paradigms where intense light exposure is used to damage the photoreceptors in the retina, thereby creating a model of retinal degeneration. Such models are valuable for studying diseases like retinitis pigmentosa (RP) or age-related macular degeneration (AMD), as well as for probing the capacity of the

retina to regenerate after injury. In essence, prolonged or high-intensity light overstimulates photoreceptor cells (rods and cones), leading to their stress and death via apoptosis. This approach has been used in rodents for decades and has also been adapted to fish models (Thomas & Thummel, 2013). Photoreceptor injury is commonly induced using continuous light exposure, typically LED light for 48 hours, or through short-term exposure to intense light sources such as lasers or mercury lamps. These methods trigger photoreceptor apoptosis and have been widely applied in rodent and zebrafish models.

However, to date, there is only one published report of their application in medaka and that study focused exclusively on adult fish exposed to UVA light (Sayed et al., 2019). In this study, UVA exposure induced retinal degeneration in both wild-type and *p53*-deficient adult medaka. After 30 minutes of UVA exposure, a reduction in cell populations was observed, particularly in the ganglion cell layer (GCL), in both genotypes. Following 60 minutes of exposure, vacuolation in the nerve fiber layer (NFL) and further loss of cells in the inner nuclear layer (INL), outer nuclear layer (ONL), and GCL were evident in wild-type. However in this study, the time course evaluation of degenerative response assessment as well as quantification of apoptotic cell evaluation was lacking. To our knowledge, no studies have applied this method to medaka hatchlings, and there are no reports on light exposure effect on medaka retina and specifically assessing photoreceptor-specific effects or apoptotic cell responses at this early developmental stages.

1.9 Aims and Approaches

The objective of this thesis was to characterize the genetic complexity of ocular/visual traits linked to congenital eye disorders (CEDs) by integrating high-throughput behavioral screening with genetic mapping techniques. The main aims were to 1) develop high-throughput visiomotor and retinal degeneration assays; 2) evaluate retinal functions across the population panel; 3) conduct segregation and mapping analysis to pinpoint loci associated with visual trait variation.

1) **Development of high-throughput optomotor (OMR) and retinal degeneration assays for functional evaluation of the visual system.**
   - **a)** Establishment of high-throughput OMR assay with automated stripe control and integrated tracking and analysis.
   - **b)** Development of a light-induced retinal degeneration assay in medaka, enabling strain- and species-level comparisons of susceptibility.
   - **c)** OMR screening across four inbred medaka strains hatchlings, an albino mutant, and zebrafish, using varying stripe width, color and contrast.
   - **d)** Characterization of blind behavior in eyeless medaka following retinal removal.
   - **e)** Visuomotor responses assessed across developmental stages to evaluate visual maturation, trainability, and behavioral diversity.
2) **Evaluation of retinal functions across the selected medaka strains.**
   - **a)** Measurement of eye sizes across strains to investigate correlations with visual performance traits.
   - **b)** Investigation of optokinetic response (OKR) to evaluate visual tracking and motion detection capabilities.
   - **c)** Electroretinography (ERG) to assess photoreceptor and retinal circuit function across medaka strains.
3) **Conduction of segregation and genetic mapping analyses to identify loci associated with visual trait variation.**
   - **a)** Screening of OMR across 74 MIKK strains and phenotypically diverse MIKK strains for 11 pairwise crosses to generate F2 mapping populations.
   - **b)** Integration of phenotypic data with whole-genome sequencing to identify quantitative trait loci (QTLs) linked to visual behavior.

2

Results

2.1 Development of High-Throughput Assays for Screening Visual Function and Retinal Degeneration

Optomotor response (OMR) has been widely used to assess visual function and behavior in teleost species, including medaka. To investigate the developmental spectrum of the visual system, identify potential impairments, and elucidate underlying genetic mechanisms, OMR serves as a robust and scalable platform. In the context of F2 phenotyping and QTL mapping, utilizing animals at the earliest possible developmental stage is advantageous, as investigations at larval stages enable larger sample sizes and minimizes environmental variability such as differences in diet, tank density, temperature, and infection, that can significantly impact eye development and morphology. (Cleymaet et al., 2022). Therefore, performing OMR at the hatchling stage is ideal for performing F0 and F2 screening and elucidating the genetic contributions to visual system phenotypes.

However, a high-throughput OMR assay that supports large sample sizes allowing phenotyping up to thousands of individuals and incorporates an efficient, automated analysis pipeline was unavailable when I began this study. Most medaka studies utilizing OMR to date had focused on juvenile and adult stages with lower throughput (Carvalho et al., 2002; Matsuo et al., 2018; Mizoguchi et al., 2023). Furthermore, these studies have primarily employed traditional drum-style OMR assays, in which fish are placed in a tank within a rotating drum displaying moving stripes with limitations in scalability and throughput. Additionally, previous reports suggested that medaka do not consistently exhibit robust OMR until several days post-hatching (typically between 5 and 10 days, depending on the study)(Carvalho et al., 2002; Furukawa & Ijiri, 2002; Imada et al., 2010), leaving the early developmental trajectory of visuomotor function largely uncharacterized.

To address these gaps, I have developed two high-throughput OMR setups: one delivering visual stimuli from below and the other from the side using prisms. Both OMR setups will be described in more detail in the following sections.

2.1.1 Development of Linear-Pool-Style OMR Assay with Visual Stimuli from Below

The data presented in the following chapter was published in PLOS One (Suzuki et al., 2024). For a detailed description of author contributions refer to “Contributions”.

2.1.1.1 Linear-Pool-Style OMR Setup

First, I developed a linear-pool-style OMR assay, enabling high-throughput testing of individual hatchlings for 15 individual hatchlings simultaneously. In this setup, hatchlings were placed in a translucent linear-pool-style arena and OMR is triggered by moving stripes displayed on a monitor underneath (Figure 2.1 A). Using the open-source Fish Stripes software (Suzuki et al., 2024), stripe parameters such as width, speed, color, and contrast can be precisely controlled (Figure 2.1 A,B). Videos are recorded from above with a CellCam Centro 200MR camera, equipped with an infrared (IR) filter and IR illumination. Image sequences were processed using a custom image analysis macro (as described in Materials and Materials and Methods section), enabling semi-automated analysis and reducing the manual workload required for behavioral quantification. The setup allows experiments with 15 hatchlings at a time, each placed in separate lanes of the linear pool (Figure 2.1 C,D).

Hatchlings were acclimated in the linear-pool-arena for 2 hours before exposure to moving stripes. Each trial consisted of four sequential phases: stationary stripes (5 seconds), rightward-moving stripes (10 seconds), stationary stripes (5 seconds), and leftward-moving stripes (10 seconds). Prior to testing different stripe conditions, hatchlings were further acclimated with stationary stripes for 10 minutes (Figure 2.1 E).

2.1.1.2 Robust OMR Triggered with Black and White Stripes From Below in 0-1 dph Medaka

To determine whether stripe stimuli from below could induce an OMR in freshly hatched Cab strain medaka (0-1 days post-hatch, prior to self-feeding), that is commonly used as wildtype strain, I exposed 15 hatchlings to black and white stripes (16 mm width) at three different speeds: 4.3, 6.5, and 8.6 mm/s (Figure 2.2 C). Additionally, Cab hatchlings were tested with stripe widths of 8, 16, and 20 mm, at a fixed speed of 6.5 mm/s (Figure 2.2 D). The medaka Cab strain exhibited a strong OMR across all stripe speeds and widths tested. The median response rate ranged from 66.7% (16 mm, 4.3 mm/s) to 100% in all other conditions (Figure 2.2).

For evaluating visual function, one key measure is visual acuity, which is determined by the ability to distinguish fine spatial details, such as the width of moving stripes.

Therefore, responses to progressively thinner stripes (raging from 2.8 mm to 1.6 mm) were analyzed to assess the threshold at which hatchlings could no longer recognize the pattern as distinct stripes, indicating their visual resolution. Since hatchlings showed similar response to all stripe widths at speeds above 6.5 mm/s, stripe speed of 6.5 mm/s, was selected for subsequent experiments to determine the minimum stripe width detectable, serving as a measure of visual acuity.

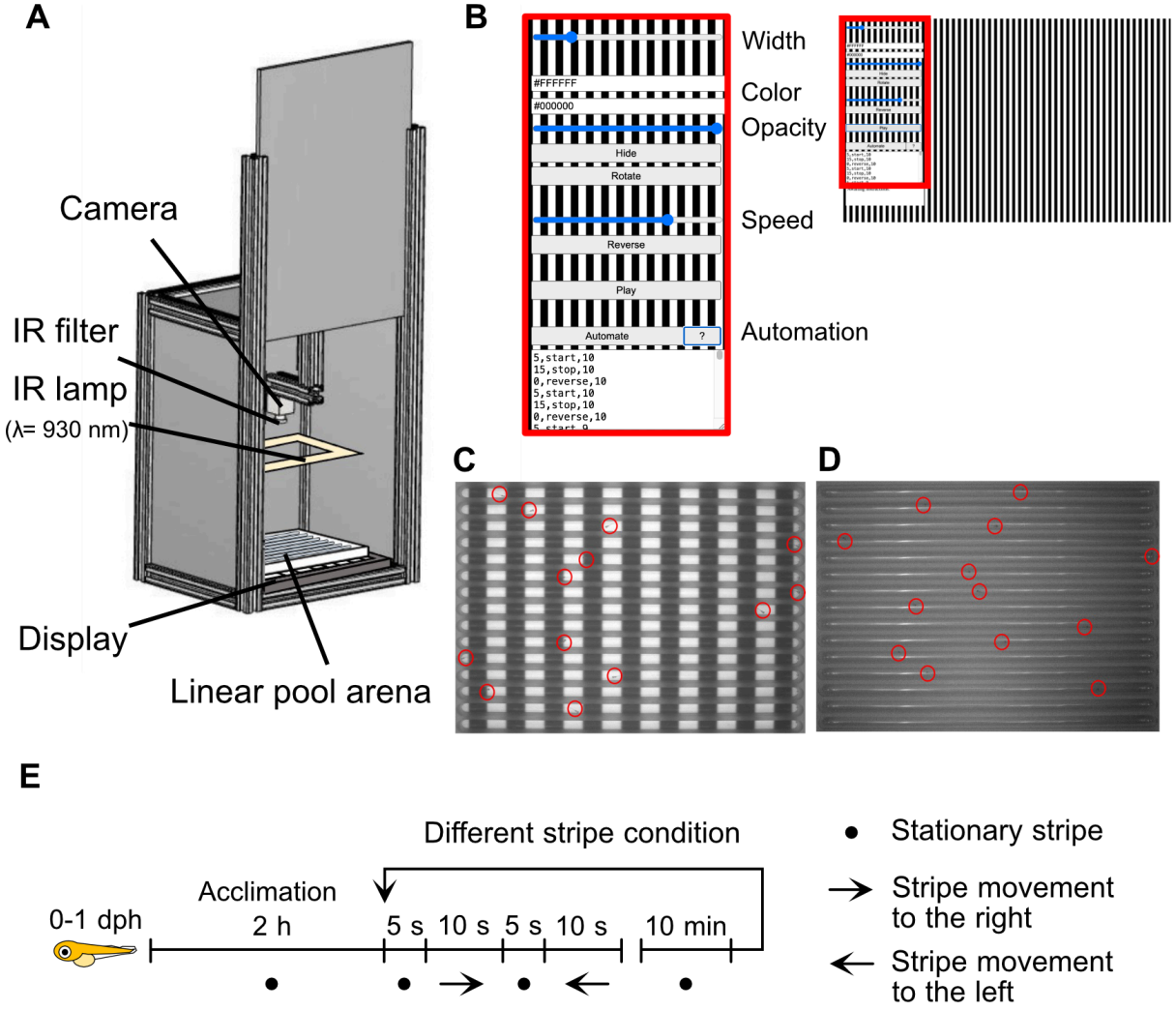


Figure 2.1 The Linear-Pool-Style Optomotor Response (OMR) Setup.

(A) Schematic representation of the linear-pool-style OMR setup: The setup consists of a PVC box containing a display screen (e.g. a tablet) with the 140 mm x 110 mm linear-pool-arena on top. Video acquisition is performed using a CellCam Centro 200MR camera, equipped with an infrared (IR) filter and illuminated by an IR lamp. The box is closed on all sides to eliminate external visual stimuli from the surrounding. (B) Stripe pattern displayed on the tablet screen. Stripe parameters: width (in pixels), stripe color 1 and stripe color 2 (defined in hex code), contrast ("opacity") (x 100%) and speed (x 50 pixel/s), are set and automated using the Fish Stripes software. (C, D) Images of the linear-pool-arena positioned over a tablet screen displaying black and white stripe pattern. The images were captured

using the camera without (C) and with IR filter (D). The lanes of the linear-pool-arena are oriented at 90° relative to the stripe movement direction. The hatchlings are indicated with red circles. (E) Experimental timeline. 0-1 days post hatch (dph) medaka were acclimatized for 2 hours in the experimental setup with stationary stripes (filled circle) shown on the display. The OMR experiment then followed a sequence of 5 seconds of stationary stripes, 10 seconds of rightward-moving stripes (arrow to right), another 5 seconds of stationary stripes, and 10 seconds of leftward-moving stripes (arrow to left). Before testing different stripe conditions, hatchlings underwent an additional 10-minute acclimation period with the respective stationary stripe pattern. Adapted from (Suzuki et al., 2024).

To ensure sufficient space for movement in response to the visual stimuli, only hatchlings positioned at least 27 mm (20% of the lane length) from the end of the lane were included. A hatchling was classified as responsive if it swam at least 27 mm (20% of the lane length) in the direction of stripe motion. Those that met this criteria in at least three out of four stripe exposures were considered for response rate calculations (Figure 2.2 B).

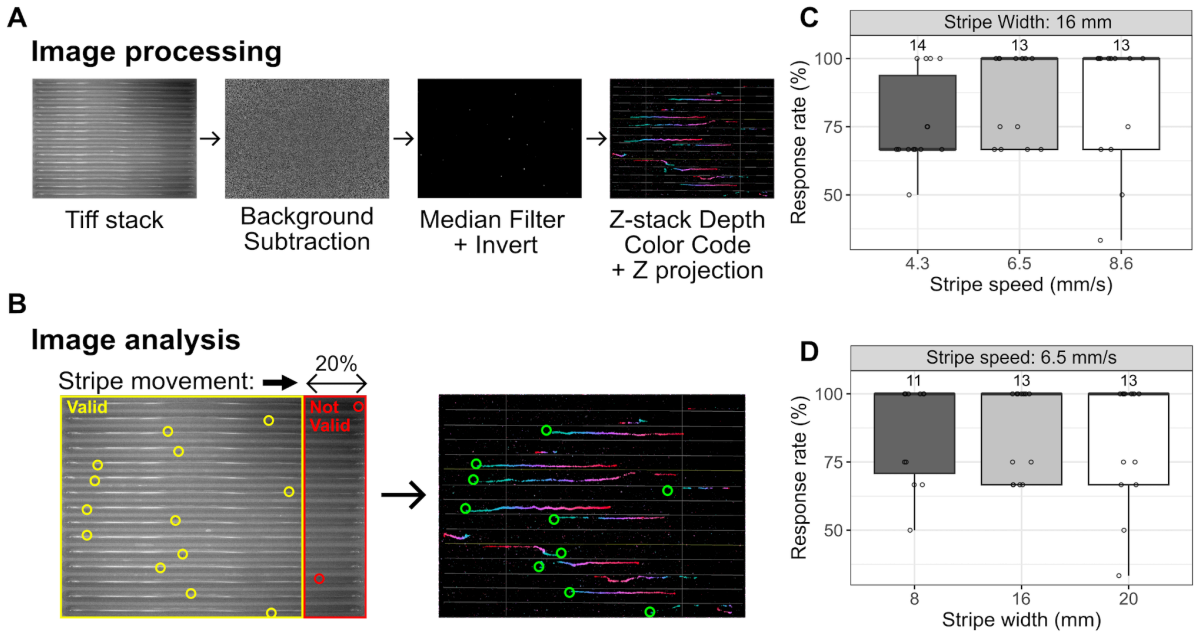


Figure 2.2 Medaka Cab Strain Exhibits a Strong Optomotor Response (OMR) to Stimulation from Below.
(A) The TIFF stack was processed using Fiji with the following steps: the first slice was duplicated and subtracted from each subsequent slice to remove background. A median filter was then applied for noise reduction, followed by image inversion. To visualize larval swimming trajectories, the Depth Color Code plugin and Z projection were used. (B) For image analysis, the first slice of the stack was utilized to classify larvae as either valid or non-valid for response rate calculations. Larvae located more than 27 mm from the end of the lane (equivalent to 20% of the lane length) were classified as valid (yellow circles), while those near the lane end were labeled as non-valid (red circles). Z projection images were used to determine whether larvae were responsive or non-responsive. hatchlings that swam in the direction of stripe motion for at least 20% of the lane length were

classified as responsive (green circles). The arrow indicates the direction of stripe movement. (C) Response rates were assessed at three different stripe speeds while maintaining a constant stripe width of 16 mm. (D) Response rates were evaluated across three different stripe widths, with a fixed stripe speed of 6.5 mm/s. The number displayed above each boxplot indicates the total number of hatchlings included in the response rate calculation. Cab strain medaka hatchlings exhibited a robust response to black-and-white stripe patterns across a speed range of 4.3 to 8.6 mm/s and stripe widths between 8 and 20 mm. Statistical analysis was conducted in RStudio using pairwise t-tests with Bonferroni correction. Adapted from (Suzuki et al., 2024).

2.1.1.3 Training Improves Sensitivity to Narrow Stripes

To establish the optimal conditions and define the screening parameters that allow phenotyping of the entire MIKK panel for differences in OMR, I examined both the trainability of Cab and HdrR hatchlings, which is another wildtype strain, in response to stripe width through repeated exposure.

Each group, consisting of 13-15 larvae, was presented with moving stripes of varying widths: 2.8, 2.4, 2.0, and 1.6 mm. The Cab strain hatchlings demonstrated a stronger response to the wider stripes of 2.8, 2.4, and 2.0 mm, than the HdrR strain and for both strains, responsiveness declined as stripe width decreased. Cab hatchlings responded at median rates of 66.7% for 2.8 mm, 16.7% for 2.4 mm, 29.2% for 2.0 mm, and 0.0% for 1.6 mm. In contrast, HdrR larvae exhibited a 50.0% response at 2.8 mm, but majority of HdrR hatchlings failed to respond to narrower stripes resulting in median response rate of 0.0%.

To test whether OMR response to stripe width is trainable, I exposed 13-15 hatchlings per strain to 16 mm-wide stripes more than 10 times before presenting narrower stripes. After “training”, Cab strain hatchlings showed a modest improvement in response rate to narrower stripes after training (median: 66.7% for 2.4 mm), but not apparent change for 2.8 mm (66.7%), 2.0 mm (29.2%), or 1.6 mm (0.0%) (Figure 2.3).

By contrast, HdrR strain hatchlings exhibited a significant improvement in response to thin stripes after training, with median response rates of 75.0% (2.8 mm), 66.7% (2.4 mm), 50.0% (2.0 mm), and 33.3% (1.6 mm) (Figure 2.3). Notably, HdrR hatchlings outperformed Cab strain hatchlings at 2.0 and 1.6 mm, suggesting strain-dependent differences in trainability.

2.1.1.4 OMR Response to Colored Stripes Varies Among Medaka Strains

To determine whether color sensitivity can be assessed using the linear-pool-style OMR setup, I exposed 15 hatchlings per strain (Cab, HdrR, and three inbred strains from MIKK panel) to moving stripes (16 mm width, 6.5 mm/s) of different colors: black/white, blue/white, green/white, and red/white (Figure 2.4).

Cab strain hatchlings responded robustly across all colors, with a median response rate of 100% in all conditions. HdrR strain hatchlings showed similarly high responses (Figure 2.4). In contrast, the three MIKK inbred strains exhibited reduced sensitivity to specific colors. Hatchlings from strains IP10-1 and IP69-1 showed lower response rates to blue/white (75%) and red/white (87.5%), while IP11-2 strain hatchlings exhibited even lower response to blue/white (70.8%) and red/white (75.0%). Strikingly, IP11-2 also showed decreased responsiveness to green/white (66.7%), whereas IP10-1 and IP69-1 responded strongly to green/white (100%) (Figure 2.4).

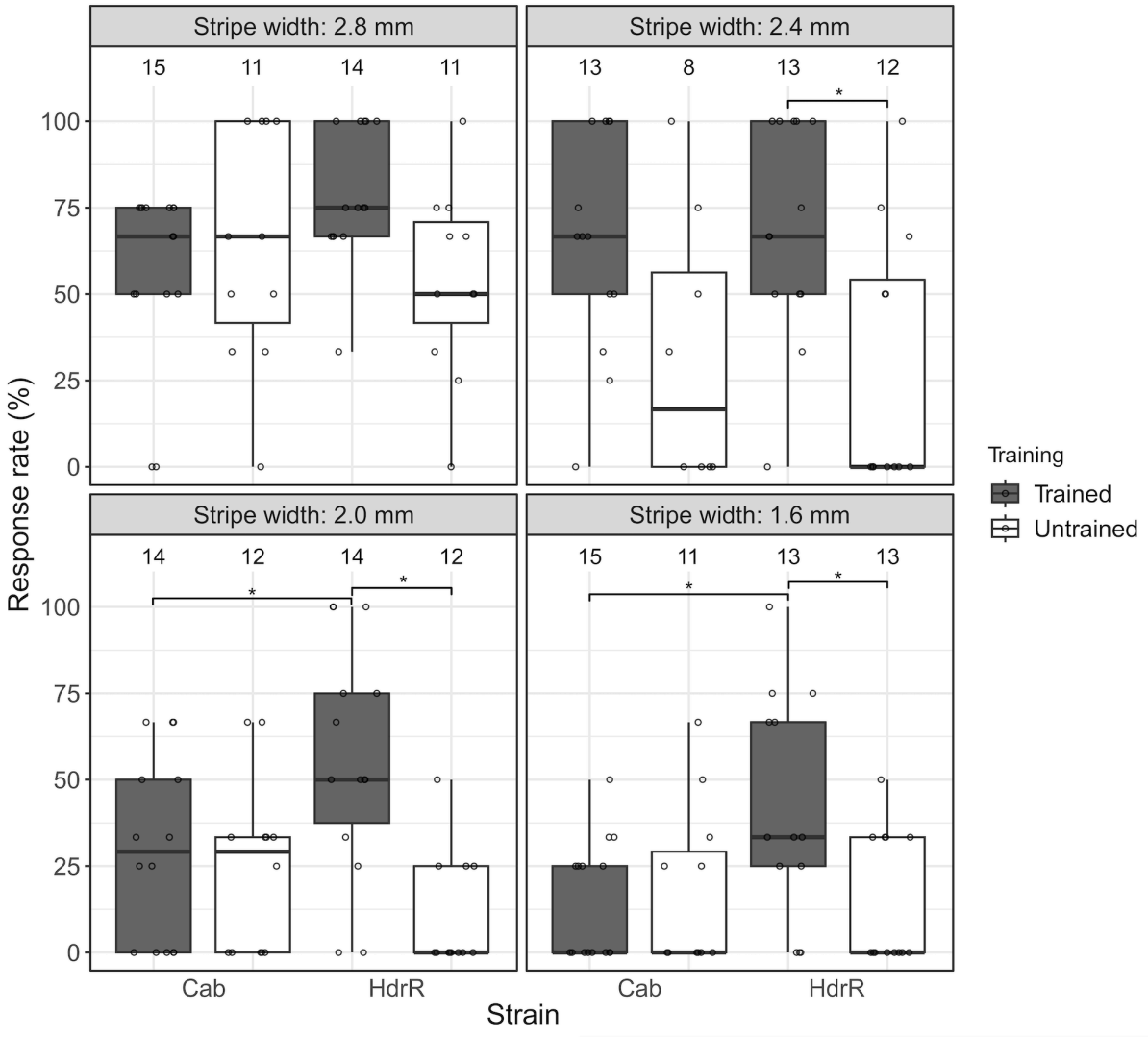


Figure 2.3 HdrR Strain Medaka Exhibit Greater Trainability in OMR Toward Narrow Stripes Compared to Cab Strain Medaka.

Response rates of Cab and HdrR strain hatchlings to moving stripe stimuli across different stripe widths (2.8, 2.4, 2.0, and 1.6 mm) under two conditions: without prior exposure to thick stripes (white) and after training with thick stripes (16 mm, grey). In both Cab and HdrR strains, response rates decreased as stripe width narrowed. Following training with thick stripes, an increase in response

rate was observed in both strains, with the HdrR strain exhibiting a more pronounced improvement. The number displayed above each boxplot indicates the total number of hatchlings included in the response rate calculation. * $p \leq 0.05$, statistical analysis performed in R, pairwise t-test, Bonferroni corrected. Adapted from (Suzuki et al., 2024).

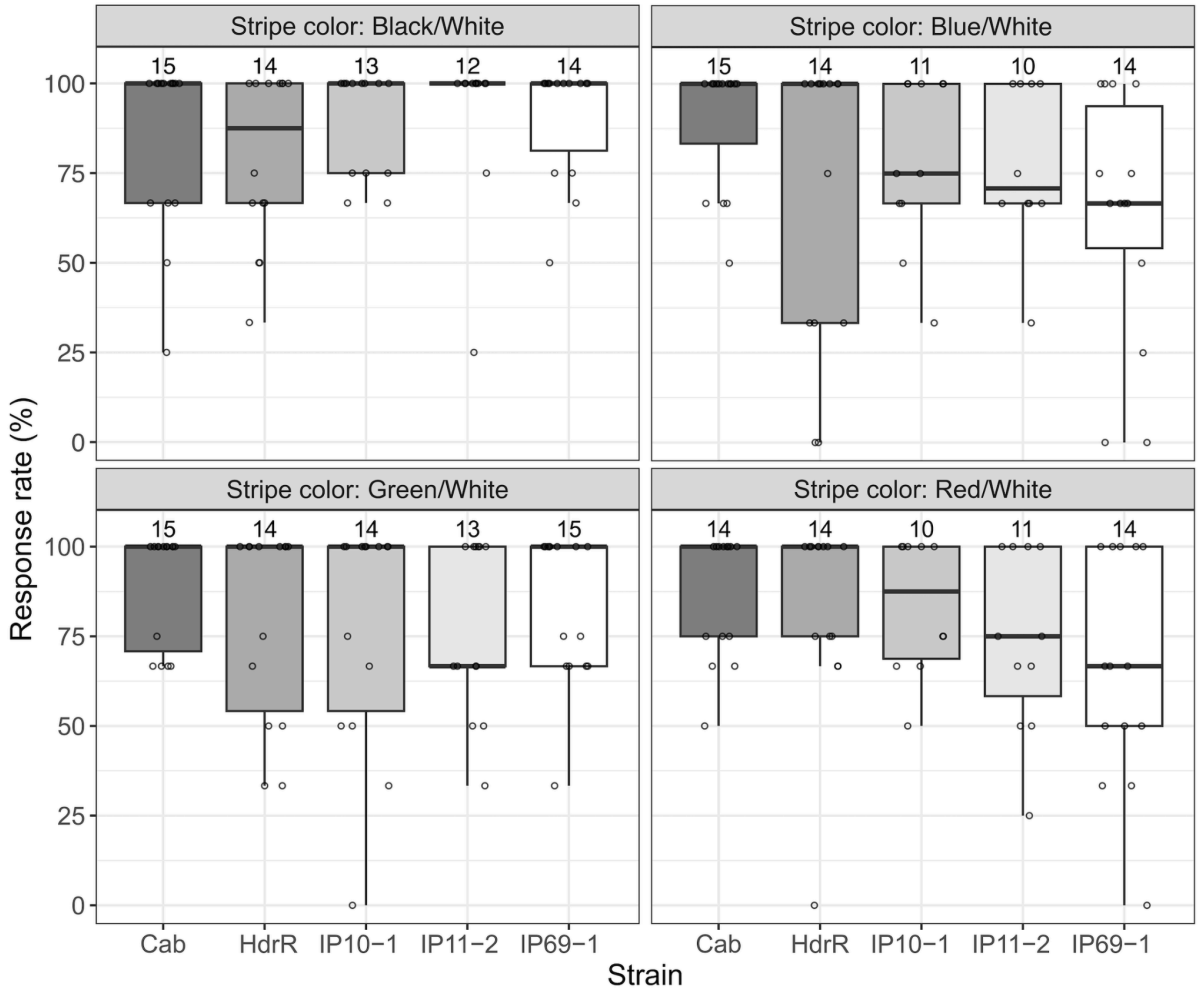


Figure 2.4 Multiple Medaka Strains Exhibit a Strong OMR to Colored Stripes.
Response rates of medaka hatchlings from five different strains (Cab, HdrR, IP10-1, IP11-2, and IP69-1) exposed to moving stripe stimuli (16 mm wide, 6.5 mm/s) with varying color combinations (black/white, blue/white, green/white, and red/white). The number displayed above each boxplot indicates the total number of hatchlings included in the response rate calculation. Statistical analysis was conducted in RStudio using pairwise t-tests with Bonferroni correction. Adapted from (Suzuki et al., 2024).

These findings indicate that OMR can be robustly induced in different medaka strains as early as 0 dph when stripe stimuli are presented from below and sensitivity to thinner stripes can be improved by training. The linear-pool-style OMR assay was shown to effectively measure spectral sensitivity. Furthermore, the results demonstrate that HdrR strain hatchlings show higher trainability in response to narrow stripes than Cab strain hatchlings, and that

MIKK strains display distinct spectral sensitivities compared to Cab and HdrR strains. These findings highlight the potential for genetic influences on visual processing in medaka.
The linear-pool-style OMR assay provides a scalable, high-throughput approach for studying visual function, trainability, and spectral sensitivity in medaka hatchlings. By precisely controlling stripe parameters and recording individual responses, this setup allows for detailed comparisons across strains.

2.1.2 Development of Infinity-Pool-Style OMR Assay With Visual Stimuli from Side

Although the linear pool arena for OMR testing was effective and throughput was increased from phenotyping single individuals to the parallel assessment of up to 15 hatchlings, it still presented several limitations. Firstly, scalability remained relatively modest, with a maximum capacity of only 15 individuals per trial. Secondly, although a semi-automated analysis pipeline was implemented, it requires manual user interference for deciding response, increasing the workload, and would benefit from further refinement toward full automation. Thirdly, the automation of image analysis was hindered by the background stripe patterns, which interfered with the detection of the hatchlings when the IR filter was absent. While the IR filter facilitated more straightforward analysis, the IR-illuminated images were of low resolution. Additionally, the reduced response rate to thinner stripes observed might potentially be due to medaka being more sensitive to visual stimuli presented from the side or above, rather than from below. Lastly, the restricted swimming area per lane within the arena further limited the range of behavioral phenotypes to mainly linear swimming behaviour that could be extracted. The brief exposure to moving stripes from below, combined with the confined space, may have led to the inclusion of startle responses in the recorded data. It is possible that some hatchlings might have reacted to abrupt changes in the light conditions caused by the stripe motion, rather than exhibiting a genuine response based on the directionality of the moving stripes.

2.1.2.1 Infinity-Pool-Style OMR Setup and Automated Detection Software

To overcome the limitation presented in the linear-pool-style OMR assay, I developed an infinity-pool-style OMR assay, utilizing monitor, prisms and petri dish. This setup effectively expands swimming space while also enabling projection of stripes from side, higher image resolution, and automated analysis.

In the infinity-pool-style OMR setup, hatchlings were placed individually in 3.5 cm diameter petri dishes, each surrounded by a prism designed to reflect stripe images from a monitor below to the side of the dish (Figure 2.5). A total of 50 dishes were placed on a monitor and subjected to stripe motions each. Stripe parameters including color, width and speed were controlled and automated using Rotating Radial Stripes Software, a custom stimulus generator developed in the lab. The software also supports scaling the size and number of the circular stripes, effectively controlling the number of dishes that can be tested simultaneously on the screen. Videos were recorded from above using two CellCam Centro 200MR cameras.

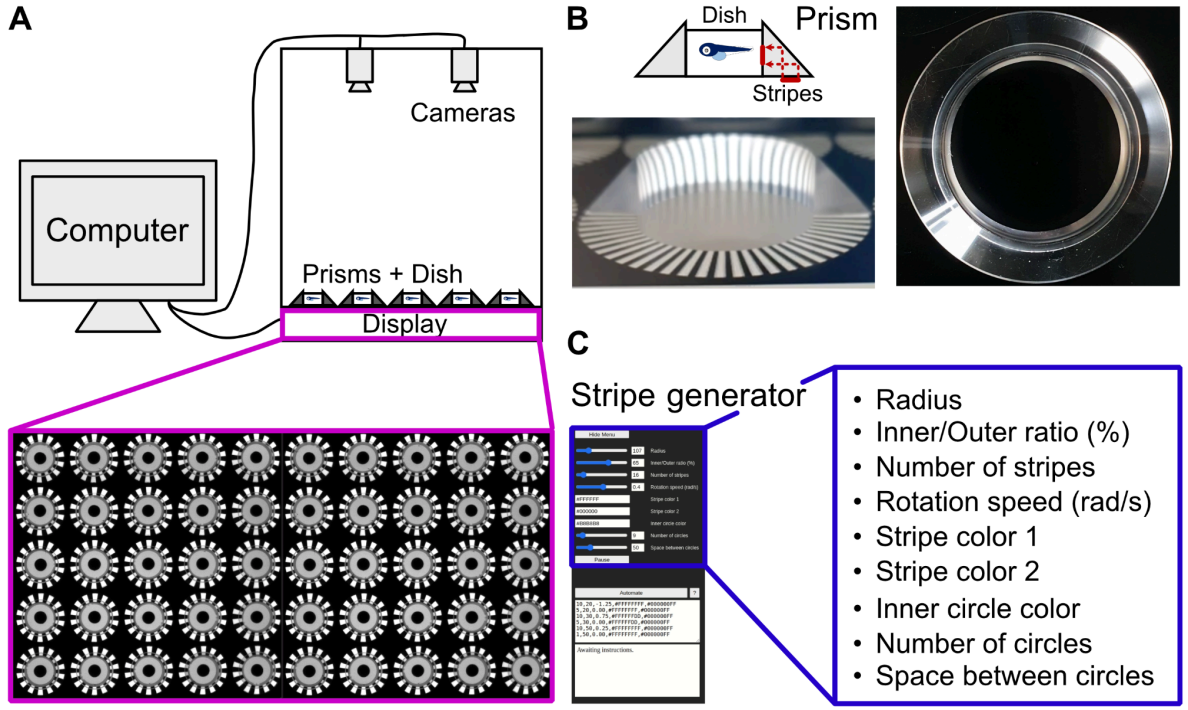


Figure 2.5 The Infinity-Pool-Style Optomotor response (OMR) Setup.
(A) Schematic representation of the high-throughput optomotor response (OMR) setup: The setup consists of an upward-facing monitor controlled by a desktop computer, with two cameras mounted above for video acquisition. 50 petri dishes (3.5 cm in diameter) and 50 prisms containing individual hatchlings are placed on top of the monitor, which displays a moving stripe pattern generated by the stripe generating software, "Rotating Radial Stripes Software". The cameras are positioned to capture hatchlings' responses from above. A black curtain was used to surround the display and the camera to eliminate external visual stimuli from the surrounding. (B) Illustration and pictures of an individual Petri dish and prism, including an example image showing the stripe pattern reflected on the side of a half-cut prism. (C) Stripe parameters such as the radius of the outer circle, inner/outer circle ratio (%), number of stripes per circle, rotation speed (rad/s), stripe colors (in hex code), inner circle color, number of circles, and spacing between circles are set and automated using the Rotating Radial Stripes Software, allowing automation of the assay.

The fish tracking software was developed in collaboration with Damjan Kalšan and Thomas Thumberger. The image analysis pipeline using the tracking software begins with the definition of regions of interest (ROIs) for each dish either manually or automatically. This is followed by background subtraction and the image processing steps, including morphological opening and dilation. Objects exceeding a size threshold are classified as hatchlings automatically (Figure 2.6 A). For each frame, the software extracts XY coordinates and angular positions of the detected objects within each dish, which allows tracking of the hatchling movements. Results can be exported as CSV files containing the coordinates or as annotated videos with overlaid tracking information (Figure 2.6 B). The detection software demonstrated over 96% detection rate (number of frames with fish-like object detected) / (total number of frames) regardless of dish positions on the monitor when tested with 0-1 dph medaka hatchlings(Figure 2.6 C, D).

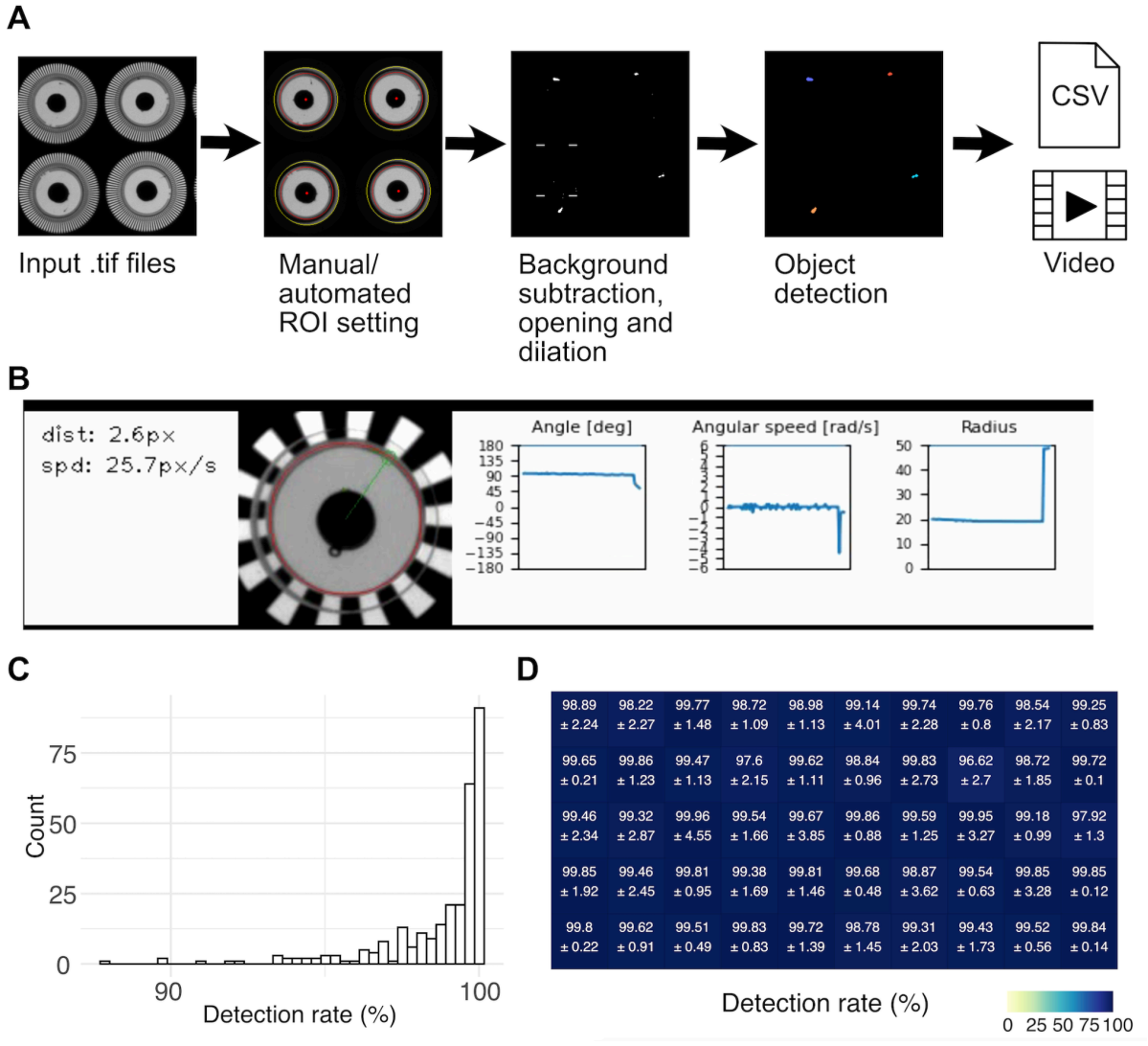


Figure 2.6 Image Processing Workflow and Detection Performance of Object Tracking Software.
(A) Pipeline of the image processing workflow used in the detection software. Regions of interest (ROIs) for each dish are defined either manually or automatically. This is followed by background subtraction and image processing, including object opening and dilation. Objects exceeding a defined threshold are identified and classified as hatchlings. For each frame, the software extracts XY coordinates and the angular position of detected objects in each dish, which can be extracted as CSV files or rendered into annotated videos. Scheme modified from a graphic originally generated by Damjan Kalšan (B) Example of the output video generated by the software, displaying each fish with a green circle marking its position and a red circle indicating the defined ROI. Angular position (degrees), angular speed (rad/s), radius, distance from the center (pixels), and swimming speed (pixels/second) are overlaid in real time. (C) Histogram illustrating the detection rate per dish calculated as the proportion of frames in which hatchlings-like objects were successfully detected, relative to the total number of frames, based on 20 experiments with 1 dph medaka. (D) Detection rate summarized across all dish positions, derived from the same set of 20 experiments with 1 dph medaka. Data are presented as mean ± standard deviation (SD).

2.1.2.2 OMR Experimental Procedure and Evaluation

To test sensitivity differences to visual stimuli, medaka hatchlings of HdrR and Cab strains were used, as well as a group of eyeless hatchlings (Cab hatchlings with eyes surgically removed at 0 dph after anesthesia) to serve as a control for blind behavior. To assess the hatchling's response to visual stimuli, a series of black and white stripes were projected, starting with thin stripes and progressing to wider stripes (ranging from 1.6 mm to 16 mm). In an initial pilot trial, 1 dph Cab strain medaka hatchlings required up to 2 minutes to begin following the stripe motion with stripe width of 16 mm (data not shown). Based on this observation, the duration for each stripe motion phase was set to 2.5 minutes to ensure sufficient time for capturing responses. Additionally, some individuals consistently swam in the direction opposite to the stripe motion. This behavior was also treated as a valid response, depending on specific criteria described in the following section.

The OMR experiment was as follows: after a 5-minute acclimation period, hatchlings were exposed to moving stripe stimuli varying in width, motion direction, and speed. Each stripe motion set included: 2.5 minutes of clockwise (CW) motion, 30 seconds pause, 2.5 minutes of counter-clockwise (CCW) motion, and 30 seconds pause. This sequence was repeated with progressively thicker stripe widths and speed adjustment (Figure 2.7 A). The stripe widths (in mm) used were: 1.2, 1.6, 2.0, 2.4, 2.8, 3.2, 3.7, 4.0, 4.4, 4.7, 5.2, 5.7, 6.0, 6.3, 6.7, 7.1, 7.6, 8.1, 8.1, 16.2, 16.2. Black and white stripes moved at a speed of 20.6°/s, except for the second round of 8.1 mm and 16.2 mm stripes, which moved at 61.8°/s (Figure 2.7 B). This stripe motion configuration served as the standard for visual acuity evaluation.

To assess hatchling's response, response values were calculated as the ratio of "angle swum by the hatchlings" to "angle moved by the stripes for every 5 seconds (Figure 2.7 C). During pause phases, the stripe velocity from the previous phase was used to calculate the response values. Since response values were calculated every 5 seconds, each 2.5-minute stripe motion phase (CW or CCW) yielded 30 data points. A hatchling was considered responsive to a specific stripe width if, in both of each CW and CCW phase, at least 15 out of the 30 response values met or exceeded the threshold. The threshold was defined as response values ≥ 0.5 (indicating swimming in the same direction as stripe motion) or ≤ -0.5 (indicating swimming in the opposite direction). In addition, the direction of swimming had to be consistent across both CW and CCW phases: if a hatchling followed the stripe motion during the CW phase, it had to do the same during the CCW phase; conversely, if it swam in the opposite direction to the stripe motion during CW, it was required to swim in the opposite direction during CCW phase as well. This directional consistency is critical, as it demonstrates that the hatchling is actively responding to changes in stripe motion direction rather than swimming randomly in one direction. The same criterion for determining responsiveness was applied across all experiments, except in the color and contrast sensitivity tests described in Section 2.1.2.4, where hatchling was defined responsive if at least 8 out of 15 response values met or exceeded the thresholds.

Response values are plotted as the heatmap in Figure 2.7 D, with each column indicating each hatchling and row as 5 second periods filled with color indicating the response values. The color of the blocks indicates, the response values of higher than 1.75 (white), 0.5 to 1.75 (yellow), 0.2 to 0.5 (green), -0.2 to 0.2 (gray), -0.5 to -0.2 (dark blue), -1.75 to -0.5 (purple) and lower than -1.75 (black). All HdrR and some Cab strain medaka hatchlings started following stripe motion at stripe width of 2.0 mm to 3.2 mm (Figure 2.7 D, response depicted in yellow in both CW and CCW phases). Once they started responding, most of HdrR strain hatchlings kept following during both CW and CCW phases in the same direction as the stripe motion (Figure 2.7 D, response depicted in yellow). A subset of Cab hatchlings demonstrated continuous swimming in "one direction" regardless of stripe motion (CW or CCW). This one directional swimming behaviour is visualized as alternating purple (or black) and yellow (or white) coloration (e.g., see the rightmost column in Figure 2.7 D). A subset of Cab strain hatchlings continuously responded in the opposite direction (e.g., see 6th column from the left in the Cab group in Figure 2.7 D, where the response depicted in purple

during both CW and CCW phases). Responses in the direction opposite to the stripe motion were also categorized as responsive, provided the hatchlings changed direction in accordance with the changes in stripe movement. In contrast to HdrR and Cab strain hatchlings, eyeless hatchlings did not show any sustained swimming behaviour or response to CW or CCW stripe motion (Figure 2.7 D Eyeless group; non-response depicted predominantly in gray).

To characterize blind behavior of the eyeless hatchlings in comparison to the wildtype hatchlings, I calculated both the total response value and the longest continuous response. Total response values were quantified by summing the absolute response values (per 5-second interval) across the entire experiment for each individual. Most HdrR and Cab hatchlings exhibited total response values exceeding 300 (Figure 2.7 E), indicating high activity throughout the experiment either by successfully following the stripe motion or by sustained swimming behavior. In contrast, eyeless hatchlings demonstrated notably lower total responses, generally around 100 (Figure 2.7 E). To evaluate response continuity, the longest sustained directional response, defined as the maximum number of consecutive 5-second intervals where the response value was either ≤ -0.5 or ≥ 0.5 indicating the response to the stripe motion in the opposite direction, and the same direction, respectively. Majority of HdrR hatchlings exhibited the maximum continuity of around 30, indicating consistent tracking of the stripe motion throughout the 2.5 minute period, with the lowest individual scoring 21. About a half of Cab hatchlings exhibited the maximum continuity of around 30 and the lowest individual scoring 16. Among eyeless hatchlings, the maximum continuity observed was 7, meaning a hatchling sustained swimming for 35 seconds in CW or CCW direction at a speed exceeding half the speed of the moving stripes. (Figure 2.7 F). From these results, a hatchling was classified as blind-like if it exhibited a continuity of response <11 and a total response value <300. (Figure 2.7 G).

Visual acuity was estimated from behavioral responses to moving stripe patterns. For individual hatchlings and for each stripe width of which it responded, stripe width in degrees (θ) was calculated by first dividing half of the stripe width (in mm) by the distance from the rim (in mm) to obtain the visual angle in radians. Then, the arctangent of that ratio was taken and multiplied by 2 to account for the full angle. Finally, the result was converted from radians to degrees by multiplying by 180 and dividing by π. In this calculation, the distance from the rim was calculated for each stripe width as the median of the top 25% most central positions (those further from the rim), selected from positional data points where the

hatchlings respond (response values ≤ -0.5 or ≥ 0.5). This approach was used to represent the furthest distance from the rim at which reliable stripe discrimination was still observed. From these points, spatial frequency (f) was defined as $1/\theta$. Visual acuity was decided to be the highest spatial frequency that each individual exhibited. If a hatchling did not respond to any of the stripe motions and was classified as "blind", its visual acuity was assigned a value of zero (Figure 2.7 G). Visual acuity was calculated for each stripe width to which the hatchlings responded, and the highest resulting value was taken as the visual acuity for each hatchling.

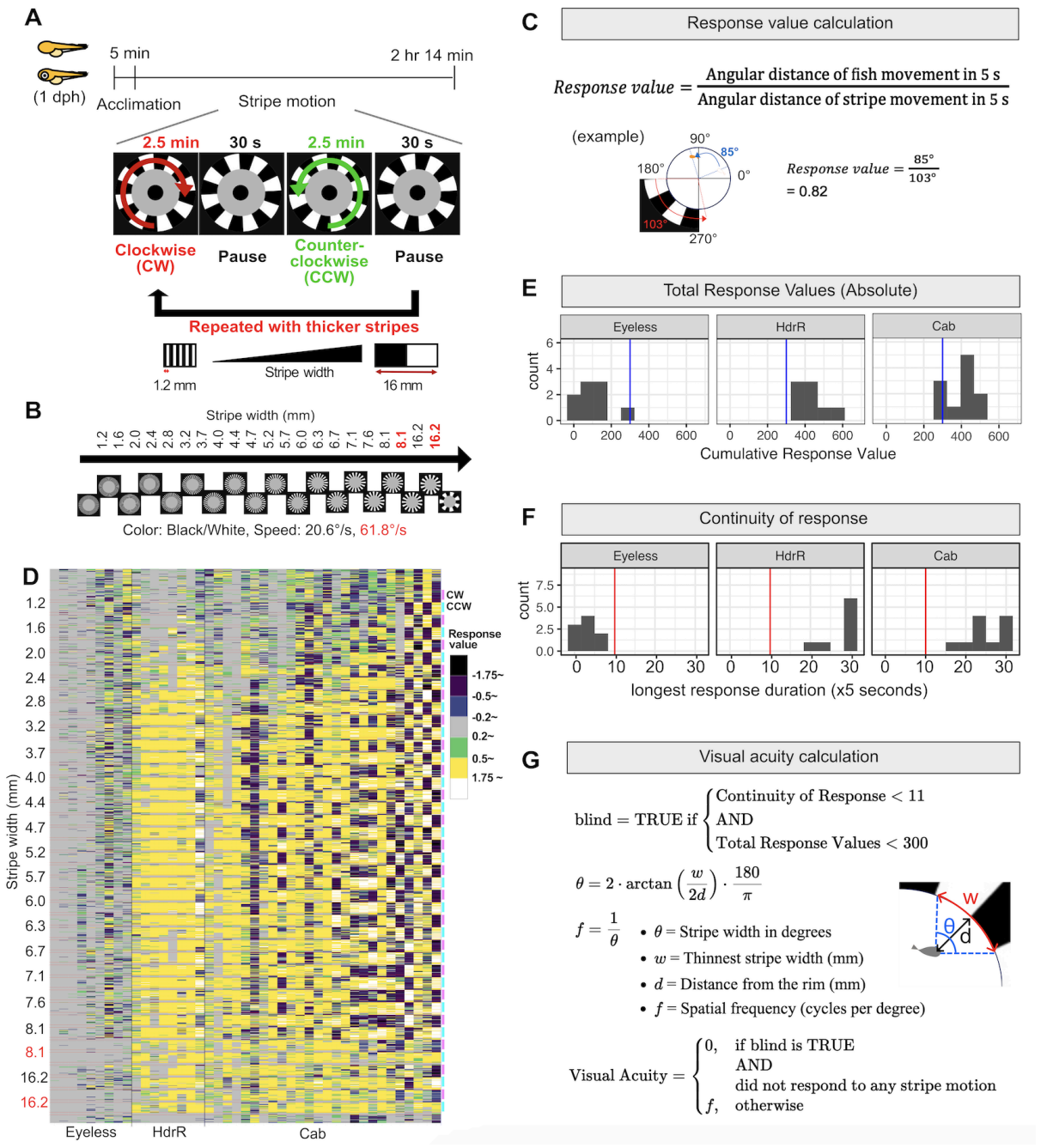


Figure 2.7 Stripe Motion Stimulus Design and Visual Acuity Evaluation.
(A) 1 dph Medaka hatchlings of HdrR (n = 8), Cab (n = 26) strains, and Eyeless (Cab strain with eyes removed surgically at 0 dph, n = 9) were used for OMR experiments. Experimental procedure for the OMR experiment. After 5 minutes of acclimation, hatchlings were subjected to the stripe motion. Each set of stripe motions consists of 2.5 minutes of clockwise (CW) motion, 30 seconds of pause, 2.5 minutes of counterclockwise (CCW) motion, and 30 seconds of pause, which was followed by further sets of stripe motions with thicker stripes and speed adjustment. (B) Stripe thickness was changed as 1.2, 1.6, 2.0, 2.4, 2.8, 3.2, 3.7, 4.0, 4.4, 4.7, 5.2, 5.7, 6.0, 6.3, 6.7, 7.1, 7.6, 8.1, 8.1, 16.2, and 16.2 mm. Black and white stripes were used. Stripe speed was set to 20.6°/s for all, except for the second round of the 8.1 mm and 16.2 mm stripe conditions, which were presented at the speed of 61.8°/s.(C) The response value was calculated as the ratio of "the angle the fish swam" to "the angle the stripes moved" every 5 seconds. Where stripe was at pause the stripe speed of previous phase was used for the response value calculation. (D) Heatmap showing response values (every 5 second as a row) of each individual hatchlings (column) at each stripe motion in each direction (CW: pink, CCW: light blue, indicated on the right side). Sections where neither CW nor CCW is indicated represent periods when stripe motion was paused. Stripe width is shown on y axis. The color of the blocks indicates, the response values: white (>1.75) and yellow (0.5 to 1.75) indicate hatchlings following the stripe motion; green (0.2 to 0.5), gray (0.2 to 0.2), and blue (-0.5 to -0.2) represent non-responsive behavior; while purple (-1.75 to -0.5) and black (<-1.75) denote hatchlings swimming in the direction opposite to the stripe motion. (E) Histogram showing total absolute response values across HdrR, Cab and Eyeless. Cumulative response values is the sum of all absolute response value measurements per individual hatchlings throughout one experiment, reflecting overall response and activity. Blue line indicates a value of 300. (F) Histogram showing the longest response duration across the entire experiment. For each hatchling and each motion phase (CW or CCW), the longest duration was defined as the maximum number of consecutive 5-second blocks in which the Response value was either ≥ 0.5, representing sustained responses in the same direction as stripe motion, or ≤ -0.5, indicating sustained responses in the opposite direction toward stripe motion, within that phase. The highest value across phases was used for each hatchling in the plot. Red line indicates a value of 10. (G) Formulas used to calculate visual acuity from behavioral responses to moving stripe patterns. Stripe width in degrees (θ) is derived from the actual stripe width (mm) and distance from the rim (mm), and spatial frequency (f) is defined as $1/\theta$. Visual acuity equals f unless the hatchling is classified as blind-like (having continuity of response < 11 and total responses < 300) and showed no response to any stripe motion, in which case visual acuity is set to zero. Visual acuity was calculated for each stripe width to which the hatchlings responded, and the highest resulting value was taken as the visual acuity for each hatchling. dph: days post hatch. min: minutes, s: second.

In the previous linear-pool-style OMR assay, Cab strain medaka exhibited reduced responsiveness following an acclimation period of less than 30 minutes, even when tested with stripe width of 16 mm. Therefore, to evaluate the influences of acclimation duration on behavioral performance in the infinity-pool-style OMR assay, I compared OMR responses of 1 dph hatchlings of HdrR and I2Cab strain (an inbred medaka line derived from the Cab strain through five generations of sibling mating) following different acclimation durations (5 minutes, 30 minutes, and 1 hour). Given the robustness of the stimulus in the infinity-pool OMR setup, I hypothesized that a 5 minutes acclimation period would yield responses

comparable to those observed with longer acclimation times, thereby improving experimental throughput relative to the previous linear-pool-style OMR setup.

As shown in Figure 2.8, hatchlings from both HdrR and I2Cab strains displayed similar OMR performance across all acclimation conditions tested. Response rates, visual acuity values, and the minimum stripe widths eliciting responses did not significantly differ between groups with different acclimation periods. These results indicate that a 5 minute acclimation period is sufficient for reliable OMR measurements in the infinity-pool-style OMR setup, supporting its use in higher-throughput behavioral assays.

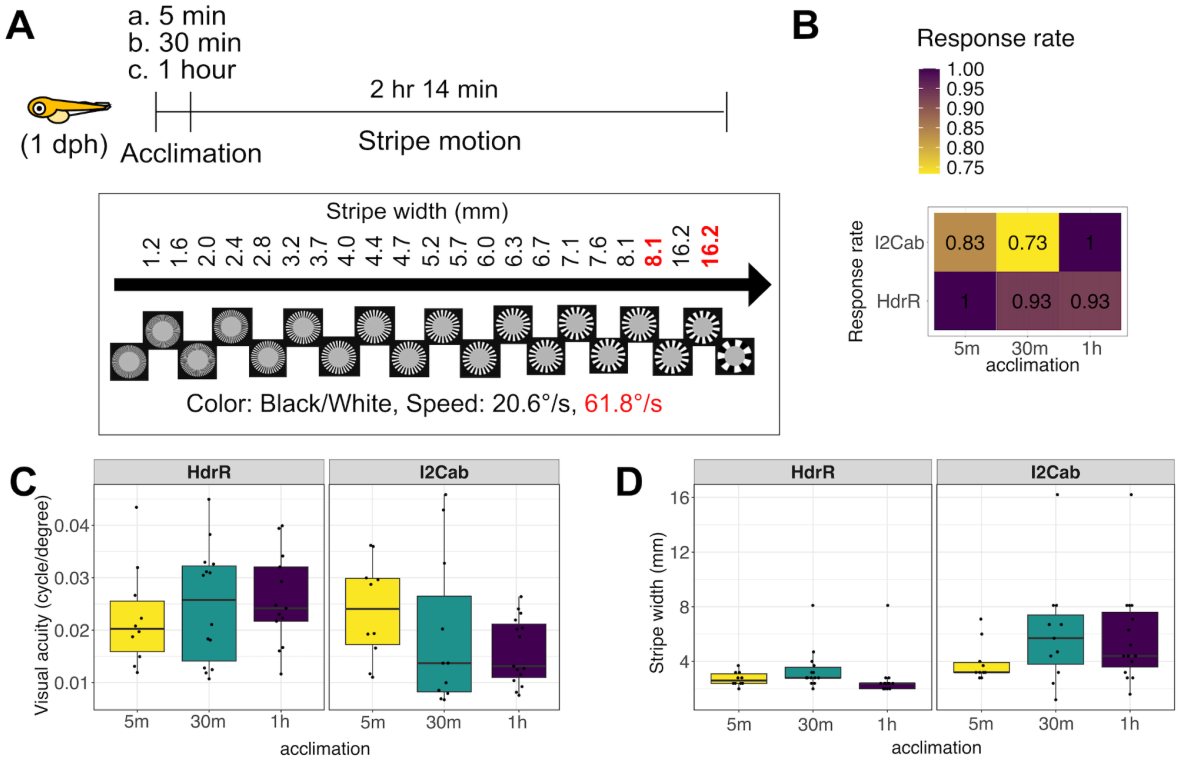


Figure 2.8 Medaka Hatchlings Exhibit Similar OMR Performance Regardless of Acclimation Duration.

(A) 1 dph Medaka hatchlings from the HdrR and I2Cab (an inbred medaka line derived from the Cab strain through five generations of sibling mating) strains were subjected to OMR assays to assess the impact of varying acclimation durations (5 minutes, 30 minutes, and 1 hour). Experiments were independently replicated three times, with more than four individuals per condition in each replicate. Stripe stimuli ranged in width from 1.2 mm to 16.2 mm and consisted of alternating black and white bands. The initial stripe motion was presented at a velocity of 20.6°/s; a second exposure to 8.1 mm and 16.2 mm stripe widths was conducted at the speed of 61.8°/s. (B) Response rate was quantified as the proportion of hatchlings considered responsive within each condition. (C) Visual acuity is presented as box plots. (D) The minimum stripe width (range 1.6 - 16 mm) eliciting a response in each hatchling was compared between HdrR and I2Cab strains. (C, D) Statistical analysis using pairwise Dunn's tests with Benjamini-Hochberg adjustment revealed no significant acclimation time dependent

differences. Each dot in each box plot represents an individual hatchling. dph: days post hatch. m: minutes, h: hour.

2.1.2.3 The Infinity-Pool-style OMR Assay Enables Cross-Strain and Cross-Species Comparison of Visual Sensitivity and Behavior

To evaluate the utility of infinity-pool-style OMR assay and to compare visual sensitivity and swimming behavior across species, five 1 dph medaka strains (HdrR, I2Cab, HO5 (all from southern population: *Oryzias latipes*), QuiH (albino mutant with mixed genetic background), Kaga strain (northern population: *Oryzias sakaizumi*)) and 5 dpf zebrafish (*Danio rerio*) larvae were subjected to the OMR assay using the standard stripe configuration: a 5 minutes acclimation period followed by 2.5 minutes of CW stripe movement, a 30-second pause, 2.5 minutes of CCW movement, and another 30 second pause. This sequence was repeated with progressively increasing stripe widths. Black and white stripes were presented in the following widths (in mm): 1.2, 1.6, 2.0, 2.4, 2.8, 3.2, 3.7, 4.0, 4.4, 4.7, 5.2, 5.7, 6.0, 6.3, 6.7, 7.1, 7.6, 8.1, 8.1, 16.2, and 16.2. Stripe speed was set to 20.6°/s for all trials, except for the second presentation of the 8.1 mm and 16.2 mm conditions, which were presented at 61.8°/s (Figure 2.9 A). Across species and strains, key behavioral metrics including “minimum stripe width” eliciting a response, “visual acuity”, “response rate” (the proportion of individuals exhibiting a response to any stripe motion relative to the total number tested in the OMR assay), “swimming speed during stripe motion and pause phases”, “one-directional swimming behavior,” and “total distance traveled”, were quantitatively assessed. These assessments revealed species and strain specific variation in visual performance (Figure 2.9 C,D,F-K), and overall responsiveness (Figure 2.9 E).

Among all groups, QuiH2 strain hatchlings exhibited the highest visual function, as indicated by the lowest minimum stripe width (median of 2.4) eliciting a response and the highest visual acuity (median of 0.0447) (Figure 2.9 C,D). Kaga, HO5 and HdrR strain hatchlings were similar in minimum stripe width (median of 2.8, 3.2, ,2.8), while I2Cab was slightly less responsive (median of 3.7), and zebrafish was the worst with median of 4.0 (Figure 2.9 C). In visual acuity, I2Cab and HdrR had similar visual acuity having median of 0.0424 and 0.0381, respectively, and Kaga and HO5 exhibit comparable values (median of 0.029 and 0.0357 respectively), while zebrafish had lowest visual acuity (median of 0.0092). These differences in the ranking of minimum stripe width and visual acuity suggest that I2Cab tend to swim closer to the center of the dish, resulting in a greater distance from the rim. As a consequence,

their visual acuity was calculated to be higher than that of other strains, despite having wider stripe widths to trigger the response. Notably, a proportion of hatchlings from the I2Cab and HO5 strains as well as Zebrafish appeared to be unresponsive having been classified as blind. Zebrafish showed limited responsiveness, with only approximately half of the individuals exhibiting measurable OMR at 5 dpf, reflecting variation in the maturation of visual circuits and physical capability within the species. In contrast, all medaka strains demonstrated robust responsiveness, with response rates exceeding 89% (Figure 2.9 E).

Analysis of swimming speed across conditions revealed that all medaka strains were able to adjust their speed closer to the stripe motion at the standard speed (20.6°/s, approximately 6.28 mm/s), having median speed of 3.5-5 mm/s, demonstrating intact visuomotor coordination (Figure 2.9 G). However, at the higher stripe velocity (61.8°/s, approximately 18.87 mm/s), while HdrR, HO5 and I2Cab increased in its speed to median of about 7.5-10 mm/s, Kaga strain hatchlings failed to adjust their speed with the stripe motion showing median speed of about 3.26 mm/s (Figure 2.9 H). In contrast, QuiH2 hatchlings maintained consistent performance at both speeds (median of 4.86 mm/s for the standard speed stripe motion and 15.7 mm/s for the faster stripe motion), suggesting higher visual tracking capabilities and robust sensorimotor integration (Figure 2.9 G, H). Zebrafish had a median speed of approximately 3 mm/s regardless of stripe speed.

During the stripe pause phases, HO5, Kaga and I2Cab strains continued to exhibit active swimming behavior, having slightly higher median speed than the others (Figure 2.9 F, I). This elevated locomotor activity in the absence of visual motion cues may reflect strain-specific differences in basal activity levels. Furthermore, HO5 and I2Cab hatchlings showed a notable tendency toward one-directional swimming, characterized by persistent movement in a single direction (either CW or CCW) throughout both motion phases (Figure 2.9 J). This behavior may indicate visual sensory deficits, motor impairments, or reduced attentional engagement with the environment. It could reflect diminished motion sensitivity or near-blindness, leading hatchlings to rely on simple light cues and exhibit wall-following or repetitive swimming patterns. Alternatively, it may result from asymmetric motor control or neurological dysfunction affecting coordinated movement. Lastly, Kaga hatchlings exhibited the greatest total swimming distance among all strains, further supporting their heightened general activity levels during the assay (Figure 2.9 K).

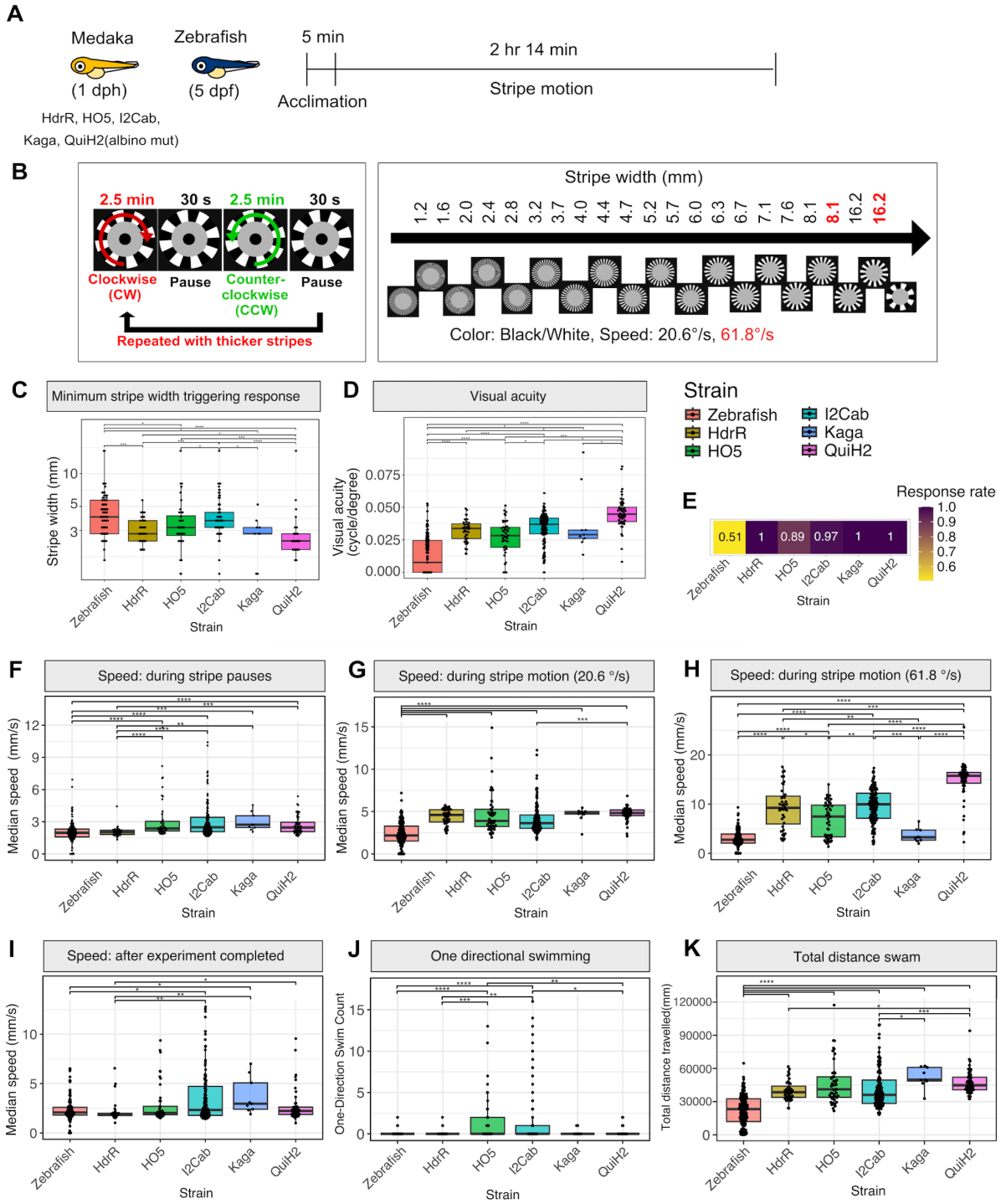


Figure 2.9 Behavioral and Visual Response Parameters Across Five Medaka Strains and Zebrafish in OMR Experiments.

(A) 1 dph five medaka strains (HdrR, I2Cab, HO5: all from southern population (*Oryzias latipes*), QuiH2: albino mutant with mixed genetic background,, Kaga strain: northern population (*Oryzias sakaizumi*) and 5 dpf zebrafish were exposed to OMR experiments. Timeline for the OMR experiment: after 5 minutes of acclimation, hatchlings were subjected to the stripe motion with different parameters. (B) Each set of stripe motions consists of 2.5 minutes of clockwise (CW) motion, 30 seconds of pause, 2.5 minutes of counter-clockwise (CCW) motion, and 30 seconds of pause, which was followed by further sets of stripe motions with thicker stripes and speed adjustment. (B) Different stripe thickness were tested: 1.2, 1.6, 2.0, 2.4, 2.8, 3.2, 3.7, 4.0, 4.4, 4.7, 5.2, 5.7, 6.0,

6.3, 6.7, 7.1, 7.6, 8.1, 8.1, 16.2, and 16.2 mm. Black and white stripe was used. Stripe speed was first 20.6°/s and for the second round of 8.1 and 16.2 mm stripes were 61.8 °/s. (C) Box plots showing the minimum stripe width to which the hatchlings of each strain responded. (D) Box plots of visual acuity for the different strains. (E) Heatmap displaying the response rate across all strains, represented as the proportion of hatchlings that responded to stripe motion relative to the total number of individuals tested in each group. (F-I) Box plots showing median swimming speed (in mm/s) during (F) stripe motion pause phases, (G) stripe movement at 20.6°/s (approximately 6.28 mm/s), (H) stripe movement at 61.8°/s (approximately 18.87 mm/s), and (I) after completion of all stripe motion phases. (J) Box plots showing the count of one-directional swimming. A response was counted as "one-directional" when a hatchling swam consistently in the same direction during both the CW and CCW stripe motion phases at a given stripe width. (e.g., a hatchling swam in the clockwise direction during both the CW and CCW stripe motion phases.) The total number of such occurrences throughout the experiment was counted per individual and summarized in box plots, reflecting the tendency for sustained directional swimming. (K) Total distance swum (in mm) during the experiment, presented as box plots. Sample size (n): Zebrafish (n = 144), HdrR (n = 40), HO5 (n = 45), I2Cab (n = 119), Kaga (n = 9), QuiH2 (n = 54). Statistical analyses were conducted in R, with Dunn's post hoc tests applied for multiple comparisons with Benjamini–Hochberg (BH) correctied. Significant differences between groups are denoted by asterisks (* $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p < 0.0001$). Each dot in each box plot represents an individual hatchling. mut: medaka mutant. dph: days post hatch. dpf: days post fertilization.

2.1.2.4 Contrast and Color Sensitivity Assessment Using Modified OMR Stimuli

To further evaluate visual function across species and strains, I assessed contrast and color sensitivity using modified stripe configuration in the OMR assay (Figure 2.10 A). Following a 5 minutes acclimation period, hatchlings were exposed to stripe motion at a fixed speed (20.6°/s) and width (8.1 mm) under various stripe contrast and color conditions. Besides seven levels of grayscale contrast and color contrast (black paired with varying intensities of blue, red, or green), 8 colored stripe combinations known to be challenging to distinguish in cases of human color blindness were tested: orange and pink, pink and gray, pink and blue, yellow and light green, light green and orange, gray and yellow, light blue and light green, and blue and purple. Each motion set consisted of 1.5 minutes of CW movement, 30 seconds of pause, followed by 1.5 minutes of CCW movement and another pause. After the grayscale contrast test and each color contrast test, there was an additional 3 minutes pause, then there was 10 minutes pause before the different color combinations were tested.

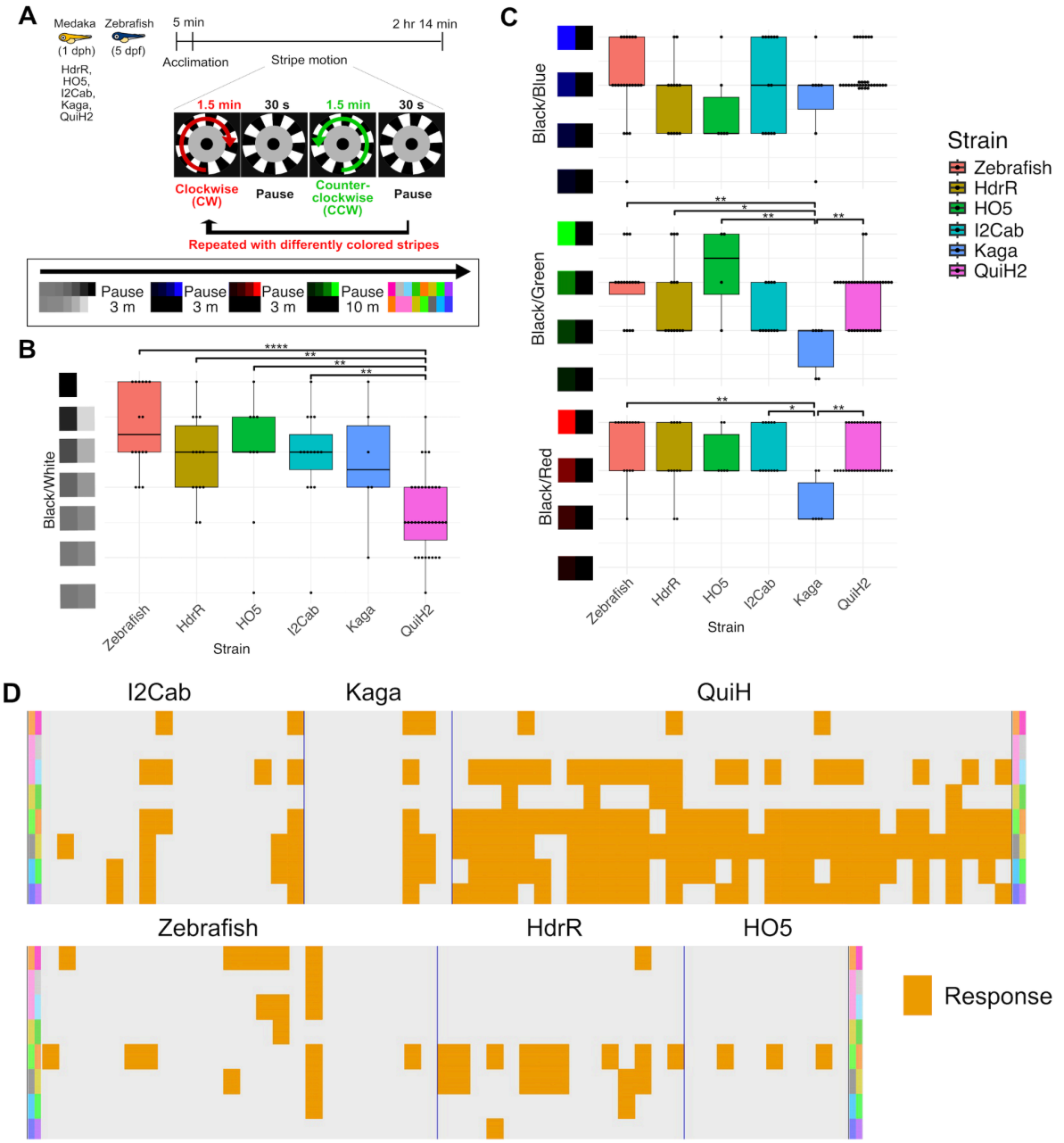


Figure 2.10 Contrast and Color Sensitivity Evaluation Across Medaka Strains and Zebrafish Using OMR.

(A) 1 dph five medaka strains and 5 dpf zebrafish were exposed to OMR experiments with different stripe contrast and color settings. Timeline for the OMR experiment for the contrast and color sensitivity: after 5 minutes of acclimation, hatchlings were subjected to the stripe motion. Each set of stripe motion consisted of 1.5 minutes of clockwise (CW) motion, 30 seconds of pause, 1.5 minutes of counter-clockwise (CCW) motion, and 30 seconds of pause, which is followed by further sets of stripe motions with higher contrasts and differential color combinations. In each experiment, stripe contrast or color was changed as gray-gray (black/white stripes with 7 different contrast levels), then pause for 3 minutest, black/blue (with 4 levels of blue), then pause for 3 minutes, black/red (with 4 levels of red), then pause for 3 minutes, black/green (with 4 levels of green), then pause for 10 minutes, colored-stripes (8 color combinations known to be difficult to distinguish for people with color weakness). The speed of stripes was 20.6°/s and the thickness was 8.1 mm. (B, C) Box plots showing

contrast threshold indicating the lowest contrasted stripe at which hatchlings responded in zebrafish and five medaka strains. Stripe response was determined as the width of the first stripe, where hatchlings responded in both CW and CCW direction. (D) Heatmap indicating individual hatchling's responses to each color combination (depicted at the row margins). The color combinations tested included, from top to bottom in the plot: orange/pink, pink/gray, pink/blue, yellow/light green, light green/orange, gray/yellow, light blue/light green, and blue/purple. Statistical analyses were conducted in R, with Dunn's post hoc tests applied for multiple comparisons with Benjamini–Hochberg (BH) corrected. Significant differences between groups are denoted by asterisks (* $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p < 0.0001$). dph: days post hatch. dpf: days post fertilization.

In the tests comparing contrast levels ranging from 1 (low) to 7 (high, black/white stripes), zebrafish, HdrR, HO5, I2Cab, and Kaga hatchlings exhibited similar responses, with a median responsiveness to stripes of contrast levels 4-6 (Figure 2.10 B). In contrast, QuiH2 strain hatchlings showed higher sensitivity to the contrast exhibiting median response at contrast level 2 (Figure 2.10 B). For the black and colored stripe combinations, no apparent differences were observed among the strains in their responses to black/blue stripes (Figure 2.10 C). Kaga strains hatchlings showed slightly better sensitivity to lower-contrast black/green, as well as black/red, stripe patterns (Figure 2.10 C). Among all tested stripes with various color combinations, QuiH2 strain demonstrated superior visual sensitivity, responding to subtle color differences except for pink and gray combination, indicating both high contrast and color discrimination capability (Figure 2.10 D). In comparison, zebrafish and other medaka strain hatchlings exhibited reduced sensitivity, showing more selective or inconsistent responses to colored stripes (Figure 2.10 D). These results reinforce the strong visual performance in QuiH2 and highlight its enhanced sensitivity to both luminance and chromatic cues.

Overall, the infinity-pool-arena OMR assay provides a robust and versatile stimuli for assessing visual function in medaka and zebrafish at hatchling stage. The setup reliably detects differences in visual acuity, contrast sensitivity, color discrimination, and behavioral phenotypes such as baseline activity and swimming speeds across strains and species. Its sensitivity allows for the identification of both subtle and pronounced visual deficits, with minimal acclimation time and scalability for even higher-throughput potential makes it ideal for large-scale screenings. Together, these results demonstrate that this OMR setup is a powerful and efficient tool for comparative visual function, genetic screening, and behavioral phenotyping.

These linear-pool-style and infinity-pool-style OMR platforms enable the simultaneous testing of up to 15 and 50 hatchlings, respectively, and incorporate semi or fully automated detection and analysis pipelines. Importantly, they demonstrate that robust OMR behavior can be reliably detected in medaka as early as 0 dph, thereby facilitating the investigation of visuomotor function at the earliest developmental stages.

2.1.3 Attempts to Establish an High-Throughput Retinal Degeneration Assay

Given the complex and multifactorial nature of retinal degeneration, screening for degenerative susceptibility across MIKK strains offers a valuable opportunity to delineate the underlying genetic mechanisms. Although genetic mutant models of retinal degeneration exist in medaka, there have been no reports to date of experimentally induced retinal degeneration models in this species (Gücüm et al., 2021; Loosli et al., 2004). Therefore I developed a protocol for inducing retinal degeneration in medaka using a metal halide lamp-based light induced retinal degeneration (LIRD) setup, adapting a well-established zebrafish LIRD model.

The experimental setup included an exposure chamber, such as a plastic dish, a beaker, or homemade chamber with optical access on three sides, enabling uniform light delivery, and was integrated with a temperature-controlled circulation system to maintain stable aquatic conditions throughout the exposure period (Figure 2.11).

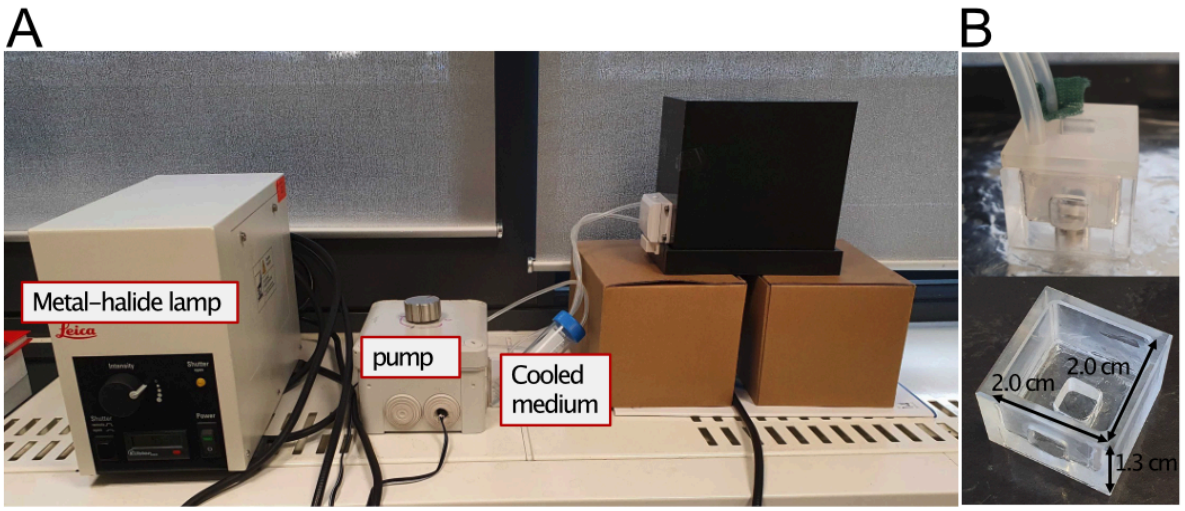


Figure 2.11 Experimental Setup for Light-Induced Retinal Injury.
(A) Photograph of the setup used to induce light injury. A metal halide lamp illuminator is positioned inside a custom-made box and placed directly over the exposure container, which may be a plastic dish, homemade chamber, or glass beaker. The tubing is connected to an external pump system that

continuously circulates cooled medium, ensuring stable temperature conditions during extended light exposure. (B) Photograph of a custom-made exposure chamber (2.0 × 2.0 × 2.0 (cm)), constructed with Polyoxymethylene. The chamber includes a square opening with glass coverslips mounted on the top, bottom, and one side, allowing light penetration and optical access during the experiment.

6 dpf zebrafish larvae and 0-1 dph Cab strain medaka hatchlings were co-exposed to intense light for 30 minutes each from the side, bottom, and top, within the same 3.5 cm polystyrene petri dishe. Following light exposure, hatchlings were fixed in 4% paraformaldehyde (PFA), cryosectioned, and subjected to immunohistochemistry. Retinal sections were stained with an anti-Zpr1 antibody (a cone photoreceptor marker), TUNEL assay for apoptotic cell detection, and DAPI for nuclear staining. The zebrafish retina exhibited a significant increase in apoptotic cells within the photoreceptor (PRC) layer as early as 24 hours post-injury (hpi), and this increase in apoptotic cells persisted through 48 hpi (Figure 2.12). Partial loss of Zpr1 staining in the retina was observed at 48 hours hpi in zebrafish, indicating complete loss of PRCs in specific retinal regions. In contrast, Cab retina showed only sparse apoptosis in the PRC layer during the same time period (Figure 2.12). The trend remained consistent even when the injury procedure was performed using a beaker. Independent experiments conducted separately on zebrafish and Cab strain medaka hatchlings consistently showed the same trend: zebrafish exhibited pronounced retinal degeneration, whereas the Cab strain medaka appeared comparatively resistant.

Because these experiments were performed using an older metal halide lamp that had been in the lab for over 6 years, I later replaced the lamp to ensure more consistent and efficient injury induction across experiments.

To determine if the resistance to light injury was strain dependent, I exposed different medaka strains (Cab, IP10-1, and IP69-1) to the LIRD setup at 0-1 dph. Following 20 minute exposure to intense light, hatchlings were fixed, cryosectioned and immunostained with an anti-Zpr1 antibody (a cone photoreceptor marker), Terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) assay for apoptotic cell detection, and DAPI for nuclear staining. As I previously observed, retina of Cab strain showed a reduced amount of TUNEL-positive cells after the injury, while the retina of IP10-1 and IP69-1 strains exhibited extensive TUNEL-positive cells in the more dorsal and central regions of the retina (Figure 2.13). Additionally, in all strains, apoptotic cells were also observed by bright DAPI staining due to chromatin condensation. These results supported the hypothesis that Cab medaka are less susceptible to the LIRD.

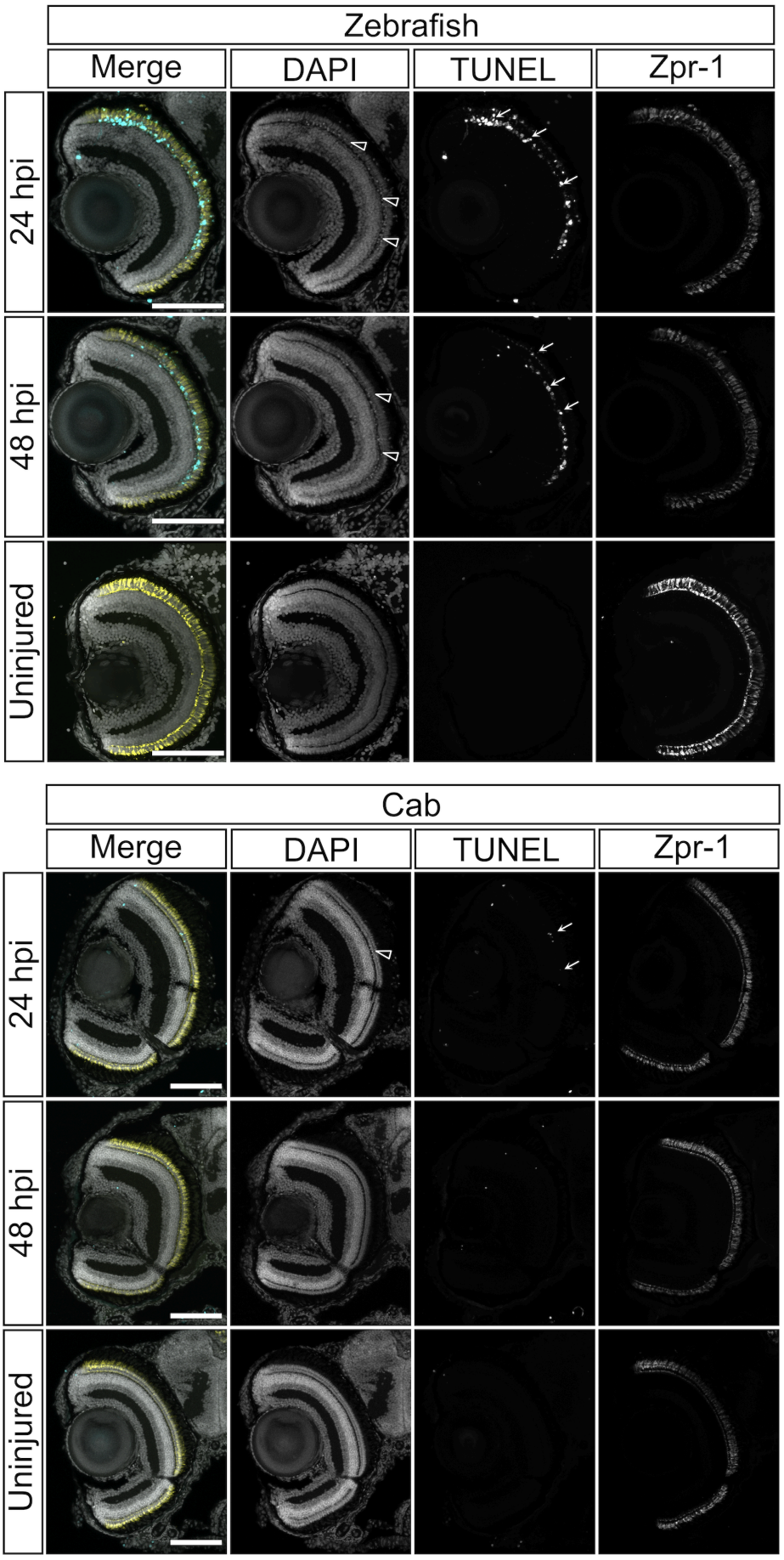

Zebrafish
Merge
DAPI
TUNEL
Zpr-1
24 hpi
48 hpi
Uninjured
Cab
Merge
DAPI
TUNEL
Zpr-1
24 hpi
48 hpi
Uninjured

Figure 2.12 Cab Medaka Exhibit Greater Resistance to Light-Induced Retinal Injury Than Zebrafish.

(A-B) Representative immunostaining of retinal cryosections from uninjured controls, 24 hours post-injury (hpi), and 48 hpi in zebrafish (A) and Cab strain medaka (B). Zebrafish larvae (6 dpf) and medaka hatchlings (0 -1 dph) were placed together in a polystyrene petri dish and exposed to metal-halide light directed sequentially from the side, bottom, and top for 30 minutes each. Following exposure, samples were fixed and processed for cryosectioning and immunostaining. Retinal sections were stained with DAPI (nuclei staining), TUNEL (staining apoptotic cells), and ZPR1 (cone photoreceptor marker) to evaluate light-induced cell death and photoreceptor integrity. Zebrafish retina exhibited a marked increase in apoptotic cells in the photoreceptor layer at 24 and 48 hpi, in contrast to the Cab strain, where apoptotic cells were infrequently observed. Scale bar: 100 µm. dph: days post hatch. dpf: days post fertilization.

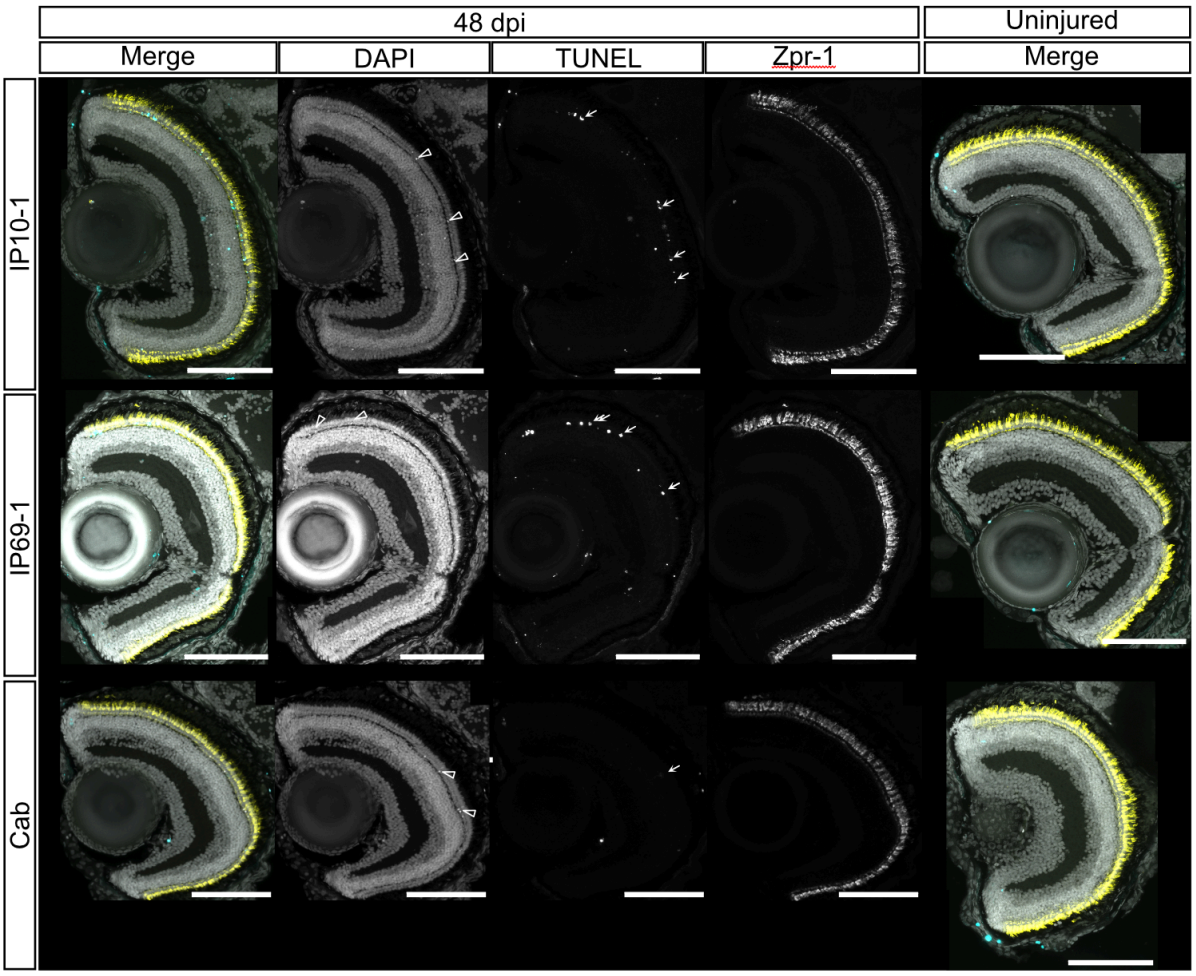


Figure 2.13 Variation in Intense Light-Induced Retinal Degeneration Among MIKK Strains.

IP 10-1, IP 69-1, and Cab strain medaka embryos were incubated at 28°C until hatching. Three days before hatching, each dish was covered with aluminum foil to shield the embryos from light. 0 dph hatchlings were exposed to intense illumination from a metal-halide lamp for 20 minutes. At 48 hours post the initial time point of exposure, hatchlings (n=2 each) were fixed in 4% PFA, cryosectioned (16 µm), and immunostained for Zpr-1 (photoreceptor cell marker), DAPI, and TUNEL (staining for apoptotic cells). Higher amount of TUNEL-positive apoptotic cells (arrows) were observed in the dorsal and central retina of both IP69-1 as well as in IP10-1 strains compared to Cab strain hatchlings.

Apoptotic cells exhibiting increased DAPI brightness due to chromatin condensation (arrowheads) were detected in Cab, IP10-1, and IP69-1 strains. Scale bar: 100 μm. dph: days post hatch.

Although the method effectively induced retinal degeneration, I encountered several limitations that affected its suitability for large-scale or high-throughput screening. I noted that survival rates after light exposure varied across strains and appeared to be different depending on the type of container used (Figure 2.14). Certain strains showed higher sensitivity when exposed in specific chambers (Figure 2.14), suggesting that both genetic and experimental conditions influenced outcomes. The severity of injury varied both within and across strains due to multiple technical and biological factors. These included fluctuations in light intensity as the lamp aged, difficulties maintaining stable water temperature, uncontrolled positioning of hatchlings during exposure, and body pigmentation differences among strains that likely influenced light penetration to the body and absorption, further contributing to uneven injury severity.

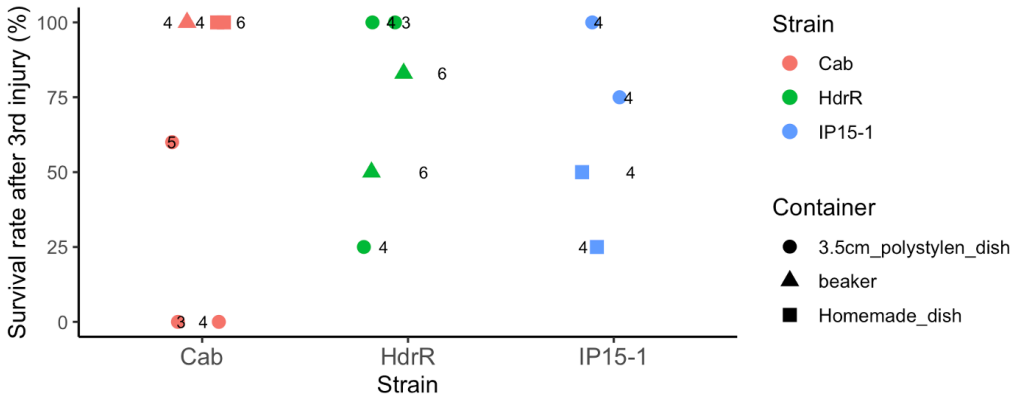


Figure 2.14 Survival Rate Differences Among Medaka Strains Depending on Container.
Dot plots show the survival rates of medaka hatchlings after three consecutive (10 to 25 minutes) exposures to the LIRD setup. Each dot represents a biological replicate (each group of 3 to 6 hatchlings). Shapes indicate the type of exposure container used: circles for plastic dishes, squares for homemade chambers, and triangles for glass beakers. Colors denote different medaka strains (Cab, HdrR, and IP15-1). Variability in survival was observed across both strains and container types, highlighting the influence of experimental setup on post-exposure viability.

To my knowledge, aside from one previous study employing UV light in adult medaka, this represents the first adaptation of a high-intensity visible light induced retinal injury model in Medaka. Although I was able to induce retinal degeneration successfully at hatchling stage medaka, I consistently observed that Cab strain medaka hatchlings exhibited greater resistance to light-induced retinal damage compared to zebrafish. These above-mentioned

sources of variability made it challenging to establish a uniform injury condition that ensured survival across all strains while still inducing retinal damage in some. As a result, the approach was not well suited for high-throughput screening within the MIKK panel.

Further investigation into the differential resistance to light-induced retinal damage among strains could provide valuable insight into the molecular mechanisms underlying susceptibility, as well as inform our understanding of the evolutionary basis for these differences. The observed variability in susceptibility to retinal degeneration across strains presents an opportunity for future investigation, which will be further discussed in the Discussion section.

2.2 Assessment of Visual Function and Behavioral Phenotypes Across the MIKK Panel Using the Infinity-Pool-Style OMR

With the successful establishment of a high-throughput OMR assay capable of distinguishing diverse visual and visuomotor phenotypes across strains, large-scale screening of genetically and phenotypically diverse populations became feasible. In this section, I apply the OMR assay to further phenotype the MIKK panel, comprising over 80 inbred medaka lines with naturally occurring genetic variation, which serves as a powerful resource for dissecting the genetic architecture underlying variation in visual function and behavior.

2.2.1 Developmental Stage and Visual Experience Influence Strain-Specific OMR Performance

Before screening entire MIKK panel, I conducted a pilot experiment using three medaka strains (HdrR, IP58-2, and IP72-2) in order to characterize the visual function across developmental stages.

Hatchlings were repeatedly subjected to OMR with black and white stripe varying stripe width at 1, 4, 7, and 10 dph. The stripe configuration was the same as previous (Figure 2.9 A), in short, after 5 minutes acclimation period, hatchlings were exposed to alternating 2.5 minute clockwise (CW) and counter-clockwise (CCW) stripe motions, separated by 30 second pauses, across a fixed sequence of increasing stripe widths from 1.6 mm to 16 mm. The two standard speeds were tested: 20.6°/s (Group a) and 41.2°/s (Group b) (Figure 2.15

A). An additional untrained group (Group c) was tested only at 10 dph using 41.2°/s to examine the training effect by repeated visual exposure (Figure 2.15 A).

Individual behaviour (response values) differs during the experiment and across developmental stages (Figure 2.15 B-C). Under the standard stripe speed (20.6°/s, group a), HdrR strains hatchlings exhibited slightly greater sensitivity to thinner stripe widths than IP58-2 and IP72-2 strain hatchlings at 1 dph (Figure 2.16 B). However, its performance did not improve with repeated testing at later stages (4, 7, and 10 dph), on the contrary, HdrR strain hatchlings showed less sensitivity toward the thinner stripes and a drop in response rate following repeated exposure, from 74% at 1 dph to an average of 50% between 4 and 10 dph (Figure 2.16 I), suggesting decreased motivation or engagement with repeated exposure. IP72-2 strain hatchlings showed modest gains in stripe width detection but little change in overall visual acuity (Figure 2.16 A, B). Interestingly, one-directional swimming behavior (depicted as alternating yellow and purple blocks in Figure 2.15 B-C) emerged in HdrR and IP72-2 strain hatchlings from the second trial at 4 dph under lower-speed conditions (20.6°/s), supporting the possibility of a potential habituation or loss of engagement with repeated exposure. In contrast, IP58-2 demonstrated a steady improvement in both visual acuity and sensitivity to finer stripes across development (Figure 2.16 A, B). Not only that the majority of IP58-2 strain hatchlings constantly followed stripe motion throughout the experiment (Figure 2.15 B-C, depicted as continuous yellow blocks). Both IP58-2 and IP72-2 maintained high response rates throughout development, averaging 95% and 83%, respectively (Figure 2.16 I), although IP72-2 tended to swim predominantly in one direction and responded only to the thickest stripes.

At the stripe speed twice as fast (41.2°/s, group b), HdrR again performed best among the three strains at 1 dph, with increased sensitivity through 4 to 7 dph, followed by a decline at 10 dph (Figure 2.16 C, D). One-directional swimming at 4 dph was also reduced in HdrR strain hatchlings, indicating that faster motion may help maintain attention or engagement (Figure 2.15 B-C). The response rate in HdrR remained high until 7 dph with an average of 87%, then dropped to ~50% by 10 dph (Figure 2.16 I). In contrast, both IP58-2 and some IP72-2 strain hatchlings showed continuous improvements in both visual acuity and stripe width sensitivity across developmental stages. IP58-2 showed a response rate of average 61% until 4 dph, which then increased to 100% at 7 dph onward, while IP72-2 maintained a steady response rate of average 75% throughout development (Figure 2.16 I). Nearly a half of

IP72-2 strain hatchlings displayed one-directional behavior at the faster speed from 4 dph onward (Figure 2.15 B-C).

Across all strains, higher stripe speeds required wider stripe widths to elicit robust responses. At the higher stripe speed, IP72-2 and HdrR maintained higher response rates compared to the lower speed, suggesting that faster stripe motion more effectively sustained attention in their developing visual systems.

Compared trained (Group b) and untrained (Group c) hatchlings at 10 dph, IP58-2 and IP72-2 strain hatchlings, trained individuals exhibited greater or similar sensitivity to finer stripes and significantly higher visual acuity than their untrained counterparts, indicating a positive effect of repeated exposure (Figure 2.16 E, F). An increase in visual acuity alone may reflect improved body orientation in response to the stimulus, as acuity increased despite similar minimum stripe widths between trained and untrained individuals, possibly because trained fish swam farther from the rim toward the center of the arena, resulting in higher calculated visual acuity. In contrast, trained HdrR hatchlings performed worse than untrained ones constantly following stripe motion in HdrR (with response rate of untrained being 100% while trained being 50%), with lower acuity and poorer sensitivity to stripe width, again suggesting that repeated exposure reduced responsiveness, possibly due to habituation or lowered motivation. IP72-2 strain hatchlings showed a response rate of 100% in the untrained group and 70% in the trained group, whereas IP58-2 exhibited similar response rates regardless of prior exposure (untrained: 90%, trained: 100%) (Figure 2.16 E, F). In untrained hatchlings tested only at 10 dph (Group c), while both IP58-2 and IP72-2 displayed weaker responses or swam in the opposite direction (as shown by prominent purple blocks in Figure 2.15 C), suggesting an experience-dependent component to successful stripe tracking.

When comparing untrained hatchlings that are subjected OMR at 1 dph and 10 dph, HdrR and IP58-2 strain hatchlings showed comparable visual acuity between the two stages but both strains hatchlings showed slightly better sensitivity to the thinner stripes at 10 dph than 1 dph indicating the visual development (Figure 2.16 G, H). IP72-2 showed significantly improved stripe width sensitivity at 10 dph (Figure 2.16 G, H). Response rates increased with age across all strains: from 80% to 100% in IP72-2, from 64% to 90% in IP58-2, and from 85% to 100% in HdrR (Figure 2.16 I).

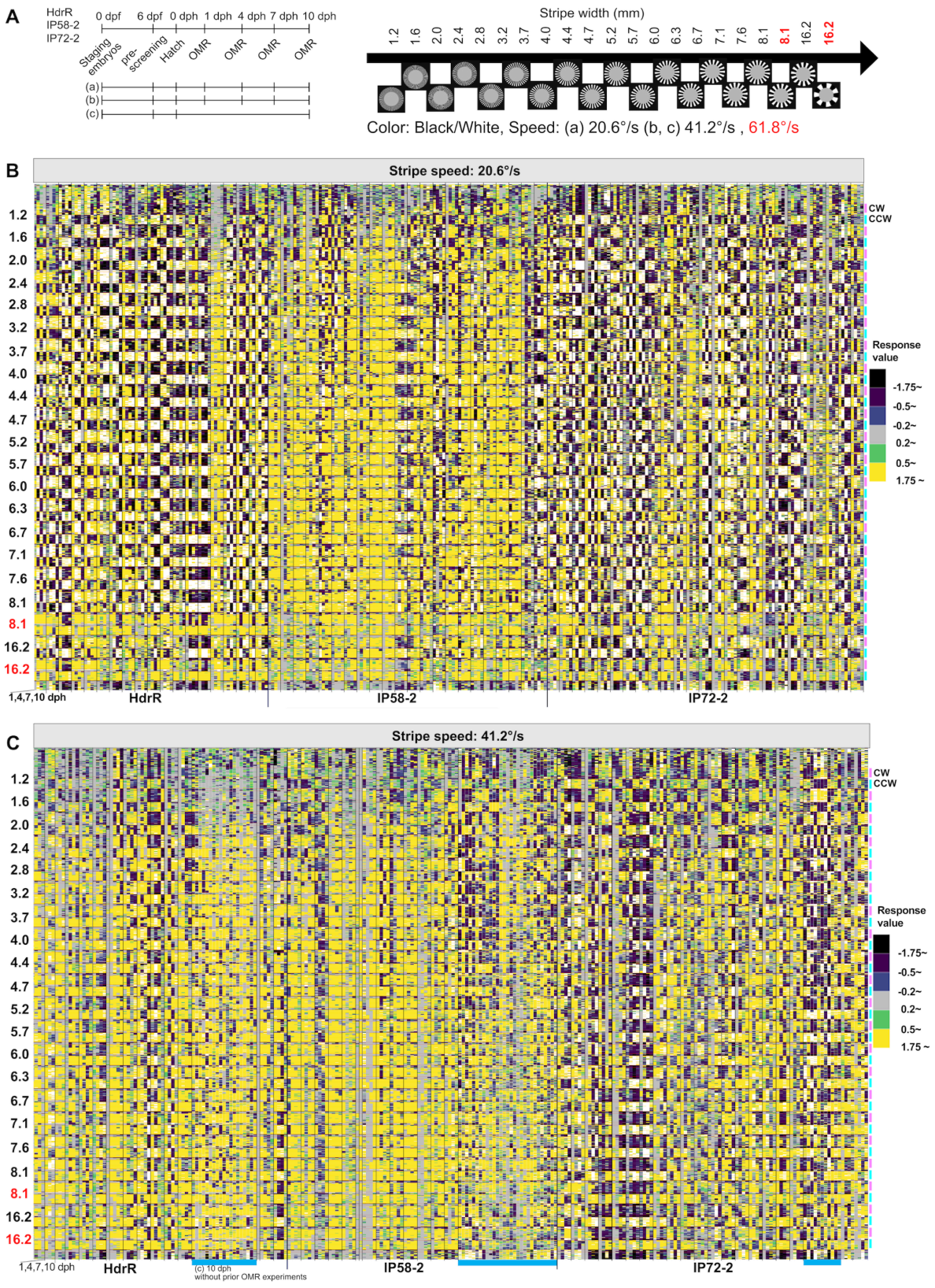


Figure 2.15 OMR Assay in Medaka Hatchlings Across Developmental Stages.

(A) Timeline of a pilot experiment conducted using the traditional inbred strain HdrR and two MIKK panel strains, IP58-2 and IP72-2. Hatchlings were repeatedly subjected to OMR assays at 1, 4, 7, and 10 days post-hatching (dph) using a stripe motion speed of 20.6°/s (group a) or 41.2°/s (group b). An additional group was tested only at 10 dph using the 41.2°/s speed (group c) Two independent

experiments were done for each group. After a 5 minutes acclimation period, hatchlings were exposed to alternating stripe motions consisting of 2.5 minutes of clockwise (CW) motion, a 30-second pause, 2.5 minutes of counter-clockwise (CCW) motion, and another 30-second pause. This sequence was repeated with progressively thicker stripe widths. Stripe thicknesses were presented in the following order: 1.2, 1.6, 2.0, 2.4, 2.8, 3.2, 3.7, 4.0, 4.4, 4.7, 5.2, 5.7, 6.0, 6.3, 6.7, 7.1, 7.6, 8.1, 8.1, 16.2, and 16.2 mm. The stripe motion speed was 20.6°/s for group (a), and 41.2°/s for groups (b) and (c), except for the second round of 8.1 mm and 16.2 mm stripes, which were presented at an increased speed of 61.8°/s. (B-C) Heatmaps showing individual response values (rows) for each hatchling at each developmental stage (columns). Columns are separated by thick black lines to distinguish strains, and by thin black lines to indicate individual hatchlings. Within those separated by thin stripes, columns are arranged from left to right in the order of developmental stages: 1, 4, 7, and 10 dph. Light blue line at the column mergin indicates for those individual hatchlings (each column) tested only at 10 dph (group c). Each stripe motion phase is indicated by color-coded bars on the right of the heatmap: CW in pink and CCW in blue; unmarked regions correspond to pause phases. Stripe width is represented on the y-axis. Heatmap color scale represents response values: white (>1.75) and yellow (0.5 to 1.75) indicate hatchlings following the stripe motion; green (0.2 to 0.5), gray (0.2 to 0.2), and blue (-0.5 to -0.2) represent non-responsive behavior; while purple (-1.75 to -0.5) and black (<-1.75) denote hatchlings swimming in the direction opposite to the stripe motion. HdrR and IP72-2 exhibited reduced responses to repeated OMR exposure and a tendency toward one-directional swimming (alternating yellow and purple blocks) over time. In contrast, IP58-2 showed increased visual sensitivity with age, as indicated by stronger responses to thinner stripe widths.

These results demonstrate that first-time exposure often elicits higher response rates regardless of developmental stage, and that even at 1 dph, there are clear strain-specific differences in visual performance. As the visual system matures, the stimulus must be sufficiently strong (e.g, faster stripe motion) to maintain responsiveness after repeated testing. However, higher speeds reduce pattern resolution, requiring wider stripes to elicit responses.

Taken together, 1 dph with stripe speed of 20.6°/s appeared optimal for visual screening for MIKK panel, as it yields higher response rate, reduces variability introduced by post-hatching factors such as feeding status, developmental speed, rearing environment, oxygen levels, and lighting conditions. This early time point offers a reliable window for comparing visual function across strains with minimal confounding factors.

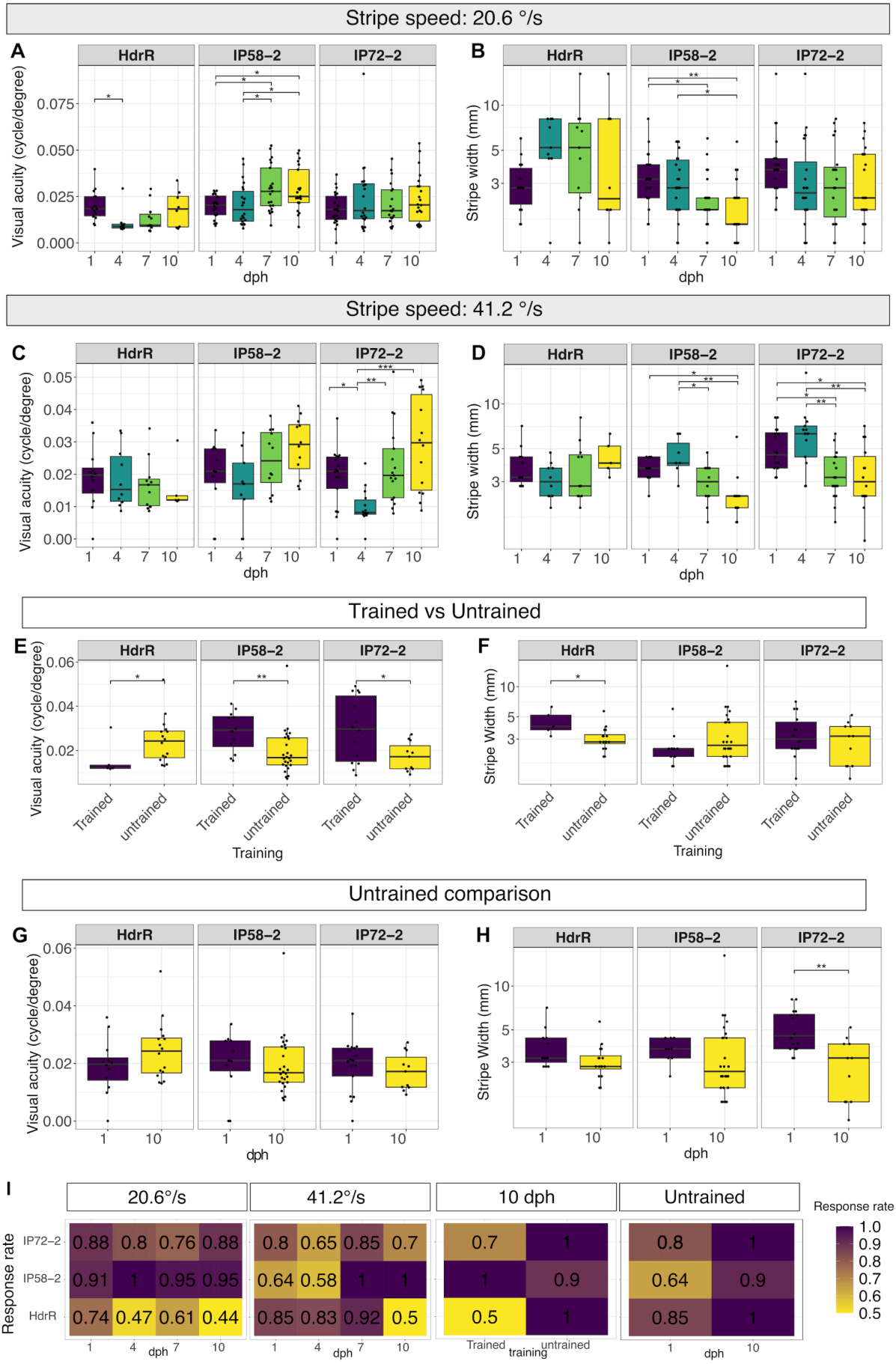

Stripe speed: 20.6 °/s
A
B
HdrR
IP58–2
IP72–2
Visual acuity (cycle/degree)
Stripe width (mm)
dph
Stripe speed: 41.2 °/s
C
D
Trained vs Untrained
E
F
Trained
untrained
Training
Untrained comparison
G
H
I
20.6°/s
41.2°/s
10 dph
Untrained
Response rate
IP72–2
IP58–2
HdrR
0.88 0.8 0.76 0.88
0.91 1 0.95 0.95
0.74 0.47 0.61 0.44
0.8 0.65 0.85 0.7
0.64 0.58 1 1
0.85 0.83 0.92 0.5
0.7 1
1 0.9
0.5 1
0.8 1
0.64 0.9
0.85 1

Figure 2.16 Developmental and Experience-Dependent Changes in Visual Acuity and Response Sensitivity in Medaka Hatchlings.

Visual performance outcomes in medaka strains (HdrR, IP58-2, and IP72-2) across developmental stages and visual experience conditions, assessed using the optomotor response (OMR) assay. Hatchlings were exposed to moving black-and-white stripes that varied in stripe width as in (Fig 2.7 A,B). Visual acuity and the minimum stripe width that elicited a response were assessed at 1, 4, 7, and 10 days post-hatching (dph) using two different stripe speeds: 20.6°/s (Group a) and 41.2°/s (Group b), except for the second round of 8.1 mm and 16.2 mm stripes, which were presented at an increased speed of 61.8°/s. An additional group of untrained hatchlings (Group c) was tested only at 10 dph to assess the effect of repeated visual exposure. (A-D) Visual acuity (A, C) and the minimum stripe width that elicited a response (B, D) across developmental stages (1, 4, 7, and 10 dph) in medaka strains tested with a stripe speed of 20.6°/s (Group a) and 41.2°/s (Group b). (E-F) Comparison of visual acuity (E) and minimum stripe width responded (F) between 10 dph trained (Group b, repeatedly tested) and untrained (Group c, tested only at 10 dph) hatchlings, highlighting the effect of accumulated visual experience. Training-related improvements were observed in IP72-2 and IP58-2 strain hatchlings, where trained hatchlings exhibited higher visual acuity, while trained HdrR strain hatchlings reduced visual response compared untrained group. (G-H) Comparison of visual acuity (G) and stripe width sensitivity (H) between untrained hatchlings at 1 dph (from Group b) and 10 dph (from Group c), reflecting visual system maturation in the absence of prior exposure. IP72-2 strain hatchlings showed significant developmental increase in visual sensitivity between 1 and 10 dph, while similar trends were less evident in the other strains. (I) Response rate across all groups and time points, shown as the proportion of hatchlings that responded to stripe motion at each age. Strain-dependent differences in visual sensitivity are already detected at 1 dph. Developmental comparisons revealed clear improvements in visual performance over time in IP58-2 and IP72-2, including enhanced acuity and sensitivity to finer stripe widths. In contrast, HdrR exhibited reduced or similar responses across stages.

Statistical analyses were conducted in R, with Dunn's post hoc tests applied for multiple comparisons with Benjamini–Hochberg (BH) corrected. Significant differences between groups are denoted by asterisks (* $p<0.05$, ** $p<0.01$, *** $p<0.001$). dph: days post hatch.

2.2.2 Visual Function and Visual Stimuli Driven Behaviour Across MIKK Panel

To delineate the loci that contributes to the diversity in visual function, an F2 segregation analysis was conducted in the following steps: first characterization of 74 MIKK strains for visual function and visual stimuli driven behaviours, then seven F0 strains with different visual and behavioural phenotypes were crossed to produce heterozygous F1 offspring, which were then intercrossed to generate a genetically and phenotypically diverse F2 population. Phenotypic evaluation of the F2 individuals, combined with whole genome sequencing, enabled the identification of trait-associated loci through genotype-phenotype correlations and association testing (Figure 2.17).

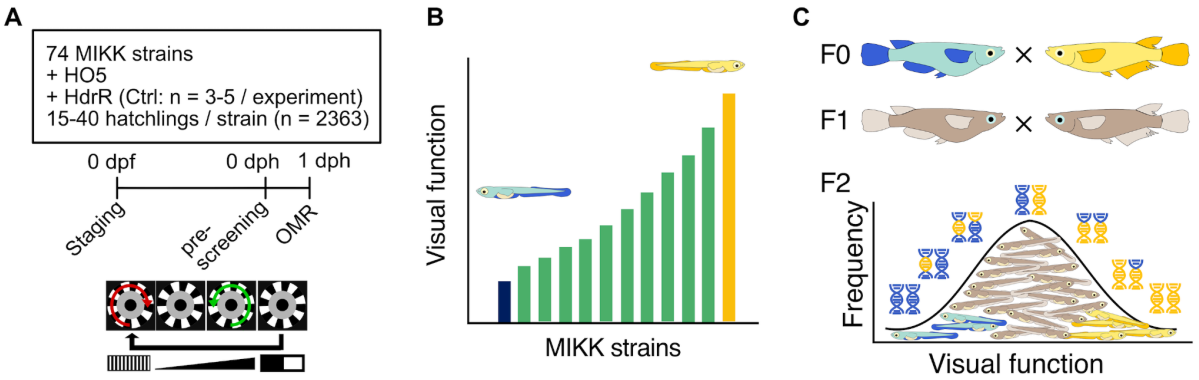


Figure 2.17 F2 Segregation Analysis of Visual Function and Related Behavioral Phenotypes.
(A) 74 MIKK strains and 2 traditional inbred strains were subjected to OMR assay with varying stripe width to evaluate their visual function. (B) The minimum stripe width eliciting an optomotor response, adult eye morphology, and response rate were key criteria in selecting inbred strains from the MIKK panel for strategic crosses and subsequent segregation analysis (C) 'F0' denotes the parental generation used for crossing; 'F1' refers to the first-generation offspring, which are theoretically uniform and heterozygous; and 'F2' represents the second-generation offspring, which exhibit genetic and phenotypic diversity due to recombination. Phenotypic variation within the F2 mapping population is expected to span the range observed in the two parental strains.

To characterize variation in visual performance among the MIKK panel, 74 inbred strains were assessed using distinct metrics (minimum stripe width responded and visual acuity). Visual acuity, quantified by the spatial frequency (in cycles per degree, Figure 2.18 A), as well as minimum stripe width triggering the response (Figure 6.5 A), indicators of visual sensitivity to spatial detail, varied across MIKK panel, with median values of visual acuity ranging from 0.0042 (IP138-1) to 0.035 (IP84-2) (Figure 2.18 A). To further investigate phenotypic diversity, I analyzed additional behavioral and developmental traits: swimming speed measured under four conditions (during stripe motion at 20.6°/s (~6.28 mm/s, Figure 6.6 B), 61.8°/s (~18.87 mm/s, Figure 2.18 B), during pause phases within the experiment (Figure 6.6 D), and 10 minutes post-experiment (Figure 6.6 D), total distance swam (mm, Figure 2.18 C), one-directional swimming (count of one directional swimming behaviour, Figure 6.5 B), and developmental age at experiment (1 dph) (in dpf, Figure 6.6 C). Swimming speed in response to stripe motion at 61.8°/s exhibited considerable strain-specific variability, with median values ranging from 2.5 mm/s (IP130-2) to 13.5 mm/s (IP8-2), indicative of differential visuomotor integration and underlying locomotor capacity (Figure 2.18 B). Total distance swam also exhibited substantial variation, with median values ranging from 24 (IP130-2) to 80.1 (IP50-2) meters, reflecting differences in overall locomotor activity and physical capacity (Figure 2.18 C).

These phenotyping results revealed substantial variation in not only visual function but also locomotor activity and the capacity to modulate movement in response to visual stimuli. Compiled from multiple technical replicates per strain, the data underscore strain-specific behavioral differences that likely reflect underlying genetic variation. Moreover, the findings highlight the robustness of this assay not only for dissecting phenotypes related to visual function, but also for capturing broader traits associated with physical fitness.

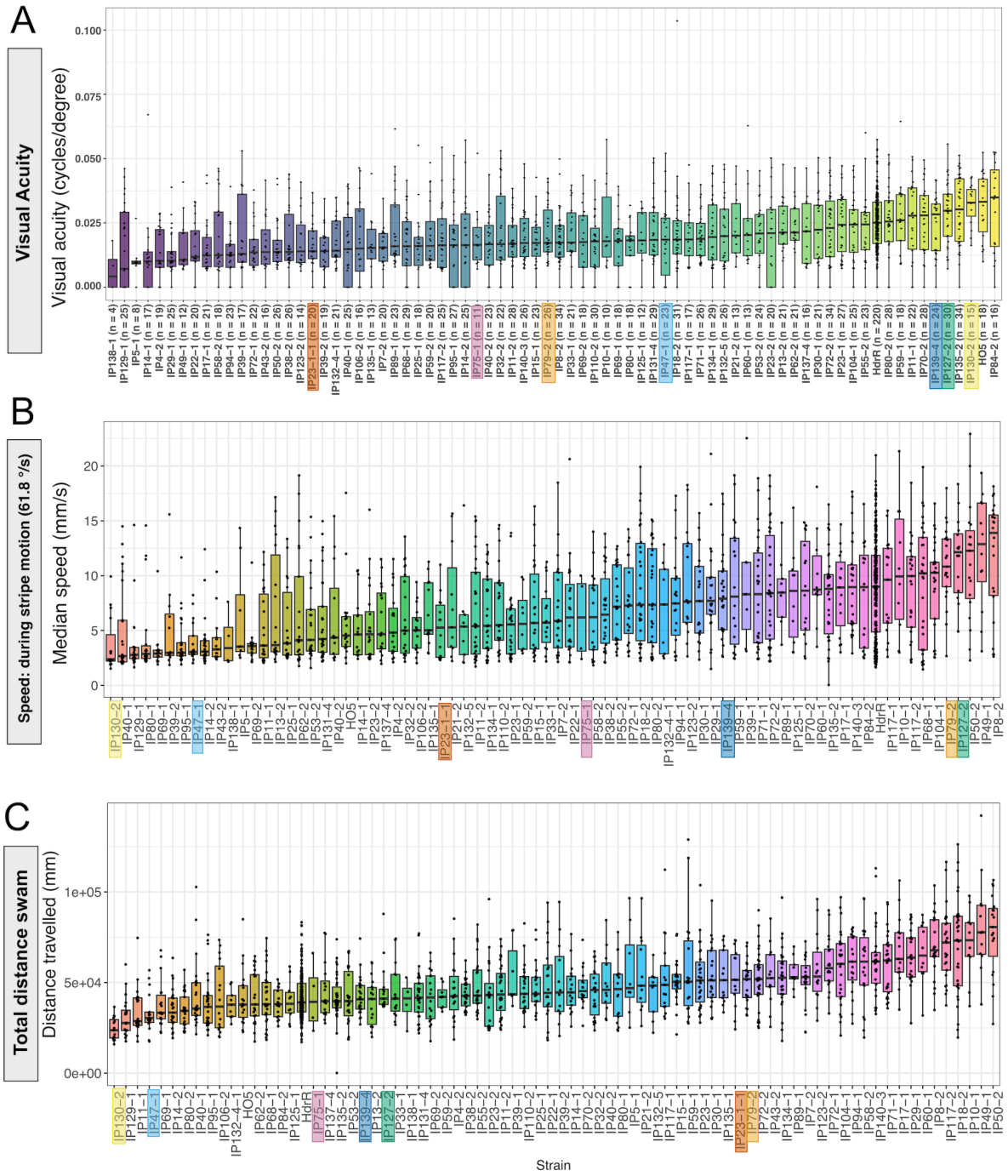


Figure 2.18 Visual and Behavioral Traits Across 74 MIKK Panel Strains.
Box plots representing behavioral metrics across 74 MIKK panel strains: (A) Visual acuity, calculated based on the spatial frequency (cycles per degree). (B) Swimming speed during stripe motion at a

speed of 61.8°/s (~18.87 mm/s). (C) Total distance swam in mm. Strains are ranked by the median value for each trait. Colors of the box plot represent the median values for each trait. Data were aggregated from multiple technical replicates per strain. Strains selected for crossing are highlighted with color.

To evaluate if a day-dependent difference is observed for the two visual performance metrics, measurements for the reference strain HdrR used in the OMR assay were compared across different experimental days. While the minimum stripe width that elicited a response remained relatively consistent (Figure 2.19 A), visual acuity exhibited some variability (Figure 2.19 B) across various experimental days. While the minimum stripe width remained relatively consistent, visual acuity exhibited some variability, suggesting that responsiveness to stripe thickness is stable, whereas the measured distance from the arena rim may be influenced by additional factors, such as the hatchling's initial position within the arena. To further assess whether the developmental time required until hatching influenced visual performance, visual acuity was examined by days post-fertilization (dpf). It should be noted that all hatchlings included in the experiments were phenotyped at 1 day post hatch (dph), providing a controlled developmental time point for OMR comparison. No consistent trend or developmental effect was observed (Figure 2.19 C), suggesting that visual function remains comparable at 1 dph despite differences in time required till hatch.

The high-throughput OMR setup enabled efficient and reliable phenotyping of visual function alongside a range of behavioral and developmental traits across the MIKK panel. All 74 inbred strains exhibited substantial phenotypic diversity, spanning from differences in developmental speed (days till hatch) to variation in visual sensitivity and behavioral responsiveness. This broad spectrum of variation provides a valuable foundation for downstream genetic mapping and functional analysis.

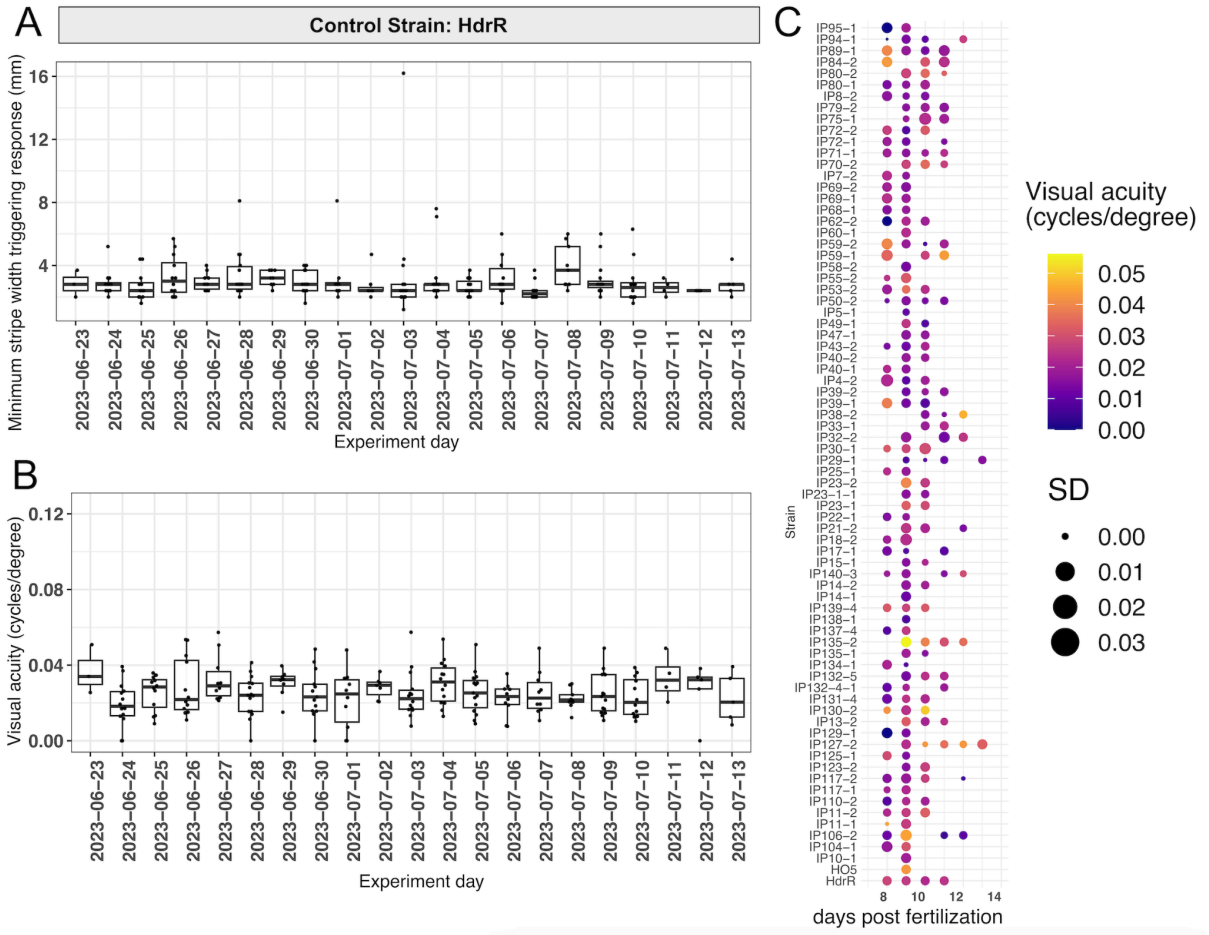


Figure 2.19 Visual Acuity Remains Stable Regardless of Experimental Day or Required Time to Hatching Across the MIKK Panel Strains.
(A, B) Boxplots showing the minimum stripe width that triggered a response (A) and the corresponding visual acuity measurements from HdrR strain hatchlings (B) across different experimental days. (C) Dot plots illustrate visual acuity for each strain at various days post-fertilization (dpf); all individuals were measured at 1 day post-hatch (dph). Visual acuity remained consistent across different days post-fertilization.

2.3 Strategic Crosses and Functional Assessment of the Retina in Selected Strains

Out of the 74 phenotyped MIKK inbred strains, seven strains were selected for strategic crossing based on their phenotypic diversity in visual function and morphological characteristics at adult stages. To guide the strain selection for crossing, two primary traits were considered: minimum stripe width that elicited a response and overall response rate (Figure 2.20 A). In addition to functional performance, fecundity was also taken into account. The selected strains (IP 23-1-1, IP 47-1, IP75-1, IP79-2, IP127-2, IP130-2, and IP139-4) span phenotypic spectrum, from high to low visual sensitivity and also vary in the other behavioral phenotypes. Notable adult eye morphology observed in the F2 generation of MIKK strains

included small eyes (unilateral or bilateral) in strain IP23-1-1 (~5% assumptions from observation), large eye size in IP139-4, lens opacities and inflated cornea resembling cataract in IP130-2 (~30%), and small eyes (unilateral or bilateral) with dark body pigmentation and low activity levels in IP47-1 (~10%).

The crossing scheme consisted of 12 unique pairwise combinations among the seven selected strains, with IP139-4 (high fecundity strain, with intermediate performance in visual function) included as a common partner in all crosses. F1 offspring was collected and reared from single mating pairs for each combination, with one cross (IP130-2 x IP47-1) resulting in embryonic lethality by 24 hpi.

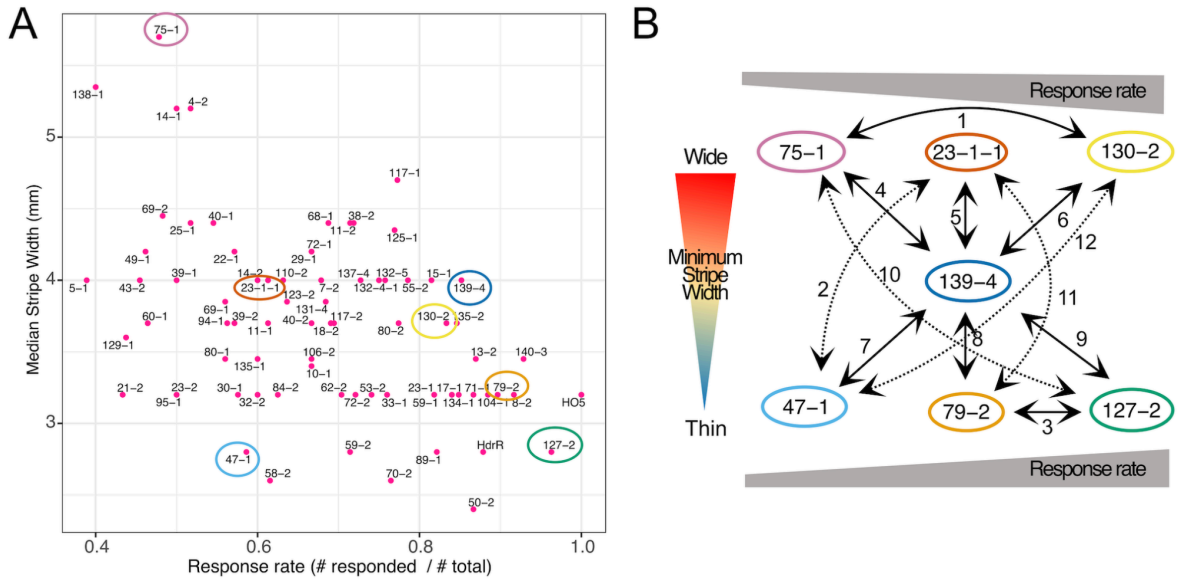


Figure 2.20 Strategic Cross Design Based on Phenotypic Variation in MIKK Strains.
(A) Visual function phenotypes across MIKK strains, shown by comparing median of minimum stripe width triggering a response and their response rates (percentage of overall individuals that responded). Strains selected for crossing (highlighted) span the full phenotypic spectrum, from high to low visual function and response rate. (B) Crossing scheme used for an F2 segregation analysis. Seven phenotypically distinct MIKK strains were crossed in 12 unique pairwise combinations (indicated by arrows, numbered from 1 to 12). All strains were paired with IP139-4 (phenotypic average), and two other strains with greater phenotypic contrast. F1 offspring were generated from single mating pairs. All crosses were successful except for cross 130-2x47-1, whose F1 embryo failed to develop further.

After initial phenotypic differences, behavioural traits were characterised in more detail for the seven selected MIKK panel strain, individual response patterns were visualized for each hatchling across different stripe motion phases in the OMR assay (Figure 2.21). The color-coded response values ranged from swimming in the same direction (white/yellow) to the opposite (black/purple) and revealed distinct, strain-specific profiles. For instance, IP75-1 strain exhibited delayed responses, primarily to thicker stripes (as indicated by yellow blocks

appearing only at stripe widths of approximately 3.7 mm and above in Figure 2.21), along with pronounced one-directional swimming behavior evident from alternating yellow and purple blocks across CW and CCW motion phases (Figure 2.21). In contrast, IP127-2 strain showed robust and consistent responses across a broad range of stripe widths, including thinner stripes (Figure 2.21). Some individuals from IP130-2 strain exhibited occasional response to thinner stripes, though inconsistently. Additionally a subset of individuals displayed a blind-like phenotype characterized by a lack of continuous response and a predominance of gray blocks (as also seen in Figure 2.7 D Eyeless), which may result from structural abnormalities such as lens opacities, partial vision loss or low locomotor capacity. A proportion of IP47-1 strain displayed little to no response, only following CW or CCW stripe motion or displaying blind-like phenotype (Figure 2.21). IP23-1-1 strain showed mixed response patterns, with signs of one-directional swimming interspersed with periods of blind-like behaviour (Figure 2.21). Notably, its performance declined at thicker stripe widths, possibly due to a restricted visual field or impaired motion detection at lower spatial frequencies (Figure 2.21). Meanwhile, IP139-4 and IP79-2 strains exhibited overall moderate to low responsiveness; while some individuals followed mid-range stripe widths, a few demonstrated one-directional swimming behavior.

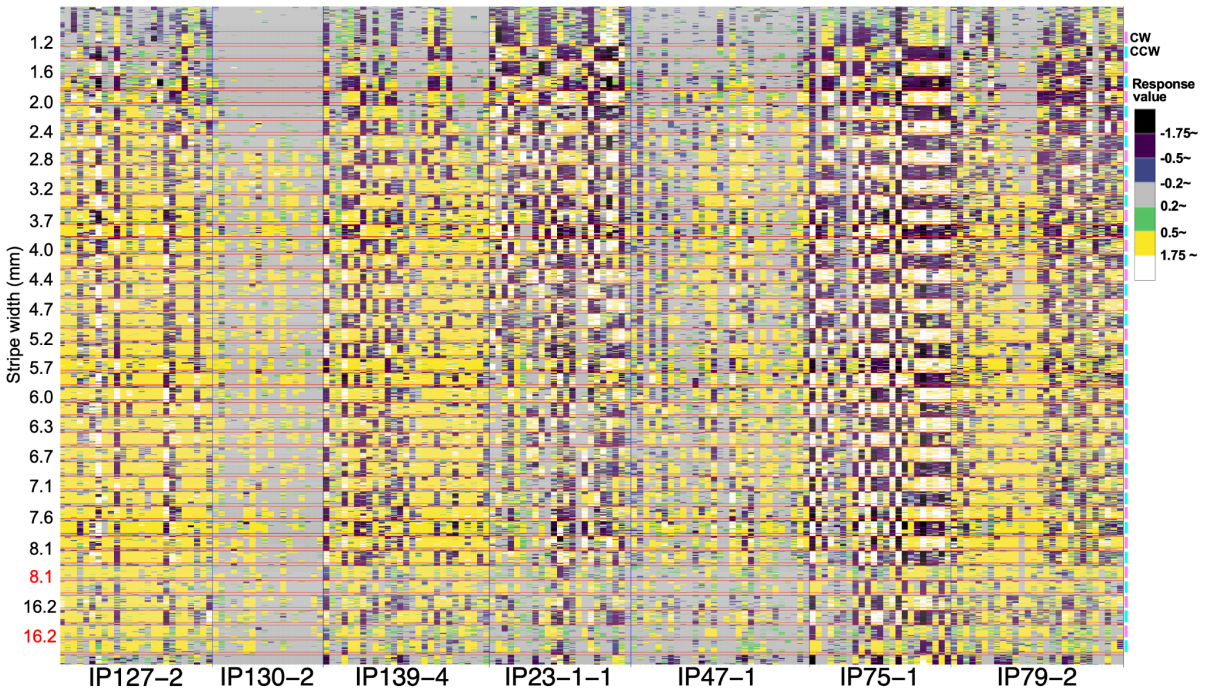


Figure 2.21 Heatmap of Response Values Highlights Behavioral Variation Among Selected MIKK Strains.

Heatmap showing individual response values at different stripe widths (rows) for each hatchling (columns) for the MIKK panel strains selected for the segregation cross setup. Each stripe motion phase is indicated by color-coded bars on the right of the heatmap: CW in pink and CCW in light

blue; unmarked regions correspond to pause phases. Stripe width is represented on the y-axis. Heatmap color scale represents response values: white (>1.75) and yellow (0.5 to 1.75) indicate hatchlings following the stripe motion; green (0.2 to 0.5), gray (0.2 to 0.2), and blue (-0.5 to -0.2) represent non-responsive behavior; while purple (-1.75 to -0.5) and black (<-1.75) denote hatchlings swimming in the direction opposite to the stripe motion. Strain-specific trends include: IP75-1 exhibited delayed responses, primarily to thicker stripes, as indicated by yellow blocks appearing only at stripe widths of approximately 3.7 mm and above, along with pronounced one-directional swimming (seen as alternating yellow and purple across CW and CCW phases). IP127-2 strain showed consistent responses, even to thinner stripes. IP130-2 responded to thinner stripes only inconsistently, with a few individuals with blind-like patterns (indicated by predominant gray blocks), which might be influenced by lens abnormalities. IP47-1 displayed weak or no responses (indicated by predominant gray blocks), suggestive of a blind-like phenotype. IP23-1-1 strain showed a mixture of one-directional swimming, moderate responses, and instances of no response (gray). Notably, its performance declined at thicker stripe widths, possibly due to a restricted visual field or impaired motion detection at lower spatial frequencies. Strains IP139-4 and IP79-2 both showed generally moderate or reduced responsiveness.

To further investigate visual sensitivity, the optokinetic response (OKR) assay was performed on 1 dph hatchlings from the selected MIKK strains and two traditional inbred strains (HdrR and I2Cab) (Figure 2.22 A). In OKR assay, an animal is immobilized while a rotating drum with moving stripes surrounds it to elicit eye movements, thereby allowing assessment of retinal function independent of locomotor capability. Hatchlings were embedded in methylcellulose and presented with black-and-white stripe stimuli varying in contrast, spatial frequency (width), and temporal frequency (speed). Configurations of stimuli were as follows, contrast sensitivity: fixed spatial frequency of 7.5 cycles/360° and angular velocity of 20°/s, with contrast ranging from 0.05 to 1.0, and spatial frequency: constant contrast of 0.7 and angular velocity of 7.5°/s, with spatial frequency ranging from 7 to 56 cycles/360°, and temporal frequency: constant contrast of 1 and spatial frequency of 20 cycles/360°, with angular velocity varying from 5 to 30°/s. Each stripe motion was presented for 3 seconds in one direction, followed by 3 seconds in the opposite direction, and again 3 seconds in the original direction. Hatchlings were exposed to all stimulus types in the same order: contrast, spatial frequency, and then temporal frequency. Eye movements were recorded under controlled light and temperature conditions, and average slow-phase eye velocities, representing the smooth rotating movement of the eye as it follows the direction of stripe motion prior to the rapid resetting saccade, were calculated using tracking software (Mueller et al., 2011). Multiple individuals per strain were tested to evaluate both inter-individual and strain-specific variation.

As shown in Figure 2.22 B-D, box plots depict average eye velocity in response to each stimuli-condition. HdrR showed consistent, high-performing optokinetic responses across all parameters and individuals, increasing the eye speed following stripe motion till 30 cycle/360° and adjusting eye velocity to stripe speed. Most hatchlings from IP127-2, IP139-2, IP79-2 strains performed well in all conditions. However, the majority of IP130-2 strains exhibited consistently lower average eye speeds compared to the other strains, with about a half of individuals failing to respond effectively to stimuli in both spatial and temporal frequency assays (faster and thinner stripe patterns). This may indicate physical difficulty in modulating eye speed or problems in visual system, potentially due to restricted visual fields, impaired retinal signaling, partial vision loss, or lens opacity. In IP23-1-1, IP47-1, IP75-1 and I2Cab, a subset of individuals exhibited reduced response in every condition tested, though most individuals responded within the normal range, suggesting impaired visual processing or sensorimotor integration in the affected individuals. These findings align with the OMR data, where the subsets of these strains exhibiting blind-like behavior (IP130-2, IP23-1-1 and IP47-1 in Figure 2.21), as well as IP75-1 strain which required markedly thicker stripe widths (Figure 2.20 A) to elicit a response in the OMR assay, also showed reduced responses in the OKR assay, collectively supporting the presence of visual functional deficits in these specific strains. Additionally, wholemount immunostaining of the IP47-1 retina using the cone photoreceptor marker zpr1 revealed an altered pattern of photoreceptor distribution compared to other strains (e.g. IP15-1, Figure 6.4), which may be one of several factors contributing to the impaired visual responses observed in OKR and OMR assays.

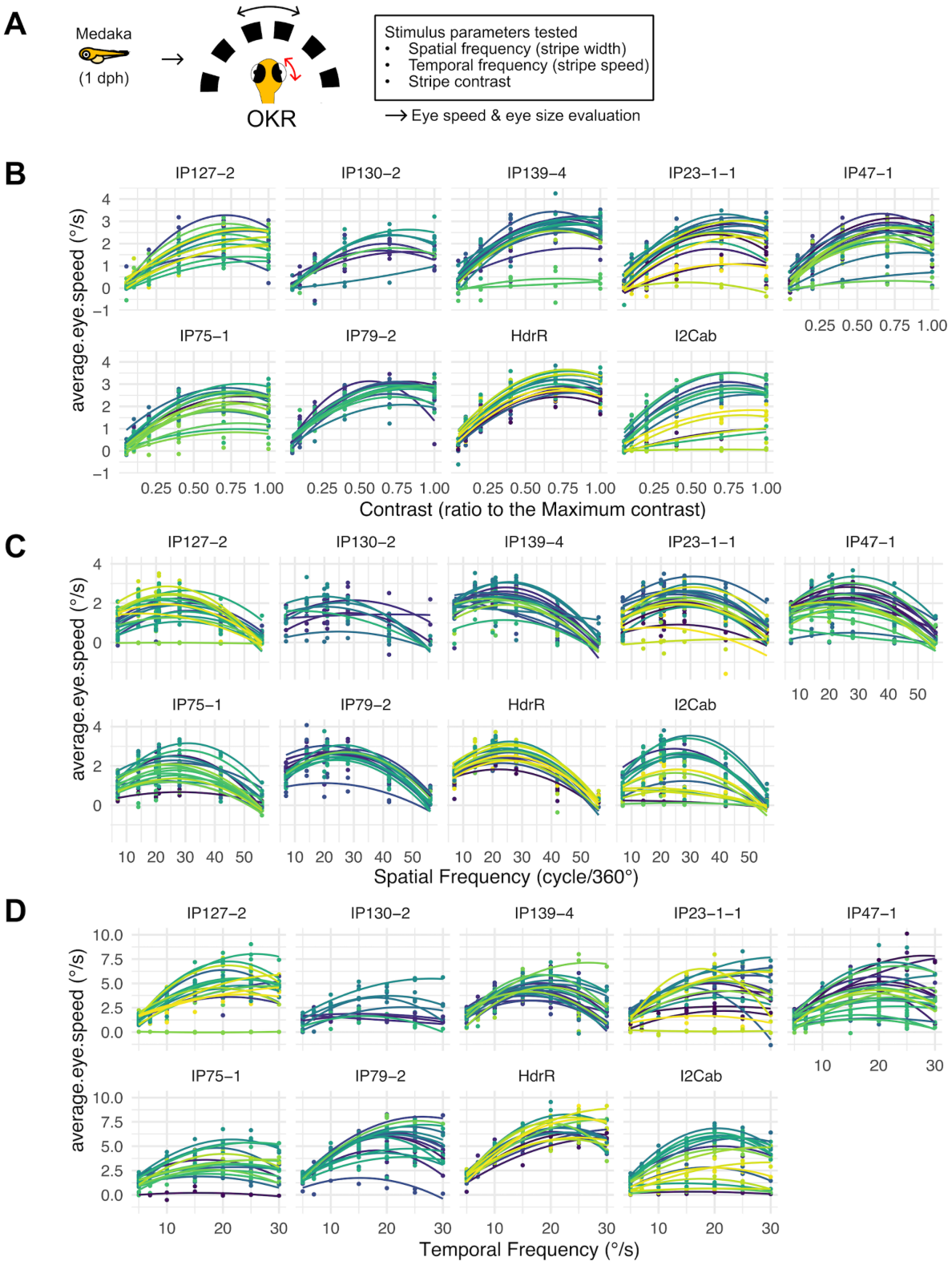


Figure 2.22 Assessment of Visual Sensitivity in MIKK Panel Strains Using an Optokinetic Response Assay.

(A) Schematic of the optokinetic response (OKR) assay. One-day-post-hatch (1 dph) medaka hatchlings from the seven selected MIKK panel strains (used for F2 segregation analysis) and two traditional inbred strains (HdrR and I2Cab) were embedded in 6% methylcellulose and exposed to moving black-and-white stripe stimuli. Stripe patterns were systematically varied in contrast, spatial frequency (stripe width), and temporal frequency (motion speed). (B–D) Dot plots with fitted line

showing average eye velocity (in °/s) of individual hatchlings in response to stripe motion under varying stimulus conditions: (B) contrast levels, (C) spatial frequency, and (D) temporal frequency. Eye size was also measured for each individual.A subset of hatchlings from strains IP130-2, IP47-1, IP23-1-1, and I2Cab exhibited reduced OKR performance across all three stimulus conditions in comparison to HdrR, indicating potential impairments in basic visual sensitivity or sensorimotor integration.

Visual acuity and function are often correlated with eye size (Caves et al., 2017). Therefore, I measured the eye size of each individual using microscopic images captured during OKR experiments (Figure 2.23 A). Retinal outlines, viewed dorsally, were extracted through thresholding and fitted with ellipses to obtain precise measurements of retinal area as well as major (retinal length along the nasal-to-temporal axis) and minor diameters (axial length). Box plots comparing the measurements across the strains (Figure 2.23 B,C) revealed that retina of IP75-1, I2Cab, and IP130-2 exhibited larger retinal areas and greater major diameters (nasal-to-temporal axis) compared to the other strains. Additionally, IP75-1 displayed a notably elongated retina in the anterior-to-posterior direction (Figure 2.23 D).

To complement behavioral and morphological analyses, electroretinography (ERG) was used to assess photoreceptor function and retinal output (Figure 2.24 A). Although this was a preliminary investigation, it aimed to test for potential differences across selected medaka strains. Because ERG responses in medaka have not been well characterized across strains, particularly at early developmental stages such as 1 dph, zebrafish were included as a comparative reference. ERG data (Figure 2.24 B-B‴) showed clear differences between zebrafish and medaka strains. As depicted in Figure 2.24 C, medaka strains exhibited significantly weak b-wave amplitudes compared to zebrafish, indicating weaker retinal responses under identical light stimuli. Furthermore, when normalized to the b-wave amplitude at first light stimulus (Figure 2.24 D), medaka strains showed a sharp decline in b-wave amplitude at log -1 intensity, suggesting photopigment bleaching and a prolonged recovery time following light exposure. HdrR, I2Cab, IP127-2 and IP79-2 strains generally exhibited lower b-wave amplitudes compared to the other strains; however, due to limited sample sizes, these differences could not be considered statistically conclusive.

Collectively, these behavioral (OKR), structural (retinal area), and electrophysiological assessments (ERG) highlight the diversity of mechanisms underlying visual function among the selected MIKK strains and their suitability for genetic mapping of visual traits.

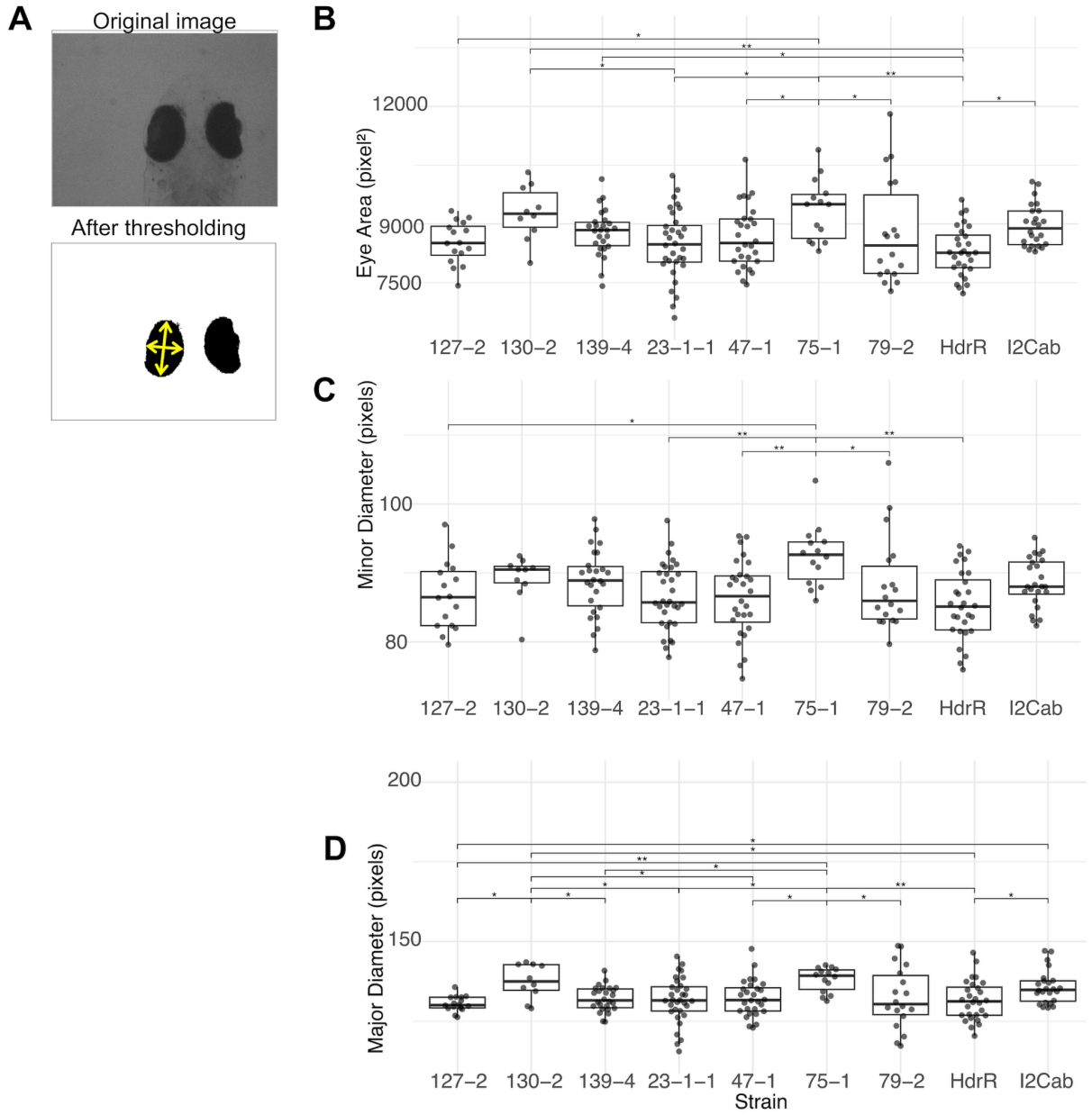


Figure 2.23 Strain-Specific Differences in Retinal Area and Diameter Measured During OKR Assays.

(A) Representative microscopic image acquired during optokinetic response (OKR) experiments. Following thresholding, the retinal outline was fitted with an ellipse, and both the area and diameters were measured for each retina (two datapoints per hatchling). (B-D) Box plots showing quantitative comparisons of retinal morphology across strains selected for F2 segregation analysis and two traditional inbred strains: (B) retinal area, (C) minor diameter, and (D) major diameter. Larger retinal area and diameters were observed in strains IP75-1, I2Cab, and IP130-2 compared to others. Statistical analyses were conducted in R, with Dunn's post hoc tests applied for multiple comparisons with Benjamini–Hochberg (BH) corrected. Significant differences between groups are denoted by asterisks (* $p<0.05$, ** $p<0.01$).

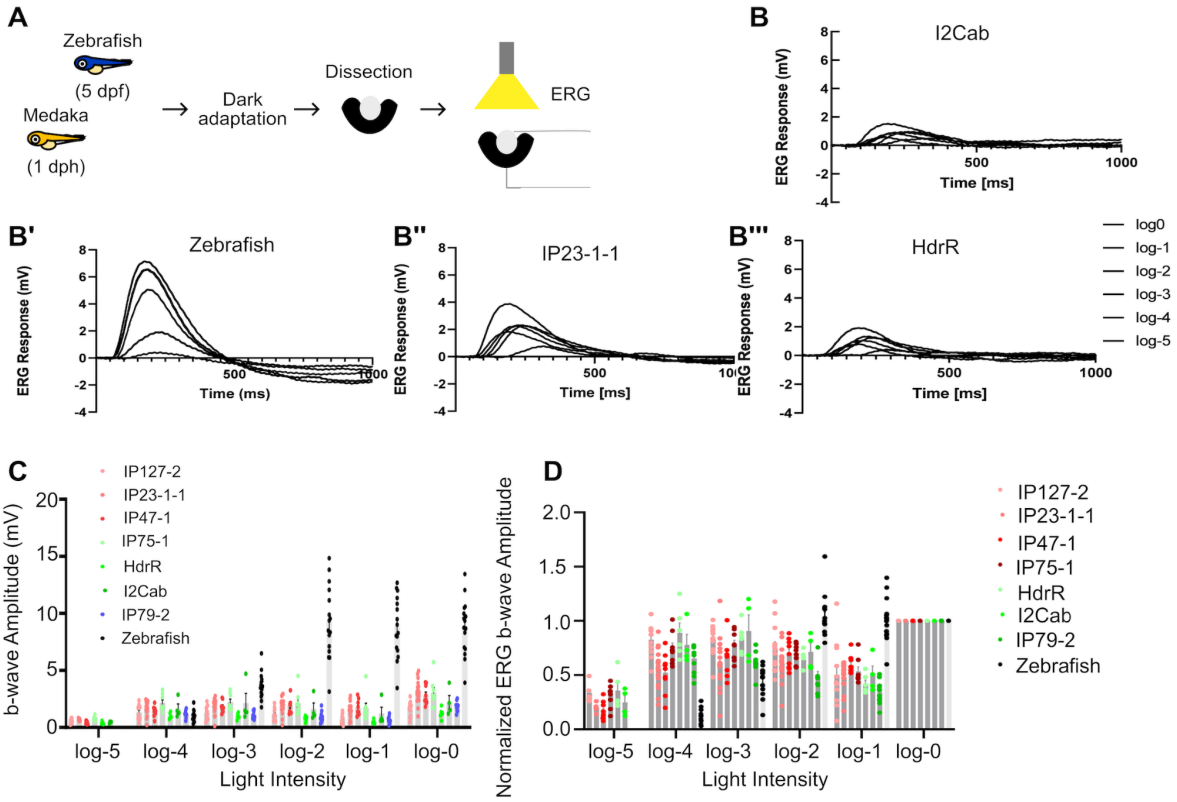


Figure 2.24 Comparison of Electroretinogram (ERG) Responses Between Zebrafish and Medaka Strains.

(A) Schematic diagram illustrating the experimental setup for ERG recordings. (B–B‴) Representative ERG responses (in mV) recorded from zebrafish and three medaka strains: I2Cab, IP23-1-1, and HdrR. (C) Column charts comparing the b-wave amplitudes between zebrafish and medaka strains. Medaka exhibited lower ERG amplitudes compared to zebrafish, indicating weaker retinal responses under identical stimulus conditions. (D) Normalized b-wave amplitudes across light intensities (log scale), normalized to the amplitude at log -0. Medaka response saturates at low light intensity while zebrafish response saturates at a much higher light intensity. Experiments were conducted by Jingjing Zang (UZH) and the figures adapted from Jingjing Zang (UZH).

2.4 Segregation Analysis

To investigate the inheritance of visual function and related behavioral traits, I analyzed phenotypes in F2 hatchlings generated from crosses between selected MIKK panel strains. 2,394 F2 hatchlings were subjected to OMR at 1 dph.

Visual acuity and sensitivity were assessed by measuring the minimum stripe width required to trigger a response and calculating visual acuity (Figure 2.25 & 6.7). Furthermore, visual stimuli driven behaviors were evaluated across 6 parameters: swimming speed during stripe motion (at speeds of 20.6 and 61.8 °/s), speed during the pause phase (during pause and after the completion of stripe motion), total distance traveled, and unidirectional swimming). Across all measured parameters, F2 phenotypes generally exhibited a wide range of variation

between those of the founder (F0) inbred strains, indicating polygenic contributions and potential additive genetic effects. However, certain traits deviated from this pattern, suggesting additional genetic complexity or non-additive influences.

These findings highlight the segregation of key visual and behavioral phenotypes in the F2 generation, supporting the utility of this approach for investigating complex genetic traits.

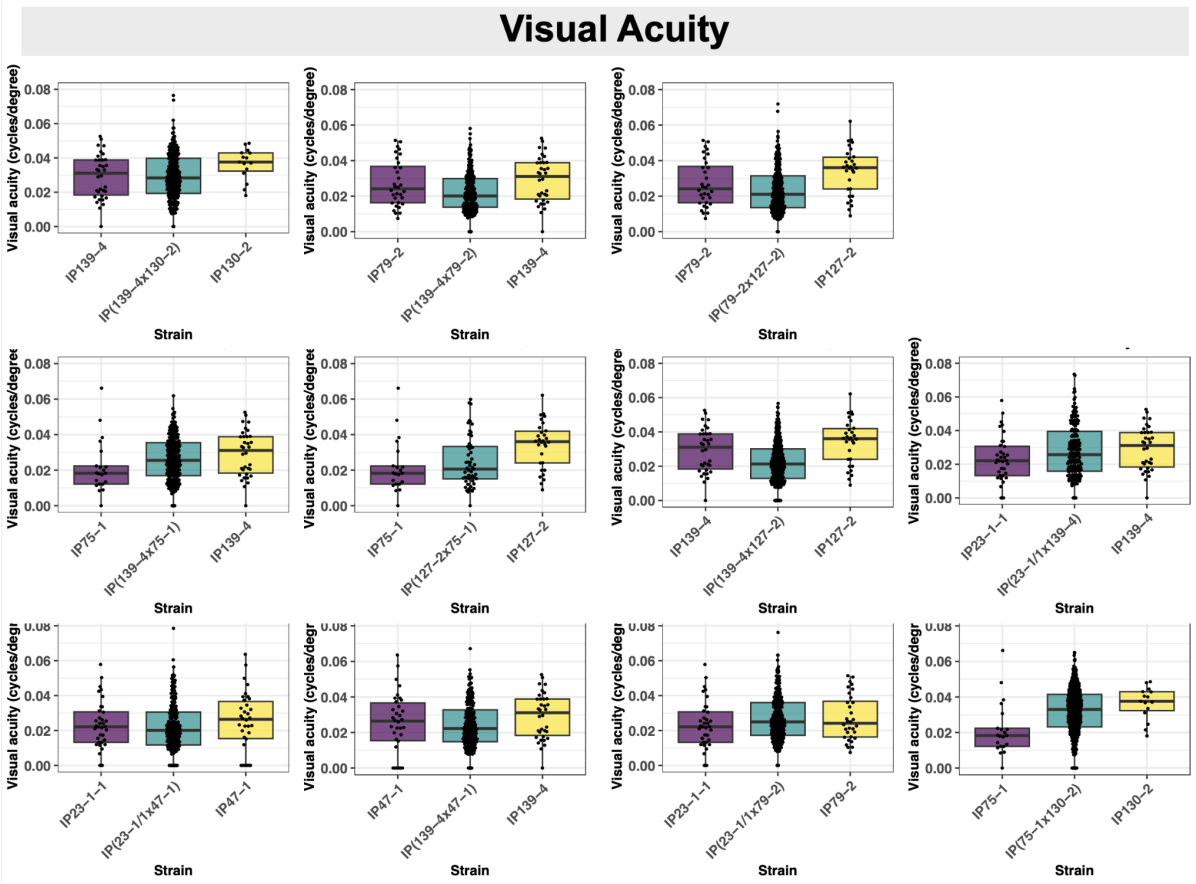


Figure 2.25 Cross-Dependent Segregation of Visual Function in F2 Offspring.
Box plots showing visual acuity of MIKK panel strains selected for strategic crossing and the respective F2 offspring.. The visual acuity in F2 hatchlings (turquoise) tend to distribute between that of the founder inbred strains (F0). Every dot corresponds to one individual hatchlings at 1 dph. dph: day post hatch

2.5 Association Testing and QTL Mapping

To associate the visual function and other investigated phenotypes of the F2 hatchlings with individual genotypes, whole-genome sequencing was performed to obtain the genetic information. After the OMR experiment, the genomic DNA of each F2 individual was extracted. Sequencing was performed with 150 bp paired-end Illumina short-read sequencing on a NextSeq2000 and P4 flow cells. Additionally genomic DNA was extracted from F1 adult brains for high coverage sequencing to identify homozygous divergent SNPs present in

the founder strains and for imputation (P3 flow cells were used). Out of 2,394 sequenced samples, 202 were excluded due to the low sequence coverage. After single imputation (Pierotti, Welz, et al., 2024), a genomic relationship matrix revealed that the F2 embryos cluster together per cross, indicating a higher relatedness (Figure 2.26).

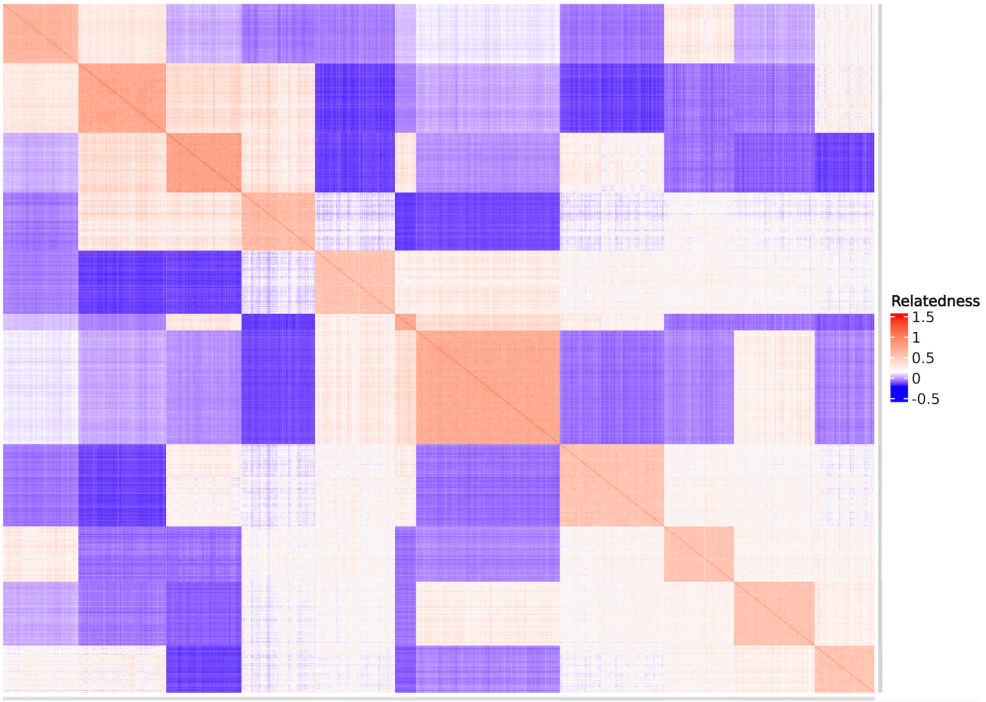


Figure 2.26 Heatmap of genomic relationship matrix using 2192 F2 samples from 11 cross. Genotype correlation levels are indicated for high relatedness in red and lower relatedness in blue. Samples are clustered based on relatedness, the clustering coincides with the cross patterns except for a few samples. Data produced together with Esther Yoo (EMBL-EBI). Figure adapted from plots generated by Esther Yoo (EMBL-EBI).

With initial QTL mapping attempts, approximately three loci demonstrated near-threshold signals when tested for the visual acuity (Figure 6.12 & Figure 2.27). Notably, five genetic loci exhibited associations that exceeded the Bonferroni-corrected significance threshold in relation to visual stimuli driven behaviors, including swimming speed during stripe motion (with the speed of 61.8°/s) and stripe pause, total distance traveled, and one directional swimming (Figure 6.12 & 2.27). These findings highlight the involvement of distinct genetic determinants in shaping complex visual-motor behaviors. These results lay the groundwork for future investigations into gene-gene interactions involved in visual and motor functions, ultimately contributing to a deeper understanding of eye development and the genetic basis of visual system disorders. Nevertheless, as the present analysis was conducted using a preliminary imputed dataset and non-optimized statistical model parameters, further validation through more comprehensive approaches will be performed in future work.

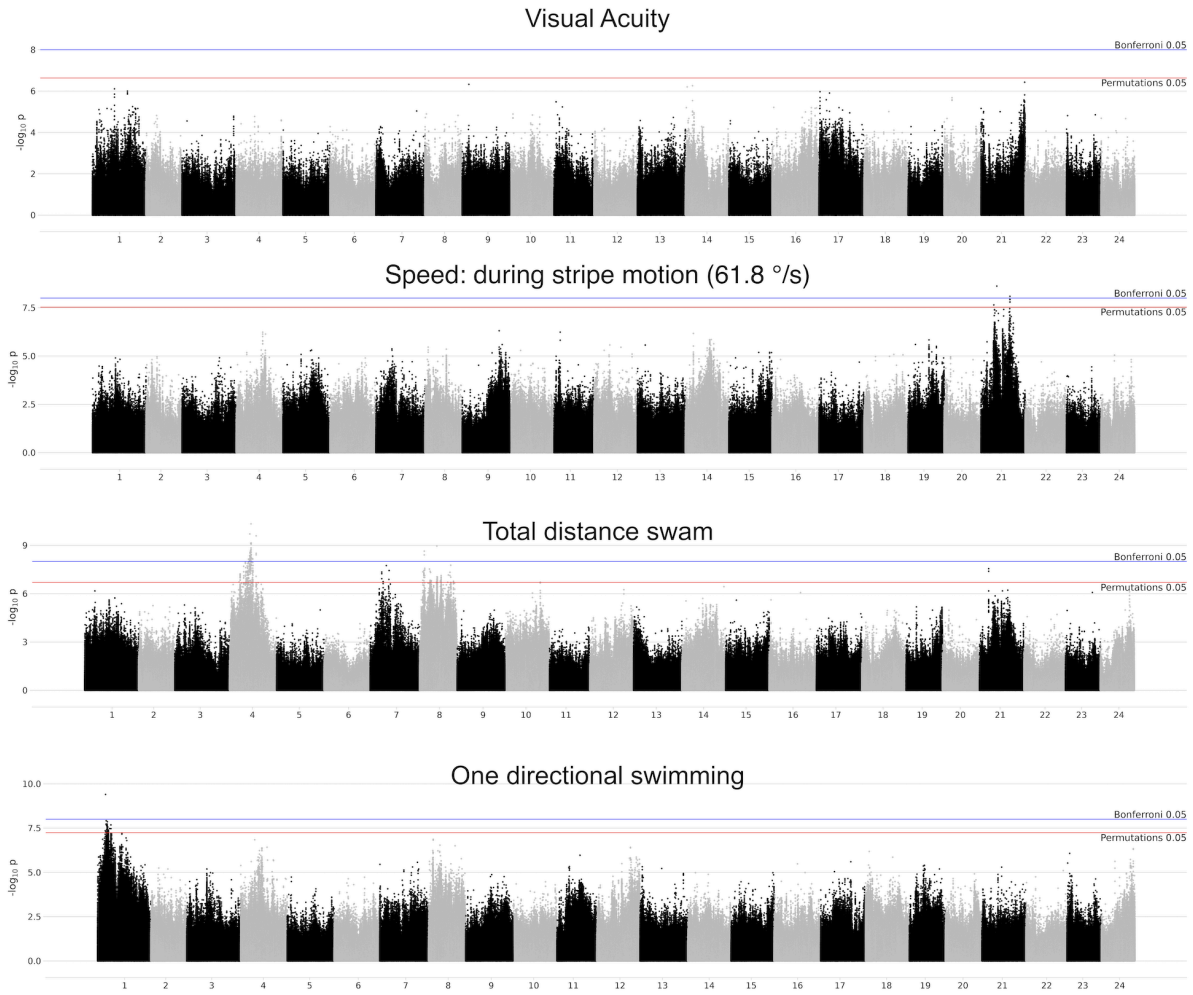


Figure 2.27 Manhattan Plots for Visual Function and Visual Stimuli Driven Behaviours.
A statistical association analysis was performed to link genetic variants with visual acuity, speed during stripe motion (61.8°/s), total distance traveled, count of one directional swimming phenotypes in the F2 population using a linear mixed model, with the strain of the parent cross and plate ID set as covariates. The phenotypes analyzed included: (top) the minimum stripe width that elicited a response, (middle) visual acuity estimated using the median distance of the top 25% of data points closest to the center, and (bottom) visual acuity estimated using the median distance of the top 10% of data points closest to the center. The significance threshold is indicated in red and was set by 10 permutations. Data produced together with Esther Yoo (EMBL-EBI). Figure adapted from plots generated by Esther Yoo (EMBL-EBI). Significance threshold indicated as red line ($\alpha = 0.05$), and with the Bonferroni-corrected significance threshold as blue line.

3

Discussion

The aim of this work was to characterize visual function and related behavioural phenotype across naturally driven medaka inbred population to eventually identify the loci that contribute to variation in visual function and its development. To achieve this, I first established two high-throughput optomotor response assays to evaluate visual system's capacity and an assay to mimic retinal degeneration in medaka. I assessed visual function (spatial frequency, color, and contrast sensitivity) as well as behavioural phenotypes of Zebrafish and various medaka strains by OMR assays. Moreover susceptibility to the light and regenerative ability were assessed across species and strains. I performed screening of OMR across the MIKK strains and phenotype various behaviours and selected 7 strains that are diverse in visual function and other behavioural traits and performed strategic cross and phenotyping and whole genome sequence of F2 generations which yielded preliminary result with five QTL. Additionally, I further analysed retinal function in the selected strains by further characterization, eye size assessment, OKR and ERG.

In the following sections, I will discuss the selections of findings and possible reasoning behind its diversity in retinal degeneration and visual function across species and strains.

3.1 Differences Between Zebrafish and Medaka in Light Susceptibility and Retinal Signaling Amplitude

3.1.1 Light Conversion to Signal in the Retina

The visual pathway is initiated when light passes through the cornea and lens, ultimately reaching the PRCs. In the absence of light, PRCs remain in a depolarized state due to the presence of cyclic guanosine monophosphate (cGMP)-gated cation channels in the outer segment membrane. These channels remain open in the dark, allowing a continuous influx of sodium (Na^+) and other cations, thus maintaining a relatively depolarized membrane potential (Yau & Hardie, 2009). Upon photon absorption, the chromophore 11-cis-retinal, covalently

bound to the opsin protein, undergoes photoisomerization to all-trans-retinal. This conformational shift activates opsin, triggering a cascade of intracellular events. The activated opsin stimulates the G protein called transducin, which in turn activates phosphodiesterase (PDE) (Cote et al., 2022). PDE hydrolyzes cGMP into GMP, resulting in a rapid decrease in intracellular cGMP levels. Consequently, the cGMP-gated channels close, leading to hyperpolarization of the PRC membrane. Although each transducing molecule activates a single PDE6 enzyme, each PDE can hydrolyze multiple cGMP molecules, resulting in significant signal amplification. To limit the duration of this response and restore the PRCs to its basal state, multiple regulatory mechanisms are engaged. In case of rhodopsin, Rhodopsin kinase phosphorylates activated rhodopsin, a process modulated by the calcium-binding protein recovering (Larhammar et al., 2009). Phosphorylated rhodopsin is then bound by arrestin, which prevents further activation of transducin and promotes rhodopsin inactivation. Concurrently, all-trans-retinal dissociates from opsin and is transported to RPE, where it is enzymatically converted back to 11-cis-retinal. The regenerated chromophore is subsequently transported back to the photoreceptor outer segment and rebinds to opsin, completing the visual cycle.

3.1.2 Light Susceptibility Differences between Zebrafish and Medaka

When the light constantly activates visual cycle, toxic byproducts such as all-trans-retinal (Kohno et al., 2013) and N-retinylidene-N-retinylethanolamine (A2E) (Radu et al., 2005) that generate reactive oxygen, cause lysosomal dysfunction, and are proinflammatory (Mata et al., 2000; Sparrow et al., 2012). Mechanistically, prolonged cascade activity disrupts intracellular ion homeostasis, notably Ca^{2+} levels. Photoreceptor Ca^{2+} regulation is tightly linked to the light response, and when rhodopsin is overactivated, Ca^{2+} imbalance can occur, activating calpains and other Ca^{2+} dependent proteases that cleave essential proteins (Vinberg et al., 2018; J. Yan et al., 2024). These are key mechanisms underlying light-induced retinal degeneration.

In my thesis, I showed that the amount of apoptotic cells in PRCs after intense metal halide lamp exposure differs between medaka and zebrafish and medaka (Cab strain) were more resistant to LIRD. The differences could potentially be explained by the difference in visual pigment regeneration efficiency or antioxidative response. Increase antioxidative response in the retina reflects potential protective effects against retinal degeneration (Ren & Léveillard,

2022). One study indicated that medaka hatchlings may have a greater resistance to oxidative stress induced by okadaic acid possibly through targeted regulation of antioxidant enzymes such as superoxide dismutase (SOD), glutathione reductase (GR), catalase (CAT), and Glutathione peroxidase (GPx) (Figueroa et al., 2022). Although the antioxidative response in the medaka retina has not been studied, it would be worthwhile to compare it with zebrafish to gain further insights into the underlying mechanisms.

Ocular pigmentation levels are another contributing factor. Both species have pigmented RPE that can absorb light and shield PRCs. However, melanin can have positive and negative impacts. In zebrafish, it was observed that pigmented individuals under intense light actually suffered more rod damage in some conditions due to melanin granules heating and exacerbating local oxidative stress (Weber et al., 2013), an effect that is absent in albino zebrafish. This might explain the differential survival rate after the LIRD across inbred strains with differing levels of body pigmentation.

The difficulty in inducing consistent and stable retinal degeneration limited the feasibility of conducting a comprehensive screen across the entire strain panel. However, the observed variability in the degree of retinal degeneration suggests the presence of strain-specific underlying mechanisms that contribute to differences in light susceptibility. One can look at expression levels of recovering, arrestin, that modulate the visual cycles levels with qPCR after injury to evaluate the possible differences in the toxic byproducts accumulation levels. Moreover single cell or bulk RNA sequencing and mass spectrometry-based analysis can be employed to uncover transcriptomic, proteomic and metabolomic differences among the strains, providing insights into pathways and gene expression profiles associated with differential vulnerability to retinal damage.

3.1.3 Electroretinogram Differences between Zebrafish and Medaka

When comparing ERG responses across light intensities, zebrafish demonstrated high response amplitudes during log-0 to log-2 light levels, with a noticeable drop at log-3 (approximately half of the log-2 amplitude) and a further reduction at log-4 (less than one-third to half of log-3 amplitude). This pattern suggests that zebrafish ERGs at early developmental stages are primarily cone-driven, with diminished performance under dim light conditions. It is consistent with reports indicating that zebrafish rods, while

morphologically present, remain functionally immature until approximately 15 days post-fertilization (Bilotta & Saszik, 2001).

In contrast, medaka exhibited lower ERG amplitudes even at the brightest intensity (log-0), generating responses that were only one-quarter to one-third the magnitude of zebrafish under the same conditions. At log-1, medaka responses were about half of those at log-0, and responses at log-2 to log-4 remained relatively stable, approximately two-thirds of the log-0 value. This data of at log-1, medaka exhibiting a lower response than at log-2, suggests slower photoreceptor recovery (possibly slower in regenerating pigments). The amplitude in medaka was relatively low and stable across a wide range of light intensities (log-1 to log-4), pointing to a potential rod-dominated scotopic ERG profile with sensitivity to low-intensity stimuli. In cone photoreceptors, a faster alternative regeneration pathway involving Müller glia (MG) exists in some species including zebrafish (Xue et al., 2017). The observed slower recovery in medaka, especially, may indicate less efficient retinoid cycling, inefficient retinoid transport proteins, or a lack of cone-specific regeneration pathways such as the MG-dependent cycle. This could imply that medaka either rely more heavily on rods for vision at early developmental stages or that their cones possess an underdeveloped alternative pathway for chromophore recycling. This pattern may indicate a greater reliance on rod photoreceptors in medaka at early stages, in contrast to the cone-dominated responses observed in zebrafish larvae at dark adapted condition.

In previous studies using full-body electroretinography (ERG), medaka and zebrafish displayed comparable response amplitudes under similar experimental conditions (Makhankov et al., 2004). However, in the ERG experiment performed with the dissected retina, medaka showed consistently lower ERG amplitudes across most tested conditions. Several factors may account for this discrepancy. One possibility is the anatomical difference in retinal structure and sampling: although the medaka retina is slightly larger than that of zebrafish, the same-sized electrode may cover a proportionally smaller retinal area, potentially sampling a region with lower photoreceptor density or activity. Another reasoning relates to PRCs composition and opsin gene expression.

Together, these findings highlight species-specific differences in retinal sensitivity and photoreceptor function between medaka and zebrafish. Further investigation, including testing ERG response with different light sources of specific wavelength or under photopic

condition, opsin expression evaluation, and photoreceptor density quantification, will be essential to elucidate the underlying molecular and structural bases of these functional differences.

3.1.4 Visual Cycle Speed and Susceptibility to the Retinal Degeneration

Electroretinography (ERG) data indicate that medaka exhibit a longer recovery time following light stimulation compared to zebrafish. Moreover, although the data are preliminary, one can also speculate on strain-specific differences in medaka. A closer examination of the response amplitudes at log-1 and log-2 reveals that I2Cab showed less increase at log-2 compare to the IP47-1, indicating I2Cab requires longer to recover from photobleaching. Such delayed recovery may indicate a slower regeneration of the visual chromophore, specifically the conversion of all-trans-retinal back to 11-cis-retinal.

Interestingly, this slower regeneration may confer a protective advantage. Prolonged presence of all-trans-retinal in the retina is known to be toxic, promoting the generation of reactive oxygen species (ROS). In medaka, slower 11-cis-retinal regeneration may reduce the accumulation of the harmful byproducts generated while 11-cis-retinal regeneration. This mechanism could underlie the greater resistance to light-induced retinal injury observed in medaka compared to zebrafish or differences in light resistance among strains.

Another potential explanation for the observed differences in light susceptibility between zebrafish and medaka may lie in natural selection shaped by their distinct native habitats. Zebrafish hatchlings inhabit shallow freshwater streams or ponds, frequently unshaded with aquatic vegetation (Spence et al., 2007). These environments likely impose selective pressures that favor high temporal resolution and spatial acuity, potentially for detecting predators or navigating complex environments. As a result, zebrafish larvae may rely primarily on cone-driven vision during early development and may be less active under dim light conditions, given the limited utility of scotopic vision in such bright environments (Spence et al., 2007).

In contrast, medaka hatchlings are commonly found in shaded, slow-moving waters such as rice paddies or weedy ponds, where ambient light is dimmer and more diffuse. These

environmental conditions may create a selective advantage for early rod function, enabling medaka larvae to forage or detect predators in low-light settings such as murky water. Although rod photoreceptors typically develop later in many fish species, the early presence and function of rods in medaka suggest evolutionary pressure favoring scotopic vision even at larval stages. This early rod expression could provide medaka with broader spectral and intensity sensitivity, supporting survival in more variable or low-light environments (Hilgers & Schwarzer, 2019).

3.2 Characterization of Optomotor Response with Stimulus from Below in Medaka Hatchlings

In linear-OMR assay, OMR can be reliably elicited in medaka as early as 0 dph, even when visual stimuli are presented from beneath. Notably, responsiveness to narrower stripes can be enhanced through training, with the degree of improvement differing between strains. More specifically, hatchlings of the HdrR strain exhibited greater trainability to narrow stripe stimuli compared to those of the Cab strain, while medaka strains from the MIKK panel showed distinct spectral sensitivity profiles relative to both HdrR and Cab. These observations underscore the role of genetic background in shaping visual processing capabilities in medaka.

The observed trainability may be linked to enhanced retinal processing or increased selectivity of neurons for visual stimuli. In vertebrates, the retina is not merely a passive transmitter of signals to the brain. It also adapts dynamically to light conditions, sharpens visual input, and includes feedback loops that refine signal processing depending on the stimulus (Vlasiuk & Asari, 2021). During training, retinal ganglion cells may develop contrast gain control and adaptive responses, effectively boosting the computational capabilities of the retina (Gollisch & Meister, 2010; Yedutenko et al., 2020).

In mammals, cognitive trainability has been extensively studied, showing that with experience, individuals learn to better discriminate task-relevant visual cues and optimize attentional timing, leading to significant changes in early-stage sensory cortical neurons (Poort et al., 2022). Since hatchlings have also been shown to form and sustain attention to visual stimuli, it is plausible that training-induced improvements in OMR may reflect increased awareness and attentional engagement (Parker et al., 2012).

Color vision is governed by the differential absorption properties of various cone opsins. Medaka possesses eight types of cone opsins and one rod opsin, providing broad spectral coverage across the ultraviolet, blue, green, and red regions (Matsumoto et al., 2006). While the Cab and HdrR strains exhibited similar response patterns across all colored stripes tested, the MIKK inbred strains showed a modest reduction in responsiveness. These strain-specific differences may stem from variations in opsin gene expression levels.

3.3 Strain Differences in Sensory and Behavioral Diversity

Traditional inbred strains such as HdrR and Cab offer valuable insight into strain-specific visual performance and behavioral diversity. Differences in spatial, color, and contrast sensitivity, as well as behavioral strategies for tracking motion, emphasize the importance of genetic background in shaping sensory and cognitive traits. Understanding this diversity is crucial for dissecting the genetic architecture of visual function and behavior, and for developing standardized behavioral assays that account for inter-strain variability.

3.3.1 Potential Visual Impairment in Traditional Wild Type Strains (Cab)

Notable differences in optomotor behavior were observed among strains. HdrR and Cab strains exhibited a tendency toward one-directional swimming behavior, with approximately one-third of individuals following the stripe motion consistently, while others either swam in the opposite direction or remained restricted to swimming in a single direction regardless of the stimulus. In contrast, eyeless medaka, whose eyes were surgically removed at 0 dph, lacked the continuity of directional swimming and instead moved randomly. These hatchlings had experienced visual input prior to eye removal, and thus their behavior may differ from hatchlings that were congenitally blind or had never perceived light. The eyeless group also showed reduced response continuity (continuous swimming in CW or CCW direction for a maximum of 35 seconds) and lower total response values (less than 300), indicating a disruption in visual driven behavior. Interestingly, a small subset of Cab hatchlings also showed low response metrics, those with total response values around 200 and continuity scores near 10. On the other hand, the lowest continuity score recorded in HdrR individuals was approximately 20. Considering that in OKR, a few individuals of I2Cab strains (Cab strain maintained in pairs over several generations to increase inbreeding levels), no or very low responses across all contrast, spatial frequency, and temporal frequency tests. This

suggests that a subset of the original Cab population may carry visual functional defects. Supporting this possibility, Dr. Lucie Zilova once reported observing lens cloudiness in some Cab hatchlings. Taken together, the observed one-directional swimming behavior could be explained by individuals with blurred vision detecting motion from the moving stripes, but lacking sufficient visual resolution to interpret the directionality correctly, potentially triggering a panic-like response instead of a directed optomotor reaction.

3.3.2 Effect of Acclimation Time on Behaviour

Acclimation time was also assessed as a factor influencing optomotor behavior. While overall changes were modest, Cab individuals showed a slight increase in response rate after extended acclimation. However, in terms of visual sensitivity, Cab performed better after 5 minutes of acclimation than after 1 hour, suggesting a possible decrease in responsiveness due to reduced attention to the environment. On the other hand, HdrR hatchlings exhibited higher response rates overall, and their sensitivity increased with longer (1 hour) acclimation, indicating potential strain-specific differences in adaptation and awareness to environmental cues. These results also suggest that experimental complexity and cognitive load may interact with visual sensitivity and motivation differently across strains.

3.3.3 Factors Influencing OMR Behaviour

OMR performance is influenced by a diverse set of sensory, cognitive, physiological, and environmental factors. These include: (1) Visual sensitivity shaped by photoreceptor function, contrast sensitivity, and spatial/temporal resolution. (2) Light levels. (3) Cognitive factors such as habituation, attention, learning, motivation (including curiosity and exploratory drive). (4) Boredom, anxiety or stress (often reflected in erratic behavior), and arousal state (ranging from hypo to hyperactivity). (5) Physical capability, including muscle development, fatigue resistance or overall health condition. (6) Environmental factors, including water temperature and circadian rhythms.

From the OMR assay, multiple quantifiable phenotypes can be extracted, potentially correlating with the factors mentioned above. Minimum stripe width eliciting a response, serving as a proxy for visual sensitivity. Swimming speed during stripe motion especially faster stripe, reflecting both physical capability. Swimming speed during pause phases or after the experiments, can be linked to basal activity levels or fatigue. One-directional

swimming behavior, possibly indicating either visual defects like lens clouding or reduced motivation/attention. Total distance traveled, which can reflect overall responsiveness, health, or engagement. Although visual acuity was also calculated, it is important to note that the distance between the hatchling and the stripe likely has a greater influence on acuity calculation than stripe width alone. Since hatchlings do not necessarily swim at their visual limit (e.g., the furthest possible point from the stimulus), the minimum stripe width eliciting a response may serve as a more relevant indicator of visual sensitivity.

3.3.4 Strain-Specific Differences in Visual Sensitivity and Behavior

Among the medaka strains tested, Kaga, a northern Japanese population (Oryzias. Sakaizumi), demonstrated superior sensitivity to black/green and black/red combinations. Kaga strain is originated from the wild population in Kaga city, Ishikawa prefecture, Japan. The enhanced red and green sensitivity observed in Kaga strain may reflect natural selection in response to local light environments. For instance, Ishikawa area receives approximately 300 fewer sunlight hours annually compared to Aichi, where the MIKK panel was originally established (e.g., sun light hours per year in Ishikawa was 2029.8 hours and in Aichi was 2378.4 hours in 2023). Weather in Ishikawa, characterized by cloudy, snow-heavy winters, results in exposure to lower-intensity light that is enriched in longer wavelengths. Slightly elevated sensitivity to red and green light in such environments may offer a selective advantage, particularly by improving plankton detection or feeding efficiency under dim or diffuse illumination during the winter months. It would be valuable to examine opsin expression levels alongside behavioral sensitivity to specific colors using OMR.
Strains such as I2Cab and HO5 exhibited a distinct behavioral profile, with a greater proportion of individuals showing one-directional swimming and increased swimming speed even after the stimulus presentation had ended. This may suggest elevated baseline activity or a degree of disengagement from visual stimuli, possibly reflecting strain-specific differences in arousal, motivation, or sensitivity to environmental cues.

In contrast, the QuiH strain performed particularly well in response to thinner stripe patterns, exhibiting greater visual acuity and precise speed adjustment during OMR. The lack of pigmentation in albino animals is known to affect visual processing, particularly in terms of spatial resolution and light scattering. RPE plays a critical role in absorbing excess light that passes through the PRCs, thereby preventing intraocular light scattering and improving image

clarity. Studies in albino adult mice (Yeritsyan et al., 2012) or human (Moshkovitz et al., 2024) suggested visual impairment. In Zebrafish fish, albinism has no major impact on contrast sensitivity, but has reduced visual acuity (Müller, 2011). However, a recent study on Mexican Tetra found that albino mutants raised in darkness outperformed their wild-type siblings in capturing prey under dark conditions (Choy et al., 2025).

QuiH2 is originated from a cross of Quintet mutant and Heino albino mutants. Although the precise mutation site remains unknown, it has been suggested that more than five causative mutations contribute to its lack of pigmentation. In QuiH2 that lack eye pigmentation, reduced or lack of melanin in the RPE could allow more light to reflect internally, possibly leading to increased backscatter toward the PRCs from the sclera or other intraocular structures. This may partially explain the enhanced contrast sensitivity and sensitivity to specific color stripe combinations.

3.3.5 Comparisons with Zebrafish and Implications for OMR Performance

When compared to medaka, zebrafish displayed similar responsiveness to color and contrast stimuli but showed lower spatial sensitivity and had difficulty maintaining performance at higher stripe speeds. These differences may reflect both developmental and functional divergence in visual system organization between species. This developmental lag, along with differences in ecological niche and behavior, may underlie the observed limitations in spatial tracking and high-speed motion responsiveness.

3.4 OMR Assays across MIKK Strains Discovered Varying Phenotype

3.4.1 Behavioral Variation, Trainability and Sensitivity across Different Strains

MIKK strains exhibited phenotypic variation across multiple parameters, including swimming speed, total distance traveled, one-directional swimming, and visual acuity. Notably, the variation observed among MIKK strains were more prominent than that among traditional inbred strains.

Increased stripe speed elicited more robust response at later developmental stages and was more effective in maintaining response rate. Repeated exposure to OMR over the developing time course enhances sensitivity, potentially due to increased awareness. Strain-specific behavioral characteristics were also evident. HdrR exhibited sign of reduced engagement with repetitive OMR, IP58-2 remained consistently responsive, while IP72-2 frequently ignored the stripe motion and exhibited one-directional swimming behaviour. These findings highlight strain-dependent differences in stimulus engagement and attentional dynamics under repeated visual stimulation.

3.4.2 Eye Size and Visual Performance

In OKR assays, HdrR showed response profiles comparable to zebrafish, particularly in terms of temporal frequency tracking. IP127-2, IP139-4, IP79-2, and IP47-1 demonstrated strong visual sensitivity, especially under varying spatial and temporal frequencies. In contrast, IP130-2 performed poorly in both temporal and spatial frequency conditions. Given the cloudiness observed in the lens at the adult stage, it is possible that lens defects were already present at the hatchling stage. IP75-1 exhibited especially weak OKR performance, failing to track well at higher speeds. In temporal frequency test hatchlings reached peak eye velocity around 20°/s, indicating this may represent an optimal stimulus speed for medaka hatchlings at 1 dph.

Eye size varied significantly across strains and may influence visual acuity and photoreceptor spacing. I2Cab, IP75-1 and IP130-2 possessed notably larger eyes, with IP130-2 showing longer size along the anterior-posterior axis. IP75-1 and IP130-2, despite its large eyes, performed poorly in OKR and OMR, especially at higher temporal frequencies. In OMR experiments, while IP75-1 displayed one-directional swimming behavior more frequently than the others, IP130-2 failed to maintain consistent response across the stripe width. Conversely, those that had rather small eyes such as HdrR and IP127-2, often exhibited better visual acuity and higher visual sensitivity.

It is also possible that the smaller eyes observed in certain strains may actually be the optimal size, in which lens and retinal growth are coordinated well. In contrast, the larger eyes may result from overproliferation or disproportionate retinal growth. If lens development proceeds at a normal pace while the retina increases beyond the focal length supported by the lens, the resulting optical mismatch could lead to blurred or low-contrast vision. A potential next step

would be to assess correlations between OMR performance, eye size, photoreceptor density, and lens size, to further clarify how structural features relate to visual function.

3.4.3 Eye Morphology Diversity in the MIKK Panel Suggests Polygenic Inheritance

Over the past years I have been involved in rearing of MIKK panel, and throughout this work, specific eye phenotypes (large or small) were consistently observed in some strains more than the others. This indicates a specific or combinations of genetic variants influencing eye size are distributed in different frequencies among these strains. In the earlier generations of IP47-1 (F18), I intentionally collected more from individuals with particularly small eyes which ended up obtaining offsprings that predominantly exhibited the small eye phenotypes. However, in more recent generations, eye size variation has become more or less frequent, with a notable increase in asymmetric phenotypes, where only one eye is smaller than the other. This pattern suggests a complex genetic basis, possibly involving polygenic or environmental interactions. Meanwhile, in IP139-4 which originally had large eyes, shows variability in adult eye size, with some individuals maintaining large eyes while others display small eye phenotypes in the current generations (F21).

For the strategic crossing, on top of visual functional phenotype using OMR, the eye morphology observed at adult stage was considered. Notably, strains such as IP23-1-1 (wider minimum stripe width eliciting response) and IP47-1 (lower response rate) with low visual function at hatchling stage, also had some fraction of individuals with small eyes. Similarly, IP130-2 which had low visual function had defects in lens opacities and inflated cornea resembling cataract in the adult stage. The observed correlation between functional defects in vision at hatchling stages (identified with OMR) and eye morphological defects in eye at the adult stages supports the inclusion of variants associated with eye size or malformation in the eye in the panel.

3.4.4 Baseline Heart Rate and Activity Level Correlates

Heart rate of 4 dpf embryos measured at 21°C varied across MIKK strains (Welz et al., 2025). Even though the measurement was performed at earlier stages, as heart morphology is already mature at this stage basal heart rate may correlate with overall activity levels and responsiveness. The strategically selected strains represented both ends of the heart rate spectrum (Welz et al., 2025). IP130-2 exhibited the second-lowest heart rate of all MIKK

panel (median of ~84 bpm), and had low swimming speed during both stripe motion and pause phases. By contrast, IP79-2, with the highest heart rate of all MIKK panel with median of ~111 bpm, exhibited the highest swimming speed after the experiment time periods from the selected strains and the ability to adjust swimming speed in response to faster stripe motion. Most other strains, IP47-1, IP23-1, IP139-4, IP75-1, IP127-2, clustered around 95 bpm, suggesting a shared basal physiological state. These differences may reflect strain-specific metabolic or cardiovascular regulation, potentially affecting arousal state and behavioral responsiveness. Notably, higher basal heart rate may enable more dynamic speed adjustment and enhanced responsiveness to motion stimuli.

3.6 Perspective

3.6.1 Approaches to Improve the Detection of QTL

Five QTLs were identified across the behavioural metrics analyzed. The next logical step will be to further investigate the identified QTLs by functionally validating candidate genes using genome editing approaches such as CRISPR/Cas9 or base editing. Comparing the effect sizes of these edits across multiple MIKK panel strains will enable the evaluation of gene-gene interactions and strain-specific genetic backgrounds. Additionally, expression profiling of candidate genes, using in situ hybridization or assessment in single-cell RNA sequencing datasets throughout the retinal development, will provide further insights into their spatial and temporal roles during eye development. Future efforts may also include exploring additional behavioral phenotypes to perform QTL mapping to expand the candidate lists.

Although the current preliminary QTL mappings did not reveal significant loci directly linked to visual acuity, indicative of visual impairment or enhanced visual performance, further refinement of phenotyping methods for visual function and optimization of analytical parameters may enhance the detection of more vision-specific QTL. Notably the identified QTL related to “one-directional swimming behavior” and the “swimming speed during faster stripe phase” may also reflect genetic loci influencing visual function, as the one-directional swimming phenotype could possibly result from near-blind vision or habituation due to overly simplistic stimuli relative to the individual’s visual capacity. Similarly, QTL associated with swimming speed may represent genetic factors influencing general locomotor fitness as

well as visual perception capabilities, as the lack of response could further reflect deficiencies in temporal frequency sensitivity or restricted visual fields.

Stripe width and visual acuity, as used in this study, may not have served as ideal proxies for QTL mapping for visual function. This limitation could stem from the exclusion of individuals exhibiting extreme phenotypes (e.g., hatchlings with complete or near blindness), as well as the possibility that differences in response to varying stripe widths were relatively subtle, which may reduce the resolution of the trait and limit its utility for fine-scale QTL mapping. Moreover, the calculation of visual acuity is highly influenced by the distance from the dish rim (where the stripe pattern appears), introducing substantial variability. The distance cannot be controlled artificially and may have been affected by stochastic factors such as the hatchling's initial position within the arena.

To enhance the resolution of visual function metrics, more refined phenotyping approaches may be beneficial. One such strategy involves analyzing and mapping behavioral responses to each stripe width independently. For instance, one could calculate the sum or median of response values during each CW and CCW motion phases. These values can then be combined into a composite metric - either by multiplying or dividing them, or by summing the absolute response values and multiplying the total by +1 if the hatchling changed swimming direction in accordance with the stripe motion, or by -1 if it exhibited one-directional swimming regardless of stimulus direction. This metric would yield high positive values for individuals that responded appropriately to both CW and CCW stimuli, intermediate or near-zero values for those with weak or inconsistent responses, and negative values for individuals displaying strong one-directional swimming behavior. For example, calculating the metrics for a stripe width of 1.6 mm may help to enrich for hatchlings with exceptionally high visual sensitivity. These derived metrics could then be used for QTL mapping or dimensionality reduction analyses (e.g., principal component analysis) to identify underlying genetic loci. These derived traits could then be subjected to QTL mapping or dimensionality reduction analysis (e.g., principal component analysis) to identify genetic loci associated with variation in visual performance.
Additionally, responses to CW and CCW motion could be evaluated separately to determine the threshold stripe width for each direction. This strategy may help reveal direction-specific response patterns in individuals with unilateral visual impairments (e.g., fish with only one functional eye), who tend to exhibit biased swimming behavior until the stimulus becomes

sufficiently prominent to be detected by the impaired side - an estimation supported by observations of hatchlings congenitally lacking one eye.

It is worth noting that, although not previously mentioned, all F2 individuals underwent additional experiments to assess contrast and color sensitivity following the OMR test with varying stripe widths prior to the DNA extraction. These experiments included responses to five levels of grayscale contrast, as well as black and colored stripes at four contrast levels each for red, green, and blue. While this extended dataset has not yet been analyzed due to time constraints, it holds considerable potential for future QTL mapping and deeper characterization of visual function.

3.6.2 Potential Applications of the Infinity-Pool-Style OMR Assay

Future studies could further elucidate the neural mechanisms underlying distinct visuomotor phenotypes, such as one-directional swimming, blind-like behaviour, consistently following stripe motion, and occasional following behaviour, through advanced brain activity mapping techniques including c-Fos immunostaining and whole-brain imaging. Given that these behavioral phenotypes have all been observed within the Cab strain, targeted genetic modifications to introduce cumulative calcium sensors, could allow visualization of neural activity. These integrative approaches may provide critical insights into the specific neural circuitry underlying diverse behavioral responses.

Additionally, the established OMR platform offers substantial potential for assessing the effects of environmental exposures during development. For instance, it could be used to evaluate both behavioral and physiological outcomes following embryonic exposure to chemicals such as ethanol, environmental toxins like perfluorooctane sulfonate (PFOS), or emerging psychoactive compounds. Particularly psychedelics are gaining significant scientific and clinical attention for their potential therapeutic applications in treating various psychiatric disorders, but are known to produce highly variable effects in humans, suggesting a strong influence of genetic background. By leveraging the natural genetic diversity of the MIKK panel in combination with the sensitivity of high-throughput OMR assays, one could conduct comprehensive studies on gene–drug interactions, particularly those affecting sensory processing, perception, or hallucination.

Furthermore, the intergenerational effects of such exposures can also be investigated. For instance, experiments could be designed where adult fish are exposed to a given compound, and the behavioral phenotypes of their offspring are subsequently assessed. Alternatively, embryos themselves could be treated with the compound to evaluate early-life exposure effects. This assessment could begin with physiological phenotyping, such as heart rate monitoring during development, to evaluate stress responses or arousal. Following this, a battery of behavioral tests could be conducted to target distinct domains of neurobehavioral function. Specifically: Exploratory behavior in novel environments (e.g., hatchling's exploratory behaviour evaluated at first 30 minutes in the arena) reflects anxiety-like traits and spontaneous locomotor drive. Then responses to black/white background transitions serve as visual startle tests. The OMR test can evaluate sensitivity to color, contrast, and both spatial and temporal frequency. Finally, social interaction assays using animated conspecific images can be tested to possibly measure social recognition, preference or shoaling-like behaviors. Together, these approaches offer a powerful platform for studying how early or paternal exposure to compounds may shape neurobehavioral outcomes across generations, and for delineating the underlying interactions among drugs, genes, developmental processes, and epigenetic mechanisms.

3.6.3 From Medaka to Human: Implications for Genetic Diagnostics of CDEs and Drug Testing

In this thesis, I established a high-throughput OMR assay and performed behavioural phenotyping of genetically diverse MIKK panel strains. By utilizing strategic crosses, phenotyping in an F2 population, and conducting an association study, five QTLs associated with visuomotor responses were identified.

From the identified QTL regions, specific causative genes or loci can be delineated through targeted gene editing approaches, particularly given the high-throughput capacity of the established assay. This enables efficient functional validation of candidate genes and locus and facilitates fine-mapping within QTL intervals. By further analysis and testing the effect of mutation in these loci on differential strains, gene-gene interactions will potentially found, advancing our understanding of complex genetic networks relevant for human diagnostic applications. Indeed, despite the availability of whole-genome sequencing in humans, the high density of SNPs complicates the identification of causative mutations. Consequently, insights potentially gained from this study with the medaka model may contribute to the diagnosis and understanding of rare CDEs in humans.

Upon identification of potential causative mutations in human patients, these variants can be rapidly validated and functionally characterized in medaka, taking the advantage of the species' efficient genome editing capabilities and the high-throughput assays developed in this study. Additionally, the genetically diverse MIKK panel facilitates in-depth analysis of gene-gene interactions. Moreover, this infinity-pool-style OMR assay, together with MIKK panel also enables detailed exploration of gene-environment interactions, further enhancing our comprehension of CDEs.

Finally, the infinity-pool-style OMR setup developed in this study holds promise for studying modeling human diseases in medaka. Its scalability supports the functional evaluation of disease-related phenotype and possesses potential for high-throughput drug screening and toxicity profiling. Leveraging the genetic diversity present in the MIKK panel, strain-specific drug responses and side effect susceptibilities can be examined, offering a valuable preclinical tool for personalized medicine research, mechanistic studies and gene-drug interaction analysis in a vertebrate system.

Overall, the high-throughput OMR assays and experimental strategies presented in this study provide a robust foundation for future investigations aimed at dissecting gene–gene and gene–environment interactions, ultimately contributing to a deeper understanding of the genetic and environmental factors underlying variability in CEDs.

4
Conclusions

In conclusion, I have established two high-throughput OMR assay platforms and conducted a comprehensive behavioral screen across 84 MIKK panel strains, generating an information-rich dataset for further phenotyping and QTL mapping of visual, behavioral, and potentially other complex traits. Preliminary association testing identified five QTLs, providing a foundation for further validation, fine-mapping, and exploration of gene-gene interactions. Additionally, I successfully adapted LIRD model for medaka hatchlings, reportedly for the first time, revealing strain- and species- specific susceptibility to phototoxic damage. By integrating ERG, OKR, retinal function was further characterized and correlated with OMR responses. Collectively, this work offers novel insights into the visual system of genetically diverse medaka strains and establishes a robust framework for dissecting genotype–phenotype relationships in visual processing and eye development.

Ultimately, this work lays the foundation for future investigations into gene–gene and gene–environment interactions, advancing our understanding of the complex genetic and environmental factors that contribute to variability in congenital eye disorders.

5
Materials & Methods

5.1 Materials

5.1.1 Organisms

Table 5.1 Traditional inbred medaka strains, F2 fish and zebrafish used in this thesis. The KIT-MIKK strains were provided by Dr. Felix Loosli (Institute of Toxicology and Genetics, KIT, Germany). HdrR and HO5 were purchased from NBRP (https://shigen.nig.ac.jp/medaka/) and maintained at COS.

Organism	Species	Strain	Generation	Source
Medaka	*Oryzias latipes*	HdrR	F103-107	Laboratory stock, NBRP Medaka
Medaka	*Oryzias latipes*	Cab	F77-78	Laboratory stock
Medaka	*Oryzias latipes*	I2Cab	F4-5	Laboratory stock
Medaka	*Oryzias latipes*	HO5	F112-114	Laboratory stock, NBRP Medaka
Medaka	*Oryzias latipes*	QuiH2	F28	Laboratory stock
Medaka	*Oryzias latipes*	COS-MIKK strains	F22	Laboratory stock
Medaka	*Oryzias latipes*	KIT-MIKK strains	F22	Laboratory stock
Medaka	*Oryzias latipes*	IP(139-4x47-1)	F2	Laboratory stock
Medaka	*Oryzias latipes*	IP(139-4x127-2)	F2	Laboratory stock
Medaka	*Oryzias latipes*	IP(23-1/1x139-4)	F2	Laboratory stock
Medaka	*Oryzias latipes*	IP(23-1/1x47-1)	F2	Laboratory stock
Medaka	*Oryzias latipes*	IP(75-1x130-2)	F2	Laboratory stock
Medaka	*Oryzias latipes*	IP(23-1/1x79-2)	F2	Laboratory stock
Medaka	*Oryzias latipes*	IP(127-2x75-1)	F2	Laboratory stock
Medaka	*Oryzias latipes*	IP(139-4x79-2)	F2	Laboratory stock
Medaka	*Oryzias latipes*	IP(139-4x75-1)	F2	Laboratory stock

Medaka	*Oryzias latipes*	IP(79-2x127-2)	F2	Laboratory stock
Medaka	*Oryzias latipes*	IP(139-4x130-2)	F2	Laboratory stock
Medaka	*Oryzias sakaizumi*	Kaga	F3	Laboratory stock
Zebrafish	*Danio rerio*	AB/Beck	F8-9	Laboratory stock

5.1.2 Primers

Table 5.2 Primers used in this study.

Binding to	Direction	Number	Annealing	Oligo sequence 5' to 3' direction
HdrR sox3 allele	Reverse	JW10951	55.8	AATGGGCAACTTATTCTGTC
HdrR sox3 allele	Forward	JW10950	55.8	CTTTTCCAATTAATTGTATTTATT TAG
Cab sox3 allele	Reverse	JW10949	57.7	CATTTCTTGCTCTAAATGAC
Cab sox3 allele	Forward	JW10948	57.7	ATTGTATAATGGAAAAAACG

5.1.3 Buffers and Solutions

Table 5.3 Solutions and buffers used in this study.

Solution or Buffer	Ingredients
TE buffer	10 mM Tris-HCl pH 8.0, 1 mM EDTA
1x PTW	0.1% Tween20 solved in 1x PBS
Optical clearing solution (Zhu et al., 2019)	20% (wt/vol) urea, 30% (wt/vol) D-sorbitol, 5% (wt/vol) glycerol dissolved in DMSO
Fish Water	80% distilled water and 20% tap water (charcoal filtered) mix with conductivity(150 µS/cm), complemented by the addition of 5g/100L Sera Mineral Salt, trace elements using 100mg/L TraceVit supplement (Rebie) and FerroVit fluid (Rebie)
Fin Clip Buffer	400 mM Tris-HCl pH 8.0, 5 mM EDTA pH 8.0, 150 mM NaCl, 0.1 % SDS
E3 medium (Fleisch et al., 2008)	5 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$, 0.33 mM $MgSO_4$, 0.1% methylene blue
Blocking solution	4% sheep serum, 1% BSA and 1% DMSO in PTW
Bleaching solution	3% H_2O_2, 0.5% KOH in PTW
10x TAE	48.4 g/l Tris base, 11.42 ml/L Glacial acetic acid, 0.01M EDTA, adjusted to pH 8.5
1x Hatch medium	2 mg/L Methylene blue in 1x ERM

1x ERM (Embryo Rearing Medium)	17 mM NaCl, 40 mM KCl, 0.27 mM $CaCl_2$, 0.66 mM $MgSO_4$, 17 mM HEPES
16% PFA (paraformaldehyde)	160 g/l PFA, adjusted to pH 7.0
20x tricaine	4 g/L tricaine, 10 g/L $Na_2HPO_4 \cdot H_2O$ in 1x ERM, pH 7-7.5

5.1.4 Antibodies

Table 5.4 Primary antibodies used in this study.

Primary antibody	Species	Dilution	Company
anti-BrdU	rat	1:200	Abcam, ab6326
anti-Zpr1	mouse	1:200	Zebrafish International Resource Center (ZIRC)

Table 5.5 Secondary antibodies used in this study.

Secondary antibody	Species	Dilution	Company
anti-mouse-Alexa 488	goat	1:750	Life Technologies A-11029
anti-rat-Alexa 633	donkey	1:750	Invitrogen A21094

5.1.5 Consumables

Table 5.6 Consumables used in this study.

Consumable	Source
96-well plate, 2.2 ml Deepwell AB-0932	Thermo Fisher Scientific
Aluminium foil lids	Beckman Coulter
25 ml Beaker	DURAN
Cover slips	Roth
Filter tips 10 µl, 20 µl, 200 µl, 1.25 ml	Starlab
Glass needles GC100F-10	Harvard Apparatus
Glass petri dishes STERIPLAN 4 cm, 9 cm	Roth
Gloves (latex, nitrile)	Semperguard, Starlab
Injury chamber	Homemade with Polyoxymethylene (PMMA)
MatTek dish	Mattek
Parafilm	Amcor

PCR tubes	Kisker
Petri dishes 35 mm for OMR	Sarstedt
Petri dishes 35 mm, 60 mm	Greiner
Petri dishes 92 mm	Sarstedt
Phase Lock Gel – tubes light 2 ml (PLG)	QuantaBio
Pipette tips	Kisker
Plastic pipettes 2 ml, 5 ml, 10 ml, 25 ml	Sarstedt
Safe-Lock tubes 1.5 ml, 2 ml	Eppendorf
Scissor	WPI
Serological pipettes 2 ml, 5 ml, 10 ml, 25 ml	Sarstedt
Slide glasses	epredia
Transfer pipets 3ml, 5ml	Thermofisher scientific
Tubes 15 ml, 50 ml	Greiner
Videoprinterpaper, Mitsubishi K61B	medpaper

5.1.6 Chemicals, Kit, and Reagents

Table 5.7 Chemicals, kit, and reagents used in this study.

Chemical or Reagent	Source
150 mM Tris-HCl	Sigma-Aldrich
2-propanol	Honeywell
5x Q5 PCR-Buffer	New England Biolabs
6x DNA Loading Dye	New England Biolabs
Acetone	Fisher Chemical
Agarose	VWR Avantor
Borax anhydrous	Fluke
Bovine Serum Albumin	Sigma Aldrich
BrdU(5-Bromo-2-deoxyuridine)	Sigma Aldrich
Calcium chloride ($CaCl_2$)	Merck
D-sorbitol	Sigma Aldrich

Chemical or Reagent	Source
DAPI (4',6-Diamidino-2-Phenyindole, Dilactate; 10 µg/ml in DMSO stock)	Roth
DMSO (Dimethyl sulfoxide)	Roth
dNTPs 10 mM	Sigma-Aldrich
Ethanol 99%	Fisher Chemical
Ethidium bromide (EtBr)	Roth
Ethylenediamine tetraacetic acid (EDTA)	AppliChem
Gene Ruler DNA ladder mix	Thermo Fisher Scientific
Glycerol	Merck
HEPES (*N*-2-Hydroxyethylpiperazin-*N*'-2-ethansulfonsäure)	Roth
Hydrochloric acid (HCl)	Merck
Hydrogen peroxide (H_2O_2)	ITW Reagents
In Situ Cell Death Detection Kit, TMR red	Sigma Aldrich
Magnesium chloride ($MgCl_2$)	AppliChem
Magnesium sulfate ($MgSO_4$)	Roth
Magnesium sulfate heptahydrate ($MgSO_4 \cdot 7\ H_2O$)	Roth
Methylene blue trihydrate	Sigma-Aldrich
Normal goat serum (NGS)	Sigma-Aldrich
PBS, pH 7.4	Thermo Fisher Scientific
Paraformaldehyde (PFA)	Sigma Aldrich
Phenol	Roth
Phenol-Chloroform-Isoamylalcohol (25:24:1) pH 8 (PCI)	Roth
Potassium chloride (KCl)	Merck
Potassium hydroxide (KOH)	Merck
Potassium hydroxide (KOH)	Merck
Proteinase K	Sigma Aldrich
Q5 High-Fidelity DNA polymerase (2U/µl)	New England Biolabs

Chemical or Reagent	Source
Sheep serum	Sigma Aldrich
Sodium chloride (NaCl)	Sigma Aldrich
Sodium Citrate dihydrate	Sigma Aldrich
Sodium dodecyl sulfate (SDS), 20%	Roth
Sodium pyruvate (100 mM)	Sigma Aldrich
Sucrose	ROTH
Tissue freezing medium	Leica
Tri-sodium citrate dihydrate ($C6H5Na_4O_7 \cdot 2\ H_2O$)	Sigma Aldrich
Tricaine ($C_9H_{11}NO_2 \cdot CH_4SO_3$, MS-222)	Sigma Aldrich
Tris-hydrochloride (Tris-HCl)	Sigma Aldrich
TritonX-100	Sigma Aldrich
Tween20	Sigma Aldrich
Urea	Sigma Aldrich

5.1.7 Equipment, Instruments, Microscopes and Server Resources

Table 5.8: Equipments and instruments used in this study.

Name	Supplier
CellCam Centro 200MR	CAIRN EUROPE
Centrifuge 5425	Eppendorf
Centrifuge 5430	Eppendorf
Centrifuge 5430 R	Eppendorf
Centrifuge 5810 R	Eppendorf
Cold light source for stereomicroscope KL 1500 LCD	Schott
Fish incubator	Heraeus instruments
Fish incubator	RuMed
Forceps 5, 55 Inox stainless steel	Dumont
Freezer -20°C	Liebherr

Freezer -80°C	Thermo Scientific
Fridge 4°C	Liebherr
Gel chamber	peqLab and custom-made
GM12HR41216MCN lens	Goyo Optical
Incubator 37°C, 60°C	BINDER
Infrared light filter	FI5830-55, Heliopan Filter 5830 | Ø 55 mm, Infrared Filter RG 830 (87C) 830 nm
Infrared light strips	SECURITY LINE-LED Flex Strip infrared IR 12V IP65 940nm
Leica EL6000 (Lamp HXP-R120W/45C VIS, power input 120 W, Osram Licht AG)	Leica
Leica TCS Sp8 (confocal laser scanning microscopy)	Leica
Linear-pool-arena	Home made (PLEXIGLAS® Satinice weiss (snow) WH10 DC)
Micropipette puller (Flaming/Brown)	Sutter Instruments, Model: P-97
Microwave	Samsung
Milli-Q water filtration station	Millipore Corporation
Monitor	Gigabyte M28U 71,1 cm (28") 4K Ultra HD LED Huawei
Multichannel pipette	Eppendorf
Multipette plus	Eppendorf
Needle puller P-30	Sutter Instrument Co USA
Nikon SMZ18 with a Digital Sight DS-Ri1 camera (stereomicroscopy)	Nikon
Olympus MVX10 (epifluorescence stereomicroscopy)	Olympus
Olympus SZX7	Olympus
pH-meter	Sartorius
Peristaltic pump	Home made
Pipetboy acu	Integra biosciences
Pipetman 10 µl pipette ErgoOne	GILSON
Pipetman 1000 µl pipette	GILSON
Pipetman 20 µl pipette	GILSON
Pipetman 200 µl pipette	GILSON
Power supply PowerPac Basic	Bio RAD
PowerPac Basic	Bio RAD

Prism	Homemade with polymethyl methacrylate
Rocking shaker DRS-12	neoLab
Scale Entris	Sartorius
Scale Extend	Sartorius
Spectrophotometer DS-11+	DeNovix
Synology RS4017xs	Synology
Synology RX3614xs+	Synology
Thermocycler C1000 Touch	Bio-Rad
Thermomixer Compact	Eppendorf
Thermomixer Compact/ F1.5	Eppendorf
Thermomixer F1.5	Eppendorf
Tube revolver	Thermo Fisher Scientific
UV table	Vilber Lourmat
UV-Gel Documentation System	Intas
Vortex	Scientific Industries

5.1.8 Software

Table 5.9: Software used in this study.

Software	Reference
Affinity Designer	Serif (Europe) Ltd.
Ensemble	(Howe et al., 2021; Yates et al., 2020)
Fiji	(Schindelin et al., 2012)
Geneious	Biomatters Limited
LasX	Leica
MATLAB	MATLAB
µManager	Micro-Manager-2.0.0-gamma1-20210214 (Edelstein et al., 2014)
Microsoft Excel	Microsoft
Microsoft Word	Microsoft
Python 3.9.18	Python
Prism9	Prism

RStudio	RStudio, PBC (Team R, 2020)

OpenAI's ChatGPT 4 and perplexity were used for some of the scripts generation for the analysis, proofreading and suggestions for rephrasing and literature search. DeepL was used to translate the abstract in German.

5.2 Methods

5.2.1 Fish Husbandry and Embryo Culturing

All medaka (*Oryzias latipes* and *Oryzias sakaizumi*) and zebrafish (Danio rerio) were maintained as closed stocks in a constant recirculating system under a 14-hour light/10-hour dark cycle at Heidelberg University and Karlsruhe Institute of Technology. Medaka husbandry (permit number 35–9185.64/BH Wittbrodt) was performed according to local animal welfare standards (Tierschutzgesetz §11, Abs. 1, Nr. 1) and in accordance with European Union animal welfare guidelines (Bert et al., 2016).
Embryos were kept either in 60mm petri dishes (max. 25 embryos) or 90mm dishes (max. 50 embryos) in 1x Hatch medium at 28°C. Medium was changed every Monday, Wednesday, and Friday.

For assessment of development of visual function through 1 to 10 dph using infinity pool arena, each hatchling was kept in 3.5 cm petri dish in 1x ERM, medium was changed at 1, 4, 7, 10 dph, and fed with dry food once a day.

5.2.2 Linear-Pool-Style OMR Assay

The OMR assays were performed as described in (Suzuki et al., 2024). In short, after 2 hours of acclimation in the linear-pool-arena with stationary stripes, 0-1 dph medaka were exposed to the stripe motion. One set of stripe motion had 4 phases. First, the stripes remained stationary for 5 seconds. Second, the stripes moved to the right for 10 seconds. Third, the stripes paused for another 5 seconds. Finally, the stripes moved to the left for 10 seconds. Before being exposed to the other stripe motion conditions, the hatchlings had 10 minute acclimation period with the corresponding stationary stripe pattern. In this thesis, "width" always refers to the combined width of a black and a white stripe. The stripe motions were generated using a JavaScript-based programme, called Fish Stripes

(https://junshern.github.io/fish-stripes/). Tracking and analysis were done using Fiji (Schindelin et al., 2012) as described in (Suzuki et al., 2024).

5.2.3 Setup of the Infinity-Pool-Style OMR Assay

The OMR setup is illustrated in Figure 2.5. The setup is made from extruded aluminum as pillars (width/length = 30 x 30 (mm)), all of which were fixed using electro-galvanized and passivated steel nuts and bolts, with a black curtain surrounding them to shield from the light. The display (Gigabyte M28U 71,1 cm (28") 4K Ultra HD LED | Huawei) was disassembled to position the electronic part not directly attached to the screen to prevent it from heating. A black cylinder (diameter/height = 12/10 (mm)) was glued (LOCTITE® 460) to the center of a 3.5 mm polystyrene dish (SARSTEDT 82.1135.500) using a mold. 50 dishes and prisms were placed on the reassembled monitor. Each dish was placed inside a prism (outer diameter/inner diameter = 56/35,7 (mm)) made from polymethyl methacrylate. Two cameras (CellCam Centro 200MR) containing a Goyo Optical GM12HR41216MCN lens were mounted above the monitor for video recording (in 1.8 m distance). The cameras were operated using µManager software (Micro-Manager-2.0.0-gamma1-20210214). All the stripes were generated using the Rotating Radial Stripes Software. For these experiments, where acclimation time and visual function development (performing OMR at 1, 4, 7, 10 dph) were tested, OMR setup with a single camera with 30 dishes/prisms was used.

Table 5.10 Parameters used on Rotating Radial Stripes Software.

Value	Parameter
92	Radius
64	Inner/Outer ratio (%)
7	Number of stripes
0	Rotationspeed (rad/s)
#FFFFFF	Stripe color 1
#000000	Stripe color 2
#B8B8B8	Inner circle color
50	Number of circles
5	Space between circles

Embryos in dishes were observed daily between 15:00 and 20:00. Those hatched were transferred to fresh dishes containing 1× ERM and kept in the incubator for OMR exposure the following day for those tested at 1 dph. On the day of the experiment, all hatchlings were examined under a binocular microscope for swimming ability and obvious morphological defects (e.g. immobility, no swimming respnse upon touch and abnormally large yolks). Hatchlings with abnormal behaviour or appearece. Hatchlings that failed to swim or showed clear abnormalities were excluded from the experiments or used with notes written on the metadata sheet. 5 ml 1xERM was added for each 3.5 cm petri dish. All OMR experiments with MIKK strains were initiated between 08:00 and 16:00 to minimize potential effects of circadian rhythms. Some OMR experiments using traditional inbred strains were initiated between 17:00 and 18:00. These later experiments were relatively few and distributed randomly across groups, and no major effect of experimental start time on the results was observed.

The standard stripe motion protocol with varying stripe widths was as follows: after a 5 minute acclimation period (consisting of 1 minute with a black background followed by 4 minutes with the first stationary stripe pattern), hatchlings were exposed to moving stripe stimuli with assay parameters varying in stripe width, stripe motion direction and speed. Each stripe motion set included: 2.5 minutes of clockwise (CW) motion, 30 seconds pause, 2.5 minutes of counter-clockwise (CCW) motion, and 30 seconds pause. This sequence was repeated with progressively thicker stripe widths (Figure 2.7 A). The stripe widths (in mm) used were: 1.2, 1.6, 2.0, 2.4, 2.8, 3.2, 3.7, 4.0, 4.4, 4.7, 5.2, 5.7, 6.0, 6.3, 6.7, 7.1, 7.6, 8.1, 8.1, 16.2, 16.2. Black and white stripes moved at a speed of 20.6°/s, except for the second round of 8.1 mm and 16.2 mm stripes, which moved at 61.8°/s. To assess differential acclimation time, some hatchlings were placed into the dishes at 0 minutes (immediately before starting the experiment), 25 minutes, or 55 minutes before the beginning of the stripe motion sequence. As a result, the total acclimation times for the hatchlings in each group were 5 minutes, 30 minutes, and 1 hour, respectively.

The stripe motion protocol with varying color was as follows: after 5 minutes of acclimation, hatchlings were subjected to the stripe motion. Each set of stripe motion consisted of 1.5 minutes of clockwise (CW) motion, 30 seconds of pause, 1.5 minutes of counter-clockwise (CCW) motion, and 30 seconds of pause, which was followed by further sets of stripe motions with thicker stripes. In each experiment, stripe color was changed as black/white

(started as a gray/gray pattern with one stripe progressively darkening and the other progressively lightening in contrast, in total 7 different contrast), then pause for 3 min, black/blue (with 4 shades of blue), then pause for 3 min, black/red(with 4 shades of red), then pause for 3 min, black/green (with 4 shades of green), then pause for 10 min, colored-stripes (8 color combinations known to be difficult to distinguish for people with color blindness). The speed of stripes was 20.6°/s, and the thickness was 8.1 mm.

In this thesis, "width" always refers to the combined width of a black and a white stripe. All experiments were carried out at 21-22 degrees.

5.2.4 Analysis for Infinity-Pool-Style OMR Assay

The acquired TIFF stack was separated into two TIFF stacks using Python script, as one stack with images from the left side camera and the other stack with images from the right side camera. Automated ROI or manual ROI was set for each stack. For setting ROI manually, a macro on FIJI was used to set the center of the circular ROI and measure XY coordinates, then Python script was used to generate a config.json file for using python-based fish detection software developed in this study. The hatchlings' swimming behavior was tracked using fish detection software. In short, regions of interest (ROIs) for each dish are defined either manually or automatically. This is followed by background subtraction and image processing, including object opening and dilation. Objects exceeding a defined threshold are identified and classified as hatchlings. For each frame, the software extracts XY coordinates and the angular position of the detected object in each dish, which can be extracted as CSV files. R script was used to further calculation of all the behavioral phenotypes.

Time stamp extraction: For each experiment, one of the two stacks (one video stack from the camera on the right side and the other video stack from the left side) was used to generate a csv file containing the mean gray value differences between frames. Image analysis was performed using a custom macro script in Fiji. For each video stack, the macro duplicated all slices from frame 2 to the last frame ("Stack-1") and from frame 1 to the penultimate frame ("Stack-2"). These two stacks were then subtracted from each other using the Image Calculator function to create a differential stack. The mean gray value for each slice of the resulting stack was measured. The resulting measurements were compiled into a single results table and saved as a CSV file in the specified output directory.

Actual time calculation: Calculation of the frames where stripe motion started/stopped was extracted from CSV files using a custom R script. Initially, individual .csv files containing mean gray values for each TIFF stack were loaded and combined into a single dataset. In short, the difference in mean gray values between consecutive frames was calculated to detect periods when stripe motion was stationary. Frames with mean gray values below 0.8 were considered to represent stationary stripes. To exclude noise, only periods with at least 25 consecutive stationary frames were retained for further analysis. A "Diff_frame" column was created by calculating the difference between consecutive frame numbers in the filtered dataset. This column quantified the number of frames separating one stationary period from the next. Large "Diff_frame" values (greater than 1500 frames) indicated points where the stripe started or stopped moving, reflecting a transition between motion and stationary states. These large gaps were used to precisely locate the frames immediately before and after the stripe's motion events. The resulting reduced dataset was labeled with "Start_f" and "End_f" timepoints, reshaped into a wide format, and a new column ("Actual_frame") was computed to quantify the number of frames within each stripe motion period. A "Duration" column was manually added based on a predefined experimental design. E.g. for standard stripe motion with varying stripe width, the duration was set as the first stationary period lasted 300 seconds, followed by alternating 150 seconds and 30 seconds stationary periods. This final annotated dataset was saved as a CSV file. For each experiment, the actual time was calculated for each frame as follows. A new dataframe was created containing all frame numbers from 1 to the maximum camera frame count, which was retrieved from an experimental metadata table by matching the experiment ID. This frame number dataframe was then merged with the timestamp data based on matching the start frame (Start_f) column. Missing values resulting from the join were filled using the "last observation carried forward" (LOCF) method to propagate the most recent non-missing values for the columns End_f, Actual_frame, and Duration. Subsequently, the script calculated the time in seconds per frame by dividing the Duration of each stripe motion by its corresponding actual frame count. Finally, a cumulative time column "Time" was generated by summing the frame-wise seconds to reconstruct the experimental timeline accurately.

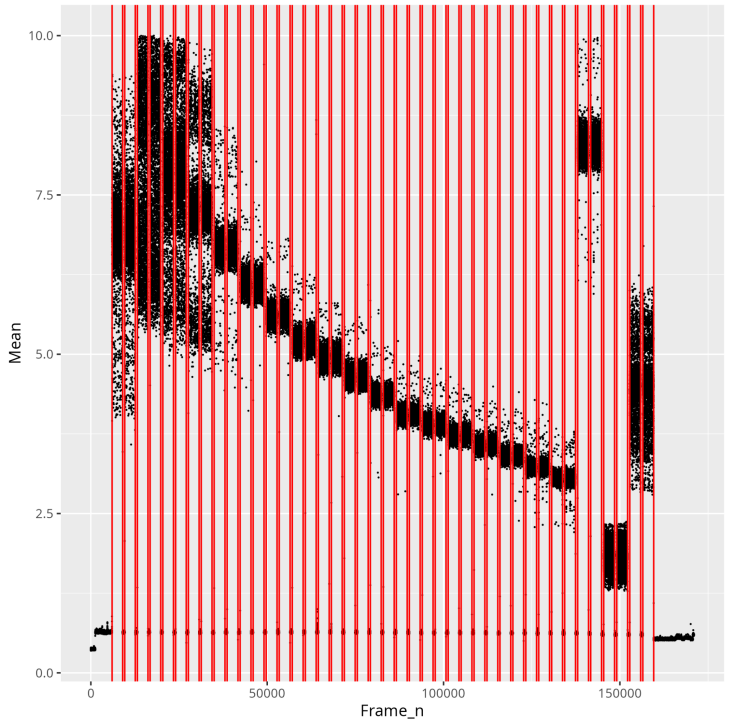


Figure 5.1 Detection of Stripe Motion Transitions Based on Mean Gray Value Differences.

An example plot showing mean gray value differences and the timing of frames where stripe motion started or stopped for one experiment. The X-axis represents the frame number, and the Y-axis represents the mean gray value differences between frames. Red lines indicate frames that were later identified as points where stripe motion changed, either starting or stopping.

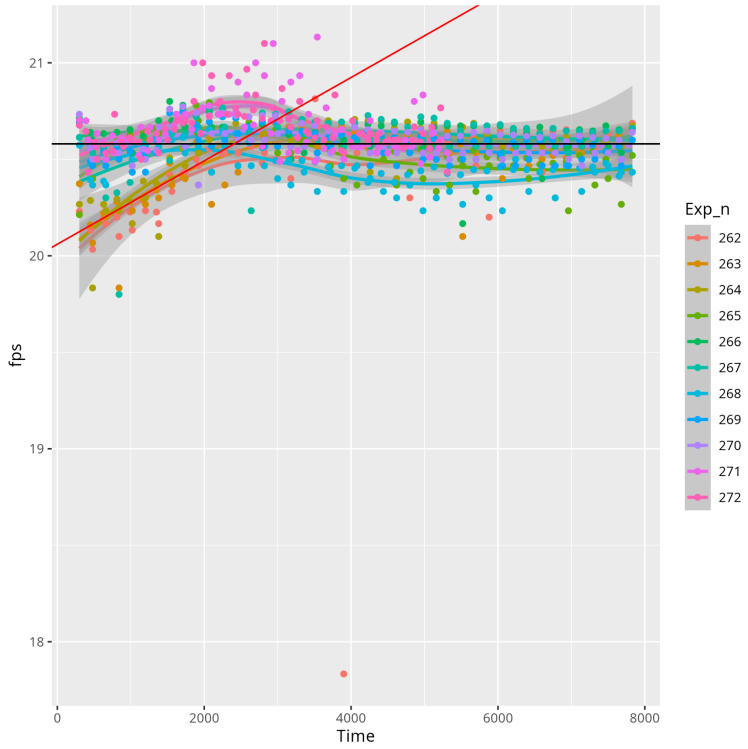


Figure 5.2 Analysis of Frames Per Second (fps) Over Time Across Multiple Experiments.

Each colored line and dot represents an individual experiment (Exp_n), showing the relationship between cumulative experimental time in seconds (X-axis) and calculated frames per second (Y-axis). Data points were generated by dividing the number of frames during each stripe motion (Actual_frame) by its corresponding duration (Duration). A smoothed trend line was fitted for each experiment. A horizontal dashed line at 20.58 fps indicates the expected frame rate, and a red regression line (slope = 2.16×10^{-4}, intercept = 20.06) represents the overall trend at the beginning of image acquisition across experiments.

Response value calculation: Response value calculations were performed using a custom R script. For each position, angular differences were calculated as follows. First, the relevant positional column and the corresponding time data were selected from the original dataset, and missing values (NAs) were removed. For each position, frame-to-frame angular differences were calculated, producing a new dataframe containing the angular difference (angle_dif) and the associated timestamp (Time).

To correct for cases where fish swam across the 0°/360° boundary, angular differences exceeding ±300° were adjusted. Values greater than 300° were replaced by 360 - angle_dif and values less than -300° by -360 - angle_dif. This correction approach was later recognized to be mathematically incorrect; the proper adjustment should subtract 360° from angle_dif when differences greater than 300°, and add 360° to angle_dif differences less than -300°. However, because the dataset was collected at 20 frames per second and real angular changes between frames were typically small (< 0.5 - 1°), this error was rare and expected to have a minimal impact on overall results and interpretation. The angular correction method will be adjusted in near future to reflect the proper mathematical approach. Given the high frame rate (20 fps) and the small angular changes per frame, the overall conclusions of this thesis remain unaffected.

A custom function fn() was implemented to calculate a response value for each time interval of 5 seconds. Within each interval, the total fish movement (fish_m) was computed as the sum of corrected angular differences. The stripe movement (stripe_m) was calculated based on stripe speed (Speed_deg.s) multiplied by the time interval duration (5 seconds). In cases where the stripe was stationary (speed = 0), the last known moving stripe speed was assigned, except during the acclimation period and after all stripe motion had concluded, where a default value of 20.6 degrees per second was used. Clockwise (CW) stripe motion was assigned a positive speed of 20.6 degrees per second, while counter-clockwise (CCW) motion was assigned a negative speed of -20.6 degrees per second. The fish response value was then

calculated as the ratio of fish movement to stripe movement, multiplied by -1 to reflect movement directionality.

Response values were computed every 5 seconds for each position (each hatchling). A positive response value indicated that the fish movement was in the same direction as the stripe motion, whereas a negative response value indicated movement in the opposite direction.

Criteria for response: A hatchling was considered responding if it showed movement following or swimming the opposite direction to the stripe pattern in both clockwise (CW) and counter-clockwise (CCW) directions. Changing the direction when stripe motion changes. Specifically, this classification required that: the count of responses (response value >= 0.5 or response value <= -0.5 (swimming the opposite direction)) in the CW direction exceeded a half (15 for stripe motion with varying stripe width, and 7 for stripe motion with color/contrast) and the count of responses in the CCW direction also exceeded the threshold (15 for stripe motion with varying stripe width, and 7 for stripe motion with color/contrast).

Minimum stripe width triggering response: The minimum stripe width that each hatchling responded was taken for plotting.

Blind: Hatchling was categorized as blind-like based on whether it failed to sustain swimming for a continuous period and did not show above threshold total absolute response values. For each hatchling and each motion phase (CW or CCW), the response duration was defined as the maximum number of consecutive 5-second blocks in which the response value was either $\leq$ -0.5 (indicating sustained responses in the opposite direction toward stripe motion) or $\geq$ 0.5, representing sustained responses within that phase. The highest value across phases (all the stripe motions with varying stripe width) was set as the **"longest response duration"** for each hatchling. **"Total response values"** (cumulative response values) were calculated as the sum of all absolute response value measurements per individual hatchlings throughout one experiment, reflecting overall response and activity. A hatchling was classified as blind-like if: **"longest response duration"** was less than 11 **AND "total response values"** was less than 300. If a hatchling was blind-like and did not respond to any of the stripe motion, it was considered as blind.

Calculating distance from the rim: Only these hatchlings that were classified as "Responded" were included in further calculations. For each hatchling, and for each stripe

width, the response direction was determined (either the hatchling followed the stripe motion in the same direction or swam against the stripe motion) by taking the median of response values throughout clockwise and counter-clockwise motion. The dataset with x,y position of each hatchling at each frame was used to calculate the distance from the rim. The distance from the rim was calculated for each hatchling for each stripe width as follows. These position values with response values above 0.5 (if hatchling followed stripe motion) or below -0.5 (if hatchling swam against the stripe motion) were retained. Out of these values, the top 25% (or 10 %) of the most inner position values were selected, ensuring only reliable and more distant positions where hatchlings could still distinguish the stripe and respond were considered. The median of "the top 25% (or 10%) of inner position values" was computed for each direction. The mean of these two values was set as "distance from the rim".

Visual acuity calculation: Visual acuity was calculated for each stripe width to which the hatchlings responded, and the highest resulting value was taken as the visual acuity for each hatchling. Stripe width in degrees (θ) is derived from the physical stripe width and distance from the rim as 2*(arctan(stripe width (mm) / 2* distance from the rim (mm))*180/pi), and spatial frequency (f) is defined as $1/\theta$. Visual acuity equals f unless the hatchling is classified as blind and showed no response to any stripe motion, in which case visual acuity is set to zero. Visual acuity was calculated for each stripe width to which the hatchlings responded, and the highest resulting value was taken as the visual acuity for each hatchling.

Response rate: Ratio of a total number of hatchlings that responded to any of the stripe motions to the total number of hatchlings in each group.

Speed & total distance: For each hatchling, median swimming speed for each stripe speed was calculated from x, y positional data. After omitting missing values, frame-to-frame distances were computed using the Euclidean (Pythagorean) formula, and cumulative distance travelled was obtained by summing these distances over time. Distances were converted to millimeters using a pixel-to-millimeter conversion factor which was set accordingly (OMR setup 5th floor: mm_per_pix = 0.349, OMR setup 4th floor: mm_per_pix = 0.355, OMR setup with single camera: mm_per_pix = 0.353). Post-experiment swimming speed was analyzed separately by filtering frames after 7830 seconds and calculating the median speed for each fish. Median swimming speeds were calculated for different stripe motion phases (pause, 20.6 degrees/second, and 61.8 degrees/second). The total distance traveled for each hatchling was calculated by taking the maximum cumulative distance

traveled and converting it from pixels to millimeters using a predefined pixel-to-millimeter scaling factor.

One-directional swimming: Hatchlings that exhibited swimming movement constantly in only one direction, regardless of stripe movement, were classified as “one-directional swimming”. This was determined by: a significant response (response value >= 0.5 or response value <= -0.5 (swimming the opposite direction)) count (greater than 20) in only one direction (either CW or CCW) in consecutive CW and CCW stripe motion. (e.g. A hatchling which showed 20 times “response value >= 0.5” (= swimming in CW) during stripe motion of CW and 20 times “response value <= -0.5” (= again swimming in CW) during stripe motion of CW at stripe width of 2.0 mm was considered doing one directional swimming for one time. If this hatchling showed the same response again at stripe width of 2.4 mm, this hatchling was considered to be doing one directional swimming for two times.)

Detection rate calculation: Detection rates for individual positions (dish/arena position) were calculated using a custom R script. For each position, the number of non-missing frames was determined by subtracting the count of NA values from the total expected number of frames. The expected frame count for each experiment was retrieved from a metadata table based on the experiment ID. The detection rate was then calculated as the percentage of non-missing frames relative to the total frame count.

5.2.5 Metal-Halide Light-Induced Retinal Injury

Zebrafish (6 dpf) larvae and Cab hatchlings (0 dph) were placed together in a polystyrene petri dish (3.3 cm diameter) filled with fish water, with one side covered in aluminum foil to reflect light. The light illuminant from Metal halide lamp (Leica, EL6000 External Fluorescence Light Source With Metal Halide Bulb, with bulb that has been used for more than 6 years) was placed directly on the dish surface and exposed from the side, bottom, and top of the dish for 30 minutes each. Water circulation was maintained using a peristaltic pump at a flow rate of 200 µL/min.

The water was replaced every time the light exposure direction was changed. After the light exposure, hatchlings and larvae were removed and examined under a microscope to assess heartbeat and pigmentation. Control specimens were kept on the bench. Hatchlings and larvae

were euthanized by immersion in 20× Tricaine for 5 minutes, then fixed overnight at 4°C in 4% PFA/PTW for cryosectioning sample preparation.

Medaka hatchlings (Cab, IP69-1, and IP10-1) were adapted to darkness by covering the dish with aluminum foil for three days in the 28°C incubator. Following dark adaptation, hatchlings were transferred to a homemade dish filled with 1× ERM, with one side covered in aluminum foil to reflect light. Additionally, the inside of the lid of the experimental box was lined with aluminum foil to enhance light reflection. Water circulation was maintained using a homemade pump at a flow rate of 4.6 mL/min, and fish were exposed to light for 20 minutes from the side of the homemade dish. After light exposure, hatchlings were removed and examined under a microscope to assess heartbeat and pigmentation. Control hatchlings were kept on the bench under standard conditions. Hatchlings and larvae were euthanized at the specified time point by immersion in 20× Tricaine for 5 minutes, then fixed overnight at 4°C in 4% PFA/PTW for cryosectioning sample preparation.

5.2.6 Cryosectioning

Hatchlings were fixed overnight at 4°C in 4% PFA/PTW in 1.5 ml Eppendorf tubes. After three washes with 1x PTW each for 10 minutes, the medium was replaced with 30% sucrose, and samples were incubated overnight at 4°C. The medium was then changed to a 1:1 solution of 30% sucrose in PTW and tissue-freezing medium (Leica) and incubated for more than one overnight at 4°C. Hatchlings were mounted in tissue freezing medium in a mold, with their head facing downward, ensuring the dorsal side was positioned upward and the ventral side facing the person mounting the sample. The blocks were solidified in liquid nitrogen. Cryosections were cut at 16 µm thickness, placed on slide glass, and dried overnight in the dark at 4°C.

5.2.7 Immunohistochemistry and TUNEL Staining on Cryosections

For incubation steps longer than 10 minutes, slides were covered with parafilm to avoid evaporation. Sections were rehydrated in 1× PTW for 30 minutes, followed by blocking for 1-2 hours in 10% normal goat serum (NGS) in 1× PTW. After two washes with 1× PTW (5 minutes each), primary antibody (1:500 dilution unless stated otherwise) in 1% NGS in 1x PTW was applied, and sections were incubated overnight at 4°C. Slides were washed six

times for 5 minutes each in 1× PTW. A secondary antibody (1:750 dilution) and DAPI (1:500 dilution from a 5 mg/mL stock in 1× PTW) were prepared in 1% NGS in PTW and incubated for 2 hours at 37°C. Sections were subsequently washed three times for 5 minutes with 1× PTW. TUNEL staining was performed following the Roche protocol. Sections were fixed in 4% PFA/PTW (pH 7.4) at room temperature for 20 minutes, followed by a 30-minute wash in 1x PTW. Samples were then incubated for 2 minutes at 65°C in 0.1 M sodium citrate solution. For the TUNEL reaction, a 50 µL reaction mix was prepared by combining 45 µL label solution with 5 µL enzyme solution, mixed well, and kept on ice until use. Slides were rinsed twice with PTW, and 50 µL of the reaction mix was added per section. Samples were covered with parafilm and incubated at 37°C for 60 minutes in the dark inside a staining box wrapped in aluminum foil. After incubation, sections were washed three times with PTW, mounted in 100 µL of 60% glycerol/PTW, and coverslipped for imaging.

5.2.8 Whole-Mount Immunohistochemistry

PFA fixed Hatchlings were placed in a mesh fitting the wells and subjected to bleaching. The bleaching solution consisted of 0.5% KOH and % H_2O_2 in 1x PTW. Hatchlings were gently shaken at room temperature for 2-3 hours until the tissue was completely clear. Following bleaching, samples were washed three times in PTW before retina dissection. The lens and superficial tissue were removed. Retinae were then transferred to PCR tubes, and the medium was replaced with 150 mM Tris-HCl (pH 9.0). Samples were heated to 70°C for 15 minutes, followed by three 1x PTW washes to ensure complete removal of residual solution. For permeabilization, samples were incubated in pre-cooled acetone at -20°C for 15 minutes, rinsed three times in PTW, and incubated in blocking solution (4% sheep serum, 1% BSA, 1% DMSO in 1x PTW) for at least 3 hours at room temperature. Primary antibody (diluted in blocking solution with 1:200 dilution) incubation was performed with rotation at 4°C overnight. The next day, samples were washed five times for 20 minutes each in PTW on a shaker before incubation in secondary antibody and DAPI (diluted in blocking solution with 1:750 dilution for secondary antibody and 1:500 dilution for DAPI) at 4°C overnight in the dark. Subsequent washes were performed five times for 20 minutes in PTW on a shaker in the dark.

5.2.9 Retinal Needle Injuries and BrdU Incorporation

Medaka 0 dph were anesthetized in 1x Tricaine in 1x ERM and placed on a wet tissue. Under binocular microscope, the right retina was stabbed three times in the dorsal part with a glass needle (0.05-0.1 mm diameter). Left retinae were used as controls. Just after the injury, hatchlings were incubated in 2.5 mM BrdU (4 ml/3.5 cm dish) and fed every day with BrdU solution, being changed every 2 days for 3 to 4 days at 28°C incubator. After washing with 1x ERM for three times, hatchlings were euthanized by immersion in 20× Tricaine for 5 minutes. The head was fixed in 4%PFA/PTW (in PCR tubes), and individual tail was transferred to 20 µl FinClip buffer/Proteinase K mix (20:1 ratio) for genotyping in 1.5 ml Eppendorf tubes.

5.2.10 BrdU Immunohistochemistry

BrdU staining was performed after staining of other antibodies. First, slides were post-fixed with 4 % PFA (pH 7) in 1x PTW for 30 minutes at room temperature. Then, sections were washed for 5 minutes for three times with 1x PTW. Antigen retrieval was performed by incubation of slides with 2N Tris-HCl/0.5% Triton X solution for 60 min at 37°C. Slides were washed for 10 minutes for three times with 1x PTW. Afterwards, pH recovery was performed with a saturated Borax solution diluted in 1x PTW for 15 min. Then, slides were washed for 5 min for three times with 1x PTW. Blocking was performed for 2 hours with 10 % NGS/1x PTW at room temperature, followed by 5 minutes washes for two times with 1x PTW. Primary antibody was applied in the respective dilutions in 1 % NGS/1x PTW at 4°C overnight. The next day, sections were washed 5 minutes for 6 times with 1x PTW. Respective secondary antibodies were added in 1 % NGS/1x PTW for 2 hours at 37°C. Then, slides were washed 5 minutes for three times with 1x PTW then mounted in 100 µL of 60% glycerol/PTW, and coverslipped for imaging.

5.2.11 Imaging

All immunohistochemistry imaging (cryosections and whole mount samples) was performed with the inverted confocal microscope Leica TCS SP8 (ACS APO objective lenses: 10x/0.3 dry, 20x/0.75 multi-immersion, 63x/1.3 glycerol with laser lines: 405 nm, 488 nm, 532 nm, 638 nm). For imaging whole mount samples, retinae were placed in MatTek dishes in clearing solution (Zhu et al., 2019) with the RPE touching the bottom of the dish.

5.2.12 Image Quantification on Retinal Regeneration

After imaging, the number of sections where the phenotype was observed was counted with 24-40 sections per strain and 2-10 sections per individual hatchlings were observed. Each phenotype was defined as follows: neurogenic cluster: more than 4 BrdU-positive cells next to each other, an increase in proliferative cells in ciliary marginal zone (CMZ): the area of BrdU-positive cells in the CMZ region was larger than that of the control, and BrdU-positive cells in all GCL, INL & ONL: more than one BrdU-positive cell observed across all GCL, INL and ONL.

5.2.13 Comparison of sox2 and sox3 Locus across MIKK Panel

Genomic regions corresponding to sox2 (ENSORLG00000011685) and sox3 (ENSORLG00000001780) loci were extracted from the MIKK panel and Cab lines and aligned to the HdrR reference genome. Sequence extraction was performed using a custom script provided by Fanny Defranoux (EBI).

5.2.14 Genomic DNA Extraction for PCR Genotyping

Tails in 20 µl FinClip buffer/Proteinase K mix (20:1 ratio) were incubated at 60°C for overnight. The next day, after the short centrifuge, 40 µl of H2O (double the amount of water relative to the amount of FinClip buffer/Proteinase K mix) was added and incubated for 15 min at 95°C (thermoblock). Samples were cooled down to RT, spun again shortly and stored at 4°C until using for PCR genotyping

5.2.15 Genotyping PCR

PCR with the commercial high-fidelity Q5 polymerase (NEB) was performed as shown in the following tables (Table 5.11 and Table 5.12) for 30 cycles. Desalted primers were obtained from Eurofins Genomics.

Table 5.11 PCR reaction mix.

Reagents	1x mix	Final concentration
Nuclease-free H_2O	15,8 µl	-
5x Q5 PCR-Buffer	5 µl	1x

dNTPs 10 mM	0,5 µl	200 µM
Forward primer 10 µM	1,25 µl	500 nM
Reverse primer 10 µM	1,25 µl	500 nM
DNA template	1 µl	0.1-200 ng
Q5 High-Fidelity DNA polymerase (2U/µl)	0.2 µl	0.008-0.02 U/ µl
Total	25 µl	

Table 5.12 PCR program.

Program	Time
98°C	2 min
98°C	30 s
Annealing temperature	30 s
72°C	1 min 15 s
72°C	5 min
12°C	10 min

5.2.16 DNA Agarose Gel electrophoresis

For genotyping 1% agarose gel (w/v) in 1x TAE was prepared and poured into trays. A comb was used to generate pockets for the sample volume in the agarose gel. The agarose gel was placed in an electrophoresis chamber filled with 1x TAE. The DNA sample was mixed with either 6x DNA Loading Dye and loaded into the solidified agarose gel. One lane was loaded with Gene Ruler DNA ladder mix as a band size reference. Electrophoresis was carried out at 80-130V. Following that, the DNA was stained by bathing the agarose gel in 0.2 µg/ml ethidium bromide (in 1xTAE) for 15-20 min. DNA bands were visualized and documented using a UV (λ=254 nm) transilluminator.

5.2.17 Optokinetic Response (OKR) Assay

The OKR was recorded by the experiment setup as previously described (Mueller & Neuhauss, 2010). In brief, 1 dph medaka hatchlings were exposed to binocular simulation with sinusoidal gratings. For contrast sensitivity testing, gratings with a spatial frequency of 20 cycles/360° and an angular velocity of 7.5 deg/s were used with varying contrast levels (5, 10, 20, 40, 70, and 100%). To assess spatial sensitivity, stimuli were delivered at an angular velocity of 7.5 /s and 70% of the maximum contrast was used with varying spatial frequency (7, 14, 21, 28, 42, 56 cycles/360°). Temporal sensitivity was evaluated by presenting gratings with maximum contrast and a spatial frequency of 20 cycles/360°, while progressively increasing temporal frequency (5, 10, 15, 20, 25, 30 deg/s).

5.2.18 Electroretinography (ERG)

White light ERGs were recorded on 5 dpf Zebrafish and 1dph medaka hatchlings' eyes, as previously described ((Sirisi et al., 2014)). In brief, larvae and hatchlings were dark-adapted for at least 30 min before recordings. Preparatory steps, including eye dissection, positioning the eye, and recording pipette, were done under a dim red light to prevent bleaching of photopigment. The eye was removed and placed on a filter paper on top of an agarose gel. The reference electrode was inserted into the agarose gel, and the recording electrode, a glass capillary with a tip diameter of 20-30 µm filled with E3, was placed on top of the cornea. ERG light generation and measurement were done as previously described in (Niklaus et al., 2024). A series of five white light stimuli of decreasing light intensities (log 0 to log -4 (zebrafish) or log -5 (medaka)) were presented to the eyes. Each stimulus lasted 100 milliseconds, with a 15-second interval between stimuli. ERG recordings were analyzed using Excel and MATLAB, with b-wave amplitudes used as an indicator of ON-bipolar cell depolarization. The first 50 milliseconds of each recording were averaged to establish baseline values, and b-wave amplitudes were statistically analyzed using Prism9.

5.2.19 Dissection of Adult Brain for Genomic DNA Extraction

Fish were euthanized with a hypothermic shock in an ice and water mix. Once the gill movement stopped and the fish did not respond to pressure applied to body, they were decapitated by a quick cut through the spine with scissors. Organs were dissected for genomic DNA extraction using forceps.

5.2.20 Genomic DNA Extraction for Whole Genome Sequencing of Medaka Hatchlings and Adults

Medaka hatchlings:

Genomic DNA was extracted from snap-frozen hatched embryos that were stored at -80°C. The hatchlings were incubated in 40 µl GeneCore DNA extraction buffer at 60°C overnight. The samples were briefly centrifuged and 2x volumes of H_2O were added before the genomic DNA was used for sequencing library preparation.

Medaka adults:

For high-coverage sequencing, DNA was extracted via phenol-chloroform-isoamyl extraction from whole brain samples. Brains were lysed in 250 µl FinClip buffer at 60°C for two overnights. The samples were transferred to a pre spun 2ml PLG tubes. Samples were mixed by repeated inversion with 125µl water-saturated phenol, then mixed with 125µl Phenol:Chloroform:Isoamyl (25:24:1, pH 8.0), followed by centrifugation at 12.000 x g for 5 minutes at room temperature. The aqueous phase is transferred to a fresh pre-spun 2 ml PLG tube and 250 µl Chloroform: Isoamyl Alcohol (24:1) was added and mixed by repeated inversion. The tubes were centrifuged at 12.000 x g for 5 minutes at room temperature and the aqueous phase was transferred to a new 1.5 ml safe lock eppendorf tube. 250 µl absolute Isopropanol was added and mixed vigorously by inverting the tubes several times for DNA precipitation. The precipitate was collected by centrifugation at 12.000 x g for 5 minutes and washed with 500 µl 70% ethanol by incubation for 2 minutes followed by centrifugation at 8000 rpm for 1 minute at room temperature. The washing step was repeated before the supernatant was discarded and the pellet was air dried for 2 hours. The pellet was dissolved in 50 µl 1x TE buffer.

5.2.21 Cross Breeding of MIKK strains for Segregation Analysis

For segregation analysis individuals of eight MIKK panel strains that exhibit extreme visual function and other phenotypes were used to set up 11 different crosses (Table 5.13). F2 embryos were collected from F1 mating groups (Table 5.14). F2 hatchlings at 1 dph were subjected to OMR as described above (5.2.3) HdrR hatchlings were used as experiment

reference control. F0 screening was done during June to July 2023 and F2 screening was done during August to November 2024.

For F2 screening, following the OMR assay, hatchlings were transferred to 2 ml deep-well plates. After removing excess liquid, the plates were sealed with aluminum foil and snap-frozen at −80 °C. The frozen plates were transferred to the EMBL GeneCore Facility on dry ice for genomic DNA extraction, sequencing library preparation and whole genome sequencing.

Table 5.13 MIKK panel strain cross for segregation analysis.

Cross name	Female			Male		
	Founder	Stock ID	N	Founder	Stock ID	N
IP(23-1/1x79-2)	IP23-1/1 F22	10716	1	IP79-2 F22	10645	1
IP(79-2x127-2)	IP79-2 F22	10645	1	IP127-1 F22	10661	1
IP(139-4x127-2)	IP139-4 F22	10650	1	IP127-1 F22	10661	1
IP(139-4x130-2)	IP139-4 F22	10650	1	IP130-2 F22	10663	1
IP(139-4x47-1)	IP139-4 F22	10650	1	IP47-1 F22	10627	1
IP(23-1/1x139-4)	IP23-1/1 F22	10716	1	IP139-4 F22	10650	1
IP(139-4x79-2)	IP139-4 F22	10650	1	IP79-2 F22	10645	1
IP(23-1/1x47-1)	IP23-1/1 F22	10716	1	IP47-1 F22	10627	1
IP(139-4x75-1)	IP139-4 F22	10650	1	IP75-1 F22	10644	1
IP(127-2x75-1)	IP127-1 F22	10661	1	IP75-1 F22	10644	1
IP(75-1x130-2)	IP75-1 F22	10644	1	IP130-2 F22	10663	1

Table 5.14 F1 mating group from which F2 eggs were collected.

Stock name	F1 Stock ID	Female	Male
		N	N
IP(23-1/1x79-2)	11103	8	2
IP(79-2x127-2)	11007	8	3

IP(139-4x127-2)	11989	18	4
IP(139-4x130-2)	11073	10	2
IP(139-4x47-1)	11026	11	3
IP(23-1/1x139-4)	11095	6	2
IP(139-4x79-2)	10990	7	2
IP(23-1/1x47-1)	10988	12	3
IP(139-4x75-1)	10986	12	3
IP(127-2x75-1)	10985	27	7
IP(75-1x130-2)	10984	12	3

5.2.22 Whole Genome Sequencing of F1 and F2 Medaka

The sequencing library preparation and whole-genome sequencing was done at the Genomics Core facility of the European Molecular Biology Laboratory (EMBL) in Heidelberg (Germany) as described in (Pierotti et al., 2024).

5.2.23 F2 Relatedness Matrix, WGS Data Analysis and Mapping

The analysis of the WGS data, the generation of the relatedness matrix and QTL mapping were performed by Esther Yoo (EMBL-EBI). The sequencing reads were first aligned to the medaka reference genome using the nf-core/sarek pipeline (v.3,5,1) (Hanssen et al., 2024) and then imputed by applying stitchimpute (Pierotti, Fitzgerald, et al., 2024). An in house implementation of a mixed linear model (flexlmm) was used for the relatedness matrix (Pierotti, Fitzgerald, et al., 2024). In brief, the relatedness matrix was generated using plink2 (Chang et al., 2015) as part of the flexlmm pipeline. For the relatedness matrix plot, F2 samples were clustered according to their genetic relatedness and the degree of relatedness is indicated by a color gradient.

The association testing was performed as a standard LOCO (leave one chromosome out) linear mixed model where the random effect is modelled as a function of the relatedness matrix. The phenotypes were inverse normalized and the random effects were regressed out of the phenotypes, genotypes, and fixed effect terms before estimating the fixed effects coefficients. Association testing was done for the phenotypes and the significance threshold

was calculated as one-tenth of the most significant p-value achieved genome-wide in 10 sample permutations.

5.3 Scripts for Analysis and Text for Stripe Generations

5.3.1 Stripe Generation Software and Fish Tracking Software

All scripts used for the stripe generation and fish tracking software will be publicly accessible via GitHub (https://github.com/dkalsan) following journal publication.

It includes the following custom Python scripts to analyze the TIFF stacks:

- **TIFF stack separator**: Separates TIFF stacks containing images from two cameras into two distinct TIFF stacks.
- **Fish trajectory detector**: Detects fish trajectories from the TIFF stacks and generates corresponding JSON files containing trajectory data.
- **JSON to CSV converter**: Converts generated JSON files into CSV format for easier analysis.
- **ROI overlay tester**: Generates an image overlay showing regions of interest (ROIs) on top of the original images to verify correct ROI placement.
- **Detection video generator**: Creates video files that visualize fish detections frame-by-frame for quality control and presentation purposes.

5.3.2 FIJI Macros, Python Scripts, and R Scripts

Following scripts were used for the preprocessing of acquired video stacks & metadata sheets, generation of csv file that contains time stamp and analysis.

It includes the following custom ImageJ macros and Python scripts to support preprocessing, metadata handling, detection, and file management workflows.

ImageJ Macros:

- **ROI selection macro**: Defines regions of interest (ROIs) for individual dishes in TIFF image stacks. Then **ROI_to_configjson.py** was used to convert it to json file.

—----------

```
//first replace PATH to the path where you'd like to save the file
run("ROI Manager...");
//remove "//" when running the first time to open the preset ROI
//roiManager("Open", "PATH/RoiSet.zip");
run("Set Scale...", "distance=0 known=0 unit=pixel");
```

```
        run("Set Measurements...", "centroid redirect=None decimal=0");
        roiManager("show all with labels");
        waitForUser; //move the circles to fit on the dishes
        roiManager("deselect");
        //roiManager("Select", newArray(0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,1,16,17,18,19,20,21,22,23,24));
        roiManager("Measure");
        saveAs("Results", "PATH/Results.csv");
        close("*");
        run("Close All");
—----------------------------------
```

- **Mean gray values calculation macro for extracting video frames where stripe motion start/stop**: Computes mean gray values within each frame across all frames of each video stack and make csv file for each stack.

Gray_values_only_tifs.ijm

```
—----------------------------------
//Set the directory for the videos to analyze and .csv files to save
        showMessage("Select Open Folder");
        openDir = getDirectory("Choose a Directory");
        showMessage("Select Save Folder");
        saveDir = getDirectory("Choose a Directory");
        list = getFileList(openDir);
//Open the set directory and make the list of all the file names
for (i=0; i<list.length; i++){
        openDir_2 = openDir + list[i];
        list_2 = getFileList(openDir_2);
   setBatchMode("hide");
//Open each file in the folder and perform macro on the stacks in the folder
        for (l=0; l<list_2.length; l++){
//check if the file is a .tif file
      if (endsWith(list_2[l], ".tif")) {
                //Open the .tif stacks in the folder
                open(openDir_2+list_2[l]);
                operation();
        }
  }
}
//Define operation
function operation(){
File_name = getTitle();
Save_name = replace(File_name, ".ome.tif", ".csv");
selectWindow(File_name);
numSlices = nSlices; //check how many slices there are in one stack
numSlices_2 = numSlices - 1;
duplicateRange = "duplicate range=2-" + numSlices;
duplicateRange_2 = "duplicate range=1-" + numSlices_2;
selectWindow(File_name);
run("Duplicate...", duplicateRange);
rename("Stack-1");
selectWindow(File_name);
run("Duplicate...", duplicateRange_2);
rename("Stack-2");
imageCalculator("Subtract create stack", "Stack-1","Stack-2");
run("Set Measurements...", "mean redirect=None decimal=3");
selectWindow("Result of Stack-1");
```

```
run("Select All");
for (n = 1; n <= nSlices(); n++){
setSlice(n);
run("Measure");
}
saveAs("Results", saveDir + Save_name );
selectWindow("Results");
run("Close");
close("*");
}
—----------------------------------------
```

- **Camera source identification macro**: Determines whether TIFF files originated from the left or right camera and generates a file used for standardized naming of the detection files. It uses a predefined square ROI placed on the right side of the image (where in the image from the right camera appears black). By measuring the mean pixel intensity in this ROI, the macro classifies the image source: if the region is dark (mean ≤ 20), it's from the right camera; if brighter, it's from the left. The results are saved as Camera_L.csv or Camera_R.csv in the corresponding folder.

Check_camera_position.ijm

```
—----------------------------------------
//set PATH
run("ROI Manager...");
roiManager("Open", "PATH/5.camera_check.roi");
// Set the directory for the videos to analyze
showMessage("Select Open Folder");
openDir = getDirectory("Choose a Directory");
list = getFileList(openDir);
// Open the set directory and make the list of all the file names
for (i = 0; i < list.length; i++) {
   openDir_2 = openDir + list[i];
   list_2 = getFileList(openDir_2);
   setBatchMode("hide");
   // Open each file in the folder and perform macro on the stacks in the folder
   for (l = 0; l < list_2.length; l++) {
      // Check if the file is a .tif file
      if (endsWith(list_2[l], "_1.ome.tif")) {
         // Open the .tif stacks in the folder
         open(openDir_2 + list_2[l]);
         operation(openDir_2); // Pass the current folder path to the operation function
      }
   }
}
// Define operation
function operation(currentDir) {
   File_name = getTitle();
   run("Set Measurements...", "mean redirect=None decimal=3");
   roiManager("Measure");
   // Get the Mean value from the Results table
   Mean = getResult("Mean", nResults - 1);
   // Set Save_name based on Mean value
   if (Mean > 20) {
```

```
        Save_name = "Camera_L.csv";
    } else {
        Save_name = "Camera_R.csv";
    }
    // Save the results with the determined Save_name in the same folder as the image
    saveAs("Results", currentDir + Save_name);
    // Close the Results table to avoid duplication
    close("Results");
    // Close the current image
    close("*");
}
—----------------------------------------
```

Python Scripts:

- **Fish trajectory file renamer (version 1)**: Renames trajectory CSV files with angular position using the appropriate experiment name. This Python script renames a detections.csv file based on the camera used (left or right) and an ID code extracted from the folder name. It looks for either Camera_R.csv or Camera_L.csv in the folder to determine the source camera, then extracts the third underscore-separated element (nnn) from the folder name (e.g., "123" from 20250501_RS_123). It renames detections.csv to R_nnn_detections.csv or L_nnn_detections.csv accordingly.

rename_detections_file.py

```
—----------------------------------------
import os
import sys

def rename_detections_file(path):
    target_filename = "detections.csv"
    camera_r_file = "Camera_R.csv"
    camera_l_file = "Camera_L.csv"

    # Check if the target file exists in the directory
    if target_filename not in os.listdir(path):
        print(f"Error: '{target_filename}' does not exist in the specified directory.")
        return

    # Determine the 'nnn' part of the new filename from the directory name
    directory_name = os.path.basename(path)
    try:
        nnn = directory_name.split('_')[2]
    except IndexError:
        print("Error: The directory name does not follow the expected pattern (e.g., contains '_').")
        return

    # Determine if Camera_R.csv or Camera_L.csv is present
    if camera_r_file in os.listdir(path):
        new_filename = f"R_{nnn}_detections.csv"
    elif camera_l_file in os.listdir(path):
        new_filename = f"L_{nnn}_detections.csv"
    else:
```

```
        print(f"Error: Neither '{camera_r_file}' nor '{camera_l_file}' found in the directory.")
        return

    # Perform the renaming
    try:
        os.rename(
            os.path.join(path, target_filename),
            os.path.join(path, new_filename)
        )
        print(f"File renamed to: {new_filename}")
    except Exception as e:
        print(f"Error renaming file: {e}")

if __name__ == "__main__":
    if len(sys.argv) != 2:
        print("Usage: python script.py /path/to/directory")
    else:
        rename_detections_file(sys.argv[1])
—----------------------------------
```

- **Fish trajectory file renamer (version 2)**: An alternative renaming script based on different CSV with XY position or naming conventions. This Python script renames a detections.csv file by identifying the source camera and extracting an ID from the folder name. It checks whether Camera_R.csv or Camera_L.csv is present to determine if the image came from the right or left camera. Then, it extracts the third underscore-separated element (nnn) from the folder name (e.g., "123" from 20250501_RS_123). Based on this, it renames detections.csv to R_nnn_xy_detections.csv or L_nnn_xy_detections.csv accordingly.

rename_detections_xy_file.py

```
—----------------------------------
import os
import sys

def rename_detections_file(path):
    target_filename = "detections.csv"
    camera_r_file = "Camera_R.csv"
    camera_l_file = "Camera_L.csv"

    # Check if the target file exists in the directory
    if target_filename not in os.listdir(path):
        print(f"Error: '{target_filename}' does not exist in the specified directory.")
        return

    # Determine the 'nnn' part of the new filename from the directory name
    directory_name = os.path.basename(path)
    try:
        nnn = directory_name.split('_')[2]
    except IndexError:
        print("Error: The directory name does not follow the expected pattern (e.g., contains '_').")
        return
```

```
    # Determine if Camera_R.csv or Camera_L.csv is present
    if camera_r_file in os.listdir(path):
        new_filename = f"R_{nnn}_xy_detections.csv"
    elif camera_l_file in os.listdir(path):
        new_filename = f"L_{nnn}_xy_detections.csv"
    else:
        print(f"Error: Neither '{camera_r_file}' nor '{camera_l_file}' found in the directory.")
        return

    # Perform the renaming
    try:
        os.rename(
            os.path.join(path, target_filename),
            os.path.join(path, new_filename)
        )
        print(f"File renamed to: {new_filename}")
    except Exception as e:
        print(f"Error renaming file: {e}")

if __name__ == "__main__":
    if len(sys.argv) != 2:
        print("Usage: python script.py /path/to/directory")
    else:
        rename_detections_file(sys.argv[1])
```

—----------

- **CSV to JSON converter**: Converts ROI files into config.json format used for dish detection. The generated config.jason file has to be moved to the folder where video stacks are stored then, detection python script was used.

ROI_to_configjson.py

—----------

```
import csv
import json
csv_file_path = 'PATH\Results.csv'
json_file_path = 'PATH\config.json'
data = {
    "dishes": [],
    "dish_setup": [5, 5]
}
with open(csv_file_path, 'r') as csv_file:
    reader = csv.DictReader(csv_file)
    for i, row in enumerate(reader):
        x = int(row['X'])
        y = int(row['Y'])
        data["dishes"].append([[x, y], 53])
with open(json_file_path, 'w') as json_file:
    json.dump(data, json_file)
print("CSV to JSON conversion complete.")
```

—----------

- **Metadata extractor**: Aggregates metadata from multiple Excel sheets into a single CSV. This script processes an Excel file containing multiple experiment metadata sheets. It filters sheets by name (RS_ or BW_), splits each into general and

well-specific data, merges them, and saves each as a separate .csv file. Finally, it combines all sheets into one master CSV for downstream analysis.

Metadata_extractor.py

—----------

```
#!/usr/bin/env python
# coding: utf-8

from pathlib import Path, PosixPath
import pandas as pd

def save_df_as_csv(df, file_path):
        df.to_csv(path_or_buf=file_path, sep=',', header=True, index=None, index_label=None, mode='w', decimal='.', na_rep='NaN', encoding='utf-8-sig')
   print(file_path.name + ' has been saved at: ' + str(file_path))

# Import all experiment sheets
excel_file = Path('Path_metadata.xlsx')
output_dir = Path('Path_output')

# Import all sheets from all metadata excel files
excel_sheets = pd.read_excel(io=excel_file, sheet_name=None, header=None, usecols='A:D')

# Only keep RS_ or BW_ sheets
excel_sheets = {sheet_name:sheet for (sheet_name, sheet) in excel_sheets.items() if any(experimenter in sheet_name for experimenter in ['RS_', 'BW_'])}

# Split data from excel sheet into two separate dataframes and combine
for sheet in excel_sheets:
   df = excel_sheets[sheet]

   # Split df
   split_pos = df[df[0] == '#######'].index[0]

   ''' df1 '''
   df1 = df.iloc[:split_pos, 0:2].reset_index(drop=True)

   ''' df2 '''
   df2 = df.iloc[split_pos+2:].reset_index(drop=True)
   # Set column titles
   df2.columns = df2.iloc[0]
   df2 = df2.drop(0).reset_index(drop=True)

   ''' merge df1 & df2 '''
   # Add df1 to df2: column 0 as new column headers, column 1 as values
   df2[df1[0]] = pd.DataFrame([df1[1]], index=df2.index)

   ''' Update original dict '''
   excel_sheets[sheet] = df2

# Save individual metadata sheets as .csv files
for sheet in excel_sheets:
   filename = output_dir.joinpath(f'{sheet}_metadata.csv')
   save_df_as_csv(excel_sheets[sheet], filename)

# Make combined dataframe & save
metadata_sheets_df = pd.concat(excel_sheets, axis=0).reset_index(drop=True)
filename = output_dir.joinpath('#combined_sheet_metadata.csv')
```

```
save_df_as_csv(metadata_sheets_df, filename)

metadata_sheets_df
—----------------------------------------
```

- **JSON to CSV converter**: Extract x, y coordinate (center of ROI) of each dish from config.json and make CSV file, embedding experiment identifiers. This script reads a config.json file from a specified directory, extracts dish coordinates, and generates a labeled CSV file. It tags each coordinate with an experiment ID (e.g., RS_001) extracted from the path, adds camera info based on the presence of Camera_L.csv or Camera_R.csv, and saves the result as a structured CSV for downstream analysis.

config-json_to_csv.py

```
—----------------------------------------
import json
import pandas as pd
import sys
import re
import os
import argparse

def main(input_dir):
    # Build the path to config.json
    config_file_path = os.path.join(input_dir, "config.json")

    # Check if config.json exists
    if not os.path.isfile(config_file_path):
        print(f"Error: config.json not found in the directory {input_dir}.")
        sys.exit(1)

    # Extract "RS_XXX" with numeric XXX from the file path
    match = re.search(r'RS_\d+(\.\d+)?', config_file_path)
    if not match:
        print("Error: The config file path does not contain 'RS_XXX' with numeric XXX.")
        sys.exit(1)
    output_prefix = match.group(0)

    # Load JSON data from file
    try:
        with open(config_file_path, 'r') as file:
            json_data = json.load(file)
    except FileNotFoundError:
        print(f"Error: File '{config_file_path}' not found.")
        sys.exit(1)
    except json.JSONDecodeError:
        print(f"Error: File '{config_file_path}' is not a valid JSON file.")
        sys.exit(1)

    # Extract [X, Y] and create a DataFrame with Position (0-based indexing)
    data = [{"X": coords[0], "Y": coords[1], "Position": i}
            for i, (coords, _) in enumerate(json_data["dishes"])]
    df = pd.DataFrame(data)

    # Add Exp_ID column with value "RS_XXX"
    df['Exp_ID'] = output_prefix
```

```
    # Check if Camera_L.csv or Camera_R.csv exists in the directory and add L_R column
    if os.path.isfile(os.path.join(input_dir, 'Camera_L.csv')):
        df['L_R'] = 'L'
        output_file = f"{output_prefix}_L_config.csv"
    elif os.path.isfile(os.path.join(input_dir, 'Camera_R.csv')):
        df['L_R'] = 'R'
        output_file = f"{output_prefix}_R_config.csv"
    else:
        df['L_R'] = 'Unknown'  # If neither Camera_L.csv nor Camera_R.csv exist
        output_file = f"{output_prefix}_config.csv"  # Default name if neither file exists

    # Set the output file path (specify the path where to save the output CSV)
    output_dir = 'PATH'  # Specify your desired output folder here
    output_file_path = os.path.join(output_dir, output_file)

    # Save the CSV file in the specified folder
    df.to_csv(output_file_path, index=False)
    print(f"CSV file has been created: {output_file_path}")

if __name__ == "__main__":
    parser = argparse.ArgumentParser(description="Process a JSON file and save a CSV output.")
    parser.add_argument("input_dir", help="Directory that contains the config.json file")

    args = parser.parse_args()

    main(args.input_dir)
—----------------------------------------
```

- **Automation command generator**: Produces terminal command text to automate detection and subsequent renaming steps. This script generates a text file containing Python command lines for processing multiple folders, automating steps like activating an environment, converting detections to CSV, renaming files, and exporting config data.

automate6-10.generating_texts_for_running_script.py

```
—----------------------------------------
import os

# Function to generate a single script for all paths
def generate_script(folder_paths, output_file="script_output.txt"):
    # Determine the directory where the current Python script is located
    script_directory = os.path.dirname(os.path.abspath(__file__))

    # Set the output file path to the same directory as the script
    output_file_path = os.path.join(script_directory, output_file)

    base_script = """cd cOMR
cd rrs-fish-detector
workon comr
python detections_to_csv.py {path} --fields time relative_theta --angle_unit deg --angle_positive
python PATH\\rename_detections_file.py {path}
python detections_to_csv.py {path}
python PATH\\rename_detections_xy_file.py {path}
python PATH\json_to_csv_cli.py {path}\n\n"""
```

```
    # Open the output file to write the content
    with open(output_file_path, 'w') as script_file:
        # Loop over each folder path in the list
        for folder_path in folder_paths:
            # Replace the placeholder with the actual folder path
            script_content = base_script.format(path=folder_path)
            # Write the script content to the file
            script_file.write(script_content)

    print(f"Generated single script file: {output_file}")

# Example usage
folder_paths = [
r"PATH_1",
r"PATH_2"
]

generate_script(folder_paths, "combined_script_output.txt")
—----------------------------------------
```

- **Batch execution script**: Initiates batch processing directly from the terminal for streamlined experiment handling. This script reads a text file of command-line blocks (e.g. from combined_script_output.txt), groups them into batches (default 8), and opens each batch in a new Command Prompt window for execution using subprocess.run().

run_automation.py

```
—----------------------------------------
import subprocess

def execute_script_blocks_in_batches(file_path, batch_size=8):
    try:
        with open(file_path, 'r', encoding='utf-8') as file:
            lines = file.readlines()

        script_blocks = []
        current_block = []

        for line in lines:
            if line.strip().startswith("cd cOMR"):
                if current_block:
                    script_blocks.append(current_block)
                current_block = [line.strip()]
            elif line.strip():
                current_block.append(line.strip())

        if current_block:
            script_blocks.append(current_block)

        total_blocks = len(script_blocks)
        for i in range(0, total_blocks, batch_size):
            batch = script_blocks[i:i + batch_size]
            print(f"\nExecuting batch {i // batch_size + 1} with up to {batch_size} blocks:")

            # Combine all blocks in the batch into a single command string
            batch_commands = []
```

```
            for block in batch:
                block_command = " & ".join(block)
                batch_commands.append(block_command)

            # Execute all commands in the batch in a single command prompt
            full_batch_command = " & ".join(batch_commands)
            print(f"\nExecuting combined batch command:\n{full_batch_command}")
            subprocess.run(f'start cmd /k "{full_batch_command}"', shell=True)

    except Exception as e:
        print(f"An error occurred: {e}")

if __name__ == "__main__":
    script_file_path = "PATH\\combined_script_output.txt"
    execute_script_blocks_in_batches(script_file_path)
——-------------------------------------
```

R scripts:

- **Analysis and plotting notebook**: Performs analysis and generates plots from processed detections.csv files.

```
# Load libraries

```{r library}

library('dplyr')
library('tidyverse')
library('ggplot2')
library('stringr')
library("gridExtra")
library("ggbeeswarm")
library("ggrepel")
library("zoo")
library ("plyr")
library("data.table")
library("colorspace")
library("viridis")
library("furrr") # For parallel processing to improve efficiency
library("future")

`%!in%` = Negate(`%in%`)

```

```{r}
#camera_count <- 80000

camera_count <- 170000

#camera_count <- 160800

name_selected_strain <- "manuscript_VA"
name_selected_strain <- "F0"
name_selected_strain <- "F2"
name_selected_strain <- "EyeLess"
Name <- name_selected_strain

```

F2 analysis - full
```{r setting paths- F2-roi53, eval = FALSE}

#Set path to folder with detection.csv files
Path_folder <- "/media/rsuzuki/T7_Shield/repository/cOMR_analysis/cOMR_analysis/2.Analyzing/Files_for_analysis/F2_detections_roi53_combined"

#Set path to folder with detection.csv files
Path_folder_2 <- "/media/rsuzuki/T7_Shield/repository/cOMR_analysis/cOMR_analysis/2.Analyzing/Files_for_analysis/xy_detections_roi53"

#Set path to the folder with .csv files with timing information
Path_folder_3 <- "/media/rsuzuki/T7_Shield/repository/cOMR_analysis/cOMR_analysis/2.Analyzing/Files_for_analysis/Time_stamps"

#Set path to metadata_sheet
Path_logbook_1 <-
"/media/rsuzuki/T7_Shield/repository/cOMR_analysis/cOMR_analysis/2.Analyzing/Files_for_analysis/metadata_sheet/#combined_sheet_metadata.csv"
```
```

```
#Set path to csv file for the stripe parameter
Path_parameter <- "/media/rsuzuki/T7_Shield/repository/cOMR_analysis/cOMR_analysis/2.Analyzing/Files_for_analysis/Stripe_info/p1-6_30s.csv"
Path_stripe_start <- "/media/rsuzuki/T7_Shield/repository/cOMR_analysis/cOMR_analysis/2.Analyzing/Files_for_analysis/Stripe_info/1-6_start_t.csv"
#Path_parameter_2 <- "" #parameter .csv:1-7
#Path_parameter_3 <- "" #parameter .csv:1-6*

# Set path to the config.csv
Path_config <- "/media/rsuzuki/T7_Shield/repository/cOMR_analysis/cOMR_analysis/2.Analyzing/Files_for_analysis/XY_center"

#Set path to save data
Save_path <- "/media/rsuzuki/T7_Shield/repository/cOMR_analysis/cOMR_analysis/2.Analyzing/Projects/F2_tetsuya_pc"

#Save_path <- "/media/rsuzuki/T7_Shield/repository/cOMR_analysis/cOMR_analysis/2.Analyzing/Projects/F2_screening_roi53_full"

Radius = 53

```
F2
```{r setting file names- F2, eval = FALSE}
Single_camera = FALSE
FPS = 20
mm_per_pix = 0.349
# Get the list of .csv files in Path_folder
file_list <- list.files(Path_folder)

# Extract experiment ID
Files_names_raw <- gsub("_detections.csv$", "", file_list)

Files_names <- Files_names_raw

Rev_Files_names <- sub("^(L|R)_(\\d+\\.?\\d*)$", "\\2_\\1", Files_names)

# Extract only the number from file names
Files <- unique(gsub( "(L_)|(R_)","",Files_names))

File_range <- paste("RS_", min(Files), "-", max(Files), sep = "")
tiff_n <- 62 #How many .tif stacks are there in one experiments

# Experiment to remove
Remove_exp <- character(0)

cleaned_files_names <- Files_names_raw
# cleaned_files_names <- Files_names

name_selected_strain <- "F2"

```
Recalculate time
 * Combine all the tiff-stack-name.csv for each videos
 * Extract those frame number at which stripe motion started/stopped to generate _timestamp_full.csv to add actual time information for the analysis
```{r recalculating time}

# Combine tiff-stack-name.csv files with mean gray values into one .csv file

for(l in 1:length(Files)){

tryCatch({

#Add the first tiff-stack-name.csv file to Time_csv_full
Time_dir_0 <-  paste(Path_folder_3, "/RS_", Files[l],".csv", sep = "")
Time_csv_full <- read.csv(file = Time_dir_0)
#Add one row that was deleted for image calculation by duplicating the last row
Time_csv_full[nrow(Time_csv_full)+1,] <- Time_csv_full[nrow(Time_csv_full),]

Time_csv_full$video <- "0"

for(i in 1:tiff_n){

Time_dir <- paste(Path_folder_3, "/RS_", Files[l], "_", i, ".csv", sep = "")

# Load .csv file
Time_csv <- read.csv(file = Time_dir)
#Add one row that was deleted for image calculation by duplicating the last row
Time_csv[nrow(Time_csv)+1,] <- Time_csv[nrow(Time_csv),]

Time_csv$video <- i

Time_csv_full <- rbind(Time_csv_full, Time_csv)

}


# Add proper frame number
Time_csv_full <- Time_csv_full %>%
  dplyr::mutate(Frame_n = row_number())
```

```
# Add column "Difference" calculating difference between rows
Time_csv_full$Difference <- c(NA, diff(Time_csv_full$Mean))

Path_save_time_csv <- paste(Path_folder_3, "/", Files[l],"_timestamp_all.csv", sep = "")
write.csv(Time_csv_full, Path_save_time_csv, row.names=TRUE)

# Make a plot for the mean gray value difference between slices through the .tiff stacks
# To see what is the average or max mean gray value difference where stripe is stationary
p <- Time_csv_full %>%
  filter(Mean < 10) %>%
  ggplot(aes(x = Frame_n, y = Mean)) +
  geom_point(size = 0.01)

# Filter those rows where mean gray value is less than 0.8 meaning stripes was stationary
Filtered <- Time_csv_full %>%
  filter(Mean < 0.8)

# Remove when stripes seemed stationary for less than 25 frames

 # Create a helper column to identify sequences
 Filtered  <- Filtered  %>% mutate(diff = c(0, diff(Frame_n)), sequence_id = cumsum(diff != 1))
 # Count the length of each sequence
 sequence_lengths <- Filtered  %>% dplyr::group_by(sequence_id) %>% dplyr::summarize(length = n())
 # Filter out sequences with length less than 25
 long_sequences <- sequence_lengths %>% filter(length >= 25)
 # Keep only rows from long sequences
 Filtered <- Filtered  %>%
   filter(sequence_id %in% long_sequences$sequence_id) %>%
   select(-diff, -sequence_id)

# Make a plot for the frames at which stripe were stationary
# To see and decide what threshold to set for extracting the frame information at which stripes were stationary
r <- Filtered %>%
  ggplot(aes(x = Frame_n, y = Mean)) +
  geom_point(size = 0.1)

# Add column "Diff_frame" in Filtered calculating the frame differences between rows
# To find out at which frames the stripe motion was started/stopped
Filtered$Diff_frame <- c(NA, diff(Filtered$Frame_n))

# Make a plot for the difference in frame numbers in Filtered df

s <- Filtered %>%
  ggplot(aes(x = Frame_n, y = Diff_frame)) +
  geom_point(size = 0.1)


# Make a dataframe that includes only those rows where stripe motion started/stopped

Filtered_3000_dif <- Filtered %>%
  filter(Diff_frame > 1500)


Filtered  <- Filtered %>%
  # Exclude those cases where there was stationary stripe during stripe motion
  filter(Diff_frame < 20 | Diff_frame > 1000) %>%
  # Filter those rows where stripe motion started/stopped and the row one next to it
  filter(Frame_n %in% c(Filtered_3000_dif$Frame_n)|
         lead(Frame_n) %in% Filtered_3000_dif$Frame_n)

# Filter to keep only the first frame, and at which stripe motion started, and the one next frame
Time_csv_full <- Time_csv_full %>%
  filter(Frame_n %in% c(1, Filtered$Frame_n+1, Filtered$Frame_n + 2))

# Remove the last row of Time_csv_full
Time_csv_full <- Time_csv_full[-nrow(Time_csv_full),]

# Add column with "Start_f" and "End_f"....
Time_point <- rep(c("Start_f", "End_f"), 42)
Time_csv_full$Time_point <- Time_point

# Add column "Start_f" and "End_f"
Time_csv_full <- Time_csv_full %>%
  select(Time_point, Frame_n) %>%
  group_by(Time_point) %>%
  dplyr::mutate(row = row_number()) %>% # Need this to make pivot_wider work
  tidyr::pivot_wider(names_from = Time_point,
              values_from = Frame_n)

# ADD column "Actual Frame" calculating the number of frames during the sripe motion
Time_csv_full <- Time_csv_full %>%
  dplyr::mutate(Actual_frame = End_f - Start_f)


# Make a plot for the number of frames in each stripe motions
t <- Time_csv_full %>%
```

```
  ggplot(aes(x = row, y = Actual_frame)) +
  geom_point(size = 0.1)

ggsave(t, path = Path_folder_3, file = paste(Files[l], "_actual_frames", ".png", sep = ""))

# ADD column "Duration"
Duration <- c(300, rep(c(150, 30), 41),150)
Time_csv_full$Duration <- Duration

Path_save_time_csv <- paste(Path_folder_3, "/", Files[l],"_timestamp_full.csv", sep = "")
write.csv(Time_csv_full, Path_save_time_csv, row.names=TRUE)

  u <- p + geom_vline(xintercept = Time_csv_full$End_f, color = "red")
ggsave(u, path = Path_folder_3, file = paste(Files[l], "_Timestamps", ".png", sep = ""))

# save the file name to Remove_exp when giving errors
}, error = function(e) {
  message("An error occurred: ", e$message)
  Remove_exp <<- paste0(c(Remove_exp, Files[l]), collapse = "|")
  print(e)

ggsave(p, path = Path_folder_3, file = paste(Files[l], "_mean_gray_value", ".png", sep = ""))
ggsave(s, path = Path_folder_3, file = paste(Files[l], "_Frame_diff", ".png", sep = ""))
ggsave(r, path = Path_folder_3, file = paste(Files[l], "_mean_gray_value_less_than_1", ".png", sep = ""))

})

}
#Check those difference in frames are less than 3000

```
Redefining the experiment files
```{r redefining the experiment files}

#Include_names <- sapply(strsplit(grep("timestamp_full.csv",list.files(Path_folder_3), value = TRUE), "_"), "[", 1)

#Remove_exp <- setdiff(Files, Include_names[-c(90:160)])

if(length(Remove_exp) != 0){

## When removing some files
# Experiment to remove
Remove_exp
# Find the indices of elements containing the defined experiment
Which_to_remove <- grep(Remove_exp, Files_names_raw)
# Remove specific experiments
cleaned_files_names <- Files_names_raw[-Which_to_remove]
# Redefine the experiments to use
Files_names <- cleaned_files_names

Rev_Files_names <- sub("^(L|R)_(\\d+\\.?\\d*)$", "\\2_\\1", Files_names)

# Extract only the number from file names
Files <- unique(gsub( "(L_)|(R_)","",Files_names))
}

```
Metadatasheet
Modify the meta datasheet
   * set class of "Exp_date, Eggs_collected and Hatched" columns as "Date"
   * add columns for "dph" and "dpf"
   * separate "Strain" column to "Strain", "Generation", "StockID" by "_"
   * filter those rows with "0-49" as values for "Position" column

```{r load metadata sheet}

#Load metadata sheet as .csv file
row_metadata <- read.csv(file = Path_logbook_1, na.strings = c("", "NA"))


# Replace "2023" with "2024" for Exp_ID == "RS_312"
row_metadata <- row_metadata %>%
  dplyr::mutate(Eggs_collected = as.character(Eggs_collected)) %>%
  dplyr::mutate(Eggs_collected = if_else(
    Exp_ID == "RS_312",
    str_replace(Eggs_collected, "2023", "2024"),
    Eggs_collected
  ))

# Replace "202411" with "202410" for Exp_ID == "RS_354"
row_metadata <- row_metadata %>%
  dplyr::mutate(Eggs_collected = if_else(
    Exp_ID == "RS_354",
    str_replace(Eggs_collected, "202411", "202410"),
    Eggs_collected
  ))
```

```
#Set the columns (Exp_date, Eggs_collected and Hatched) as class "Date"
row_metadata$Exp_date <- as.character(row_metadata$Exp_date) %>%
  as.Date(row_metadata$Exp_date, format = "%Y%m%d")
row_metadata$Eggs_collected <- as.character(row_metadata$Eggs_collected) %>%
  as.Date(row_metadata$Eggs_collected, format = "%Y%m%d")
row_metadata$Hatched <- as.character(row_metadata$Hatched) %>%
  as.Date(row_metadata$Hatched, format = "%Y%m%d")




#Calculate the dates & separate Strain information & select only positions
clean_metadata <- row_metadata %>%
 dplyr::mutate(dpf = as.numeric(difftime(row_metadata$Exp_date, row_metadata$Eggs_collected, units = "days"))) %>%
 dplyr::mutate(dph = as.numeric(difftime(row_metadata$Exp_date, row_metadata$Hatched, units = "days"))) %>%
  separate(Strain, c("Strain", "Generation", "StockID"), sep = "_") %>%
  filter(Position %in% as.character(c(0:49)))

```

```
# Calculate response value, velocity and duration of response
  * loop for each detection.csv to generate result.csv file with response values calculated
    + by running "cOMR_response_rate_ver.0.10_20FPS_20230904.Rmd"
  * loop for each result.csv file to combine all
```{r calculate response value}

Files_2 <- sub("^(L|R)_(\\d+\\.?\\d*)$", "\\2", Files_names)

#Load detection.csv file and calculate response values for each experiment
for(i in 1:length(Files_names)){

Path_detection_file <- paste(Path_folder, "/", Files_names[i], "_detections.csv", sep = "")

#Extract experiment number and position of camera (Left or Right)
Exp_n <- gsub("[RL_]", "", Files_names[i]) #Extract R, L or _ in the File_name[i]

if (Single_camera == TRUE) {
  L_R <- "single"}
else {
L_R <- str_sub(Files_names[i], 1,1)} #Extract only the first letter

File_n <- Files_2[i]

#Run rmarkdown
#rmarkdown::render("cOMR_response_rate_ver.0.10_20FPS_20240729.Rmd")
rmarkdown::render("cOMR_response_rate_ver.0.2_20FPS.Rmd")

#Save results of above script as .csv file
csv_name <- paste("results", Exp_n, metadata_csv_clean$L_R[1], sep = "_")
Path_save_csv <- paste(Save_path, "/", csv_name, ".csv", sep = "")
write.csv(final_df_m, Path_save_csv, row.names=TRUE)

#Save results of velocity as .csv file
csv_name_2 <- paste("results_speed", Exp_n, L_R, sep = "_")
Path_save_csv_2 <- paste(Save_path, "/", csv_name_2, ".csv", sep = "")
write.csv(Speed, Path_save_csv_2, row.names=TRUE)

#Save results of difference in angle for every frame as .csv file
csv_name_3 <- paste("results_Diff_angle", Exp_n, L_R, sep = "_")
Path_save_csv_3 <- paste(Save_path, "/", csv_name_3, ".csv", sep = "")
write.csv(Diff_angle, Path_save_csv_3, row.names=TRUE)

}

```


rename csv files
```{r rename csv files}
if (Single_camera == TRUE){
# Get a list of all CSV files in the folder
csv_files <- list.files(path = Save_path, pattern = "\\.csv$", full.names = TRUE)

# Loop through each file and rename it
for (file in csv_files) {
  new_name <- gsub("_single", "", basename(file))  # Remove "_single"
  new_path <- file.path(Save_path, new_name)  # Create the new file path

  # Rename the file
  file.rename(file, new_path)

  print(paste("Renamed:", basename(file), "→", new_name))
}
}
```
```

```

## combining all csv files
```{r combining all .csv file}


Files_names <- sub("^(L|R)_(\\d+\\.?\\d*)$", "\\2_\\1", Files_names)


#Combine all the results.csv files
results_full<- NULL

for(i in 1:length(Files_names)){
Path_results <- paste(Save_path, "/results_", Files_names[i], ".csv", sep = "")
result_csv <- read.csv(file = Path_results, na.strings = c("", "NA"))
results_full <- rbind(results_full, result_csv)
}

Path_save_csv <- paste(Save_path, "/results_full.csv", sep = "")
write.csv(results_full, Path_save_csv, row.names=TRUE)

# Delete the files that doesn't need any more
Path_results_files <- NULL
for(i in 1:length(Files_names)){
Path_results <- paste(Save_path, "/results_", Files_names[i], ".csv", sep = "")
Path_results_files <- c(Path_results_files , Path_results)
}

# Check if Path_save_csv exists
if (file.exists(Path_save_csv)) {
file.remove(Path_results_files)
} else {
  cat("Path_save_csv does not exist, no files will be deleted.\n")
}

#Combine all the results_speed.csv files
results_full_speed<- NULL

for(i in 1:length(Files_names)){
Path_results_2 <- paste(Save_path, "/results_speed_", Files_names[i], ".csv", sep = "")
result_csv_2 <- read.csv(file = Path_results_2, na.strings = c("", "NA"))
results_full_speed <- rbind(results_full_speed, result_csv_2)
}

Path_save_csv <- paste(Save_path, "/results_speed_full.csv", sep = "")
write.csv(results_full_speed, Path_save_csv, row.names=TRUE)

# Delete the files that doesn't need any more
Path_results_files <- NULL
for(i in 1:length(Files_names)){
Path_results <- paste(Save_path, "/results_speed_", Files_names[i], ".csv", sep = "")
Path_results_files <- c(Path_results_files , Path_results)
}

# Check if Path_save_csv exists
if (file.exists(Path_save_csv)) {
   file.remove(Path_results_files)
    cat("File deleted successfully.\n")
  } else {
    cat("Path_save_csv does not exist, no files will be deleted.\n")
  }

```


Check NA rows
```{r}

Valid_exp_pos <- clean_metadata %>%
  filter(Strain %!in% c(NA, "NaN")) %>%
  select(Position, Exp_ID)

Files_names <- cleaned_files_names

#Load detection.csv file and calculate response values for each experiment
for(i in 1:length(Files_names)){

Path_detection_file <- paste(Path_folder, "/", Files_names[i], "_detections.csv", sep = "")

#Extract experiment number and position of camera (Left or Right)
Exp_n <- gsub("[RL_]", "", Files_names[i]) #Extract R, L or _ in the File_name[i]
if (Single_camera == TRUE) {
  L_R <- "single"}
else {
L_R <- str_sub(Files_names[i], 1,1)} #Extract only the first letter
```
```

```
File_n <- Files_2[i]


#Load detection.csv file
my_csv <- read.csv(file = Path_detection_file, na.strings = c("", "NA"))

#Clean the column name only to dish number
names(my_csv) <- str_replace(names(my_csv), "dishes.", "")
names(my_csv) <- str_replace(names(my_csv), ".relative_theta", "")

# Make dataframe filtering those rows with NA
na_rows <- my_csv[!complete.cases(my_csv),] %>%
  select(!time)

na_rows <- na_rows %>%
  pivot_longer(cols = colnames(na_rows)[2:ncol(na_rows)],
               names_to = "Position",
               values_to = "angle") %>%
   dplyr::mutate(Exp_ID = paste("RS_", Exp_n , sep =""))

na_rows <- na_rows %>%
  inner_join(Valid_exp_pos)

na_rows <- na_rows[!complete.cases(na_rows),]

# Adjust position number in na_rows if the video was from camera mounted on the right side
na_rows$Position <- as.numeric(na_rows$Position)

if(L_R == "R"){
  na_rows$Position <- na_rows$Position + 25
}

# Change the Position column to character vector
na_rows$Position <- as.character(na_rows$Position)

# Function to find the longest consecutive sequence of NAs in a vector
longest_na_sequence <- function(vec) {
  consecutive_na <- rle(is.na(vec))
  max_consecutive_na <- max(consecutive_na$lengths[consecutive_na$values])
  return(max_consecutive_na)
}
# Loop through each column to find the longest consecutive sequence of NAs
longest_sequences <- sapply(my_csv[6000:nrow(my_csv),], function(col) longest_na_sequence(col))
longest_sequences <- as.data.frame(longest_sequences)
longest_sequences <- tibble::rownames_to_column(longest_sequences, var = "Position") %>%
  filter(Position %!in% c("X", "time")) %>%
  dplyr::mutate(Exp_ID = paste("RS_", Exp_n , sep =""))


# Adjust position number in longest_sequences if the video was from camera mounted on the right side
longest_sequences$Position <- as.numeric(longest_sequences$Position)
if(L_R == "R"){
  longest_sequences$Position <- longest_sequences$Position + 25
}

# Change the Position column to character vector
longest_sequences$Position <- as.character(longest_sequences$Position)


#Save results of above dataframe as .csv file
csv_name <- paste("NA_rows", Exp_n, L_R, sep = "_")
Path_save_csv <- paste(Save_path, "/", csv_name, ".csv", sep = "")
write.csv(na_rows, Path_save_csv, row.names=TRUE)

#Save results of above dataframe as .csv file
csv_name <- paste("NA_length", Exp_n, L_R, sep = "_")
Path_save_csv <- paste(Save_path, "/", csv_name, ".csv", sep = "")
write.csv(longest_sequences, Path_save_csv, row.names=TRUE)

}


if (Single_camera == TRUE){
# Get a list of all CSV files in the folder
csv_files <- list.files(path = Save_path, pattern = "\\.csv$", full.names = TRUE)

# Loop through each file and rename it
for (file in csv_files) {
  new_name <- gsub("_single", "", basename(file))  # Remove "_single"
  new_path <- file.path(Save_path, new_name)  # Create the new file path

  # Rename the file
  file.rename(file, new_path)

  print(paste("Renamed:", basename(file), "→", new_name))
}
```

```
}


Files_names <- sub("^(L|R)_(\\d+\\.?\\d*)$", "\\2_\\1", Files_names)

#Combine all the NA_rows.csv files
NA_rows_full<- NULL

for(i in 1:length(Files_names)){
Path_results <- paste(Save_path, "/NA_rows_", Files_names[i], ".csv", sep = "")
NA_rows <- read.csv(file = Path_results, na.strings = c("", "NA"))
NA_rows_full <- rbind(NA_rows_full, NA_rows)
}

Path_save_csv <- paste(Save_path, "/NA_rows_full.csv", sep = "")
write.csv(NA_rows_full, Path_save_csv, row.names=TRUE)


# Delete the files that doesn't need any more
Path_results_files <- NULL
for(i in 1:length(Files_names)){
Path_results <- paste(Save_path, "/NA_rows_", Files_names[i], ".csv", sep = "")
Path_results_files <- c(Path_results_files , Path_results)
}

# Check if Path_save_csv exists
if (file.exists(Path_save_csv)) {
   file.remove(Path_results_files)
    cat("File deleted successfully.\n")
  } else {
    cat("Path_save_csv does not exist, no files will be deleted.\n")
  }


Files_names <- sub("^(L|R)_(\\d+\\.?\\d*)$", "\\2_\\1", Files_names)

#Combine all the NA_length.csv files
NA_length_full<- NULL

for(i in 1:length(Files_names)){
Path_results <- paste(Save_path, "/NA_length_", Files_names[i], ".csv", sep = "")
NA_length <- read.csv(file = Path_results, na.strings = c("", "NA"))
NA_length_full <- rbind(NA_length_full, NA_length)
}

Path_save_csv <- paste(Save_path, "/NA_length_full.csv", sep = "")
write.csv(NA_length_full, Path_save_csv, row.names=TRUE)

# Delete the files that doesn't need any more
Path_results_files <- NULL
for(i in 1:length(Files_names)){
Path_results <- paste(Save_path, "/NA_length_", Files_names[i], ".csv", sep = "")
Path_results_files <- c(Path_results_files , Path_results)
}

# Check if Path_save_csv exists
if (file.exists(Path_save_csv)) {
   file.remove(Path_results_files)
    cat("File deleted successfully.\n")
  } else {
    cat("Path_save_csv does not exist, no files will be deleted.\n")
  }


```

Calculate xy distance
```{r Calculate xy distance}

Files_names <- cleaned_files_names
#Files_names <- Files_names_raw
Files_2 <- sub("^(L|R)_(\\d+\\.?\\d*)$", "\\2", Files_names)

#Load detection.csv file and calculate response values for each experiment
for(i in 1:length(Files_names)){

Path_detection_file_2 <- paste(Path_folder_2, "/", Files_names[i], "_xy_detections.csv", sep = "")

#Extract experiment number and position of camera (Left or Right)
Exp_n <- gsub("[RL_]", "", Files_names[i]) #Extract R, L or _ in the File_name[i]
if (Single_camera == TRUE) {
  L_R <- "single" }
if (Single_camera != TRUE) {
  L_R <- str_sub(Files_names[i], 1,1)} #Extract only the first letter
File_n <- Files_2[i]
```

```
#Run rmarkdown
 rmarkdown::render("cOMR_distance_v2.Rmd")

# Save results of above script as .csv file
csv_name <- paste("results_xy", Exp_n, L_R, sep = "_")
Path_save_csv <- paste(Save_path, "/", csv_name, ".csv", sep = "")
write.csv(Distance, Path_save_csv, row.names=TRUE)

# Save results of above script as .csv file
csv_name <- paste("results_xy_position", Exp_n, L_R, sep = "_")
Path_save_csv <- paste(Save_path, "/", csv_name, ".csv", sep = "")
write.csv(simpler_my_csv, Path_save_csv, row.names=TRUE)

}
```

rename csv files
```{r rename csv files}
if (Single_camera == TRUE){
# Get a list of all CSV files in the folder
csv_files <- list.files(path = Save_path, pattern = "\\.csv$", full.names = TRUE)

# Loop through each file and rename it
for (file in csv_files) {
  new_name <- gsub("_single", "", basename(file))  # Remove "_single"
  new_path <- file.path(Save_path, new_name)  # Create the new file path

  # Rename the file
  file.rename(file, new_path)

  print(paste("Renamed:", basename(file), "→", new_name))
}
}
```

Combine xy files
```{r combine xy file}
#Combine all the results.csv files
Files_names <- Rev_Files_names
results_xy_full <- list()

sub_results_full <- results_full %>%
 select(c("Strain" ,"Exp_ID",  "Position")) %>%
 distinct()

for(i in 1:length(Files_names)){

Path_results <- paste(Save_path, "/results_xy_", Files_names[i], ".csv", sep = "")
result_xy_csv <- read.csv(file = Path_results, na.strings = c("", "NA"))

result_xy_csv <- result_xy_csv %>%
  select("Position", "Time", "dist","Exp_ID") %>%
  left_join(sub_results_full)

results_xy_full[[i]]<- result_xy_csv
}

xy <- ldply (results_xy_full, data.frame)
Path_save_csv <- paste(Save_path, "/results_xy_full.csv", sep = "")
write.csv(xy, Path_save_csv, row.names=TRUE)

# Delete the files that doesn't need any more
Path_results_files <- NULL
for(i in 1:length(Files_names)){
Path_results <- paste(Save_path, "/results_xy_", Files_names[i], ".csv", sep = "")
Path_results_files <- c(Path_results_files , Path_results)
}

# Check if Path_save_csv exists
if (file.exists(Path_save_csv)) {
   file.remove(Path_results_files)
    cat("File deleted successfully.\n")
  } else {
    cat("Path_save_csv does not exist, no files will be deleted.\n")
  }
```

Combine xy position files
```{r}
#Combine all the results_xy_position.csv files
Files_names <- Rev_Files_names
```

```
results_xy_pos_full <- list()
config_full <- list()

sub_results_full <- results_full %>%
 select(c("Strain" ,"Exp_ID",  "Position")) %>%
 distinct()

for(i in 1:length(Files_names)){

Path_results <- paste(Save_path, "/results_xy_position_", Files_names[i], ".csv", sep = "")

result_xy_csv <- read.csv(file = Path_results, na.strings = c("", "NA"))

config_csv <- result_xy_csv %>%
  select("Position", "center_x", "center_y", "Exp_ID") %>%
  distinct()

result_xy_csv <- result_xy_csv %>%
  select("Position", "Time", "med_dist_cent","Exp_ID") %>%
  left_join(sub_results_full)

results_xy_pos_full[[i]] <- result_xy_csv
config_full[[i]] <- config_csv

}

xy <- ldply (results_xy_pos_full, data.frame)
Path_save_csv <- paste(Save_path, "/results_xy_pos_full.csv", sep = "")
write.csv(xy, Path_save_csv, row.names=TRUE)

config_full <- ldply (config_full, data.frame)
Path_save_csv <- paste(Save_path, "/config_full.csv", sep = "")
write.csv(config_full, Path_save_csv, row.names=TRUE)

# Delete the files that doesn't need any more
Path_results_files <- NULL
for(i in 1:length(Files_names)){
Path_results <- paste(Save_path, "/results_xy_position_", Files_names[i], ".csv", sep = "")
Path_results_files <- c(Path_results_files , Path_results)
}

# Check if Path_save_csv exists
if (file.exists(Path_save_csv)) {
   file.remove(Path_results_files)
    cat("File deleted successfully.\n")
  } else {
    cat("Path_save_csv does not exist, no files will be deleted.\n")
  }

#results_xy_full <- read.csv(file = Path_save_csv, na.strings = c("", "NA"))

```

Load results_full
```{r Load results_full}

# Load combined .csv file
Path_save_csv <- paste(Save_path, "/results_full.csv", sep = "")
results_full <- read.csv(file = Path_save_csv, na.strings = c("", "NA"))

Path_save_csv <- paste(Save_path, "/NA_length_full.csv", sep = "")
NA_length_full <- read.csv(Path_save_csv, row.names = NULL)[,3:5]


Path_save_csv <- paste(Save_path, "/NA_rows_full.csv", sep = "")
NA_rows_full <- read.csv(Path_save_csv, row.names = NULL)


Path_save_csv <- paste(Save_path, "/results_xy_pos_full.csv", sep = "")
results_xy_pos_full <- read.csv(Path_save_csv, row.names = NULL)



```


- Calculating Visual acuity
```{r calculation visual acuity}
rmarkdown::render("test_parameters.Rmd")
```

- **Response value calculator (cOMR_response_rate_ver.0.2_20FPS.Rmd)**: Computes behavioral response values based on experimental conditions.

````
---
title: "generate result.csv file with response values calculated"
author: "Risa"
date: "18/09/2023"
output:
  pdf_document: default
  html_document: default
---

# Adjusted to below setup if not make change in >> "Chunk"
  * 2 camera setup, >> "Combine data frame
  * for dish setup with 5x5 >> "Clean up detections.csv file" & "Calculate fish position difference in angle"
  * for videos with 20FPS >> "Clean up detections.csv file"
  * for videos with 170000 frames in total >> "Calculate detection rate"
  * Parameter.csv has 7 column >> "Calculate the response value"

# Load metadata sheet and stripe parameter sheet
```{r}
#Load metadata sheet
metadata_csv <- clean_metadata

#
######CHANGED from ver 0.8
#Load parameter ,csv file according to the stripe parameter
if (unique(metadata_csv %>%
    filter(Exp_ID %in% paste("RS", Exp_n, sep = "_")) %>%
          select(Stripe_parameter)) %in% c("1-6", "1-9", "1-9*")){
  parameter_csv <- read.csv(file = Path_parameter, na.strings = c("", "NA"))
  } else if (
    unique(metadata_csv %>%
    filter(Exp_ID %in% paste("RS", Exp_n, sep = "_")) %>%
          select(Stripe_parameter)) %in% "1-6*"){
  parameter_csv <- read.csv(file = Path_parameter_1_6_2, na.strings = c("", "NA"))
  } else if (
    unique(metadata_csv %>%
    filter(Exp_ID %in% paste("RS", Exp_n, sep = "_")) %>%
          select(Stripe_parameter)) %in% "1-7"){
  parameter_csv <- read.csv(file = Path_parameter_1_7, na.strings = c("", "NA"))
  } else if (
    unique(metadata_csv %>%
    filter(Exp_ID %in% paste("RS", Exp_n, sep = "_")) %>%
          select(Stripe_parameter)) %in% "GR1"){
  parameter_csv <- read.csv(file = Path_parameter_GR1, na.strings = c("", "NA"))
  } else if (
    unique(metadata_csv %>%
    filter(Exp_ID %in% paste("RS", Exp_n, sep = "_")) %>%
          select(Stripe_parameter)) %in% c("CV4", "CV2", "CV")){
  parameter_csv <- read.csv(file = Path_parameter_CVs, na.strings = c("", "NA"))
  } else if (
    unique(metadata_csv %>%
    filter(Exp_ID %in% paste("RS", Exp_n, sep = "_")) %>%
          select(Stripe_parameter)) %in% c("CvGr2")){
  parameter_csv <- read.csv(file = Path_parameter_CvGr2, na.strings = c("", "NA"))
  }
######

```

## Modify parameter_csv file to have rows every 5s
```{r Load stripe parameters}

parameter_csv <- as.data.frame(lapply(parameter_csv, rep, each = 6))
Start_s <- seq(0, parameter_csv$End_s[nrow(parameter_csv)] -5, by = 5)
End_s <- seq(5, parameter_csv$End_s[nrow(parameter_csv)], by = 5)
parameter_csv <- cbind(parameter_csv[,-c(4,5)], Start_s, End_s)

ncol_parameter <- ncol(parameter_csv)
```


# Recalculating correct time stamp
````

```{r recalculate the correct time stamp}

Path_save_time_csv <- paste(Path_folder_3, "/", File_n,"_timestamp_full.csv", sep = "")
# Load combined .csv file
Time_adj <- read.csv(file = Path_save_time_csv )

Frame <- data.frame(Frame = c(1:clean_metadata$Cam_count[clean_metadata$Exp_ID ==
                                          paste("RS_", Exp_n, sep ="")][1]))

Frame <- Frame %>%
  full_join(Time_adj, by = c("Frame" = "Start_f"))

# fill the NA values with the closest non-NA value.
Frame$End_f <- na.locf(Frame$End_f)
Frame$Actual_frame <- na.locf(Frame$Actual_frame)
Frame$Duration <- na.locf(Frame$Duration)

Frame <- Frame %>%
 # Add "seconds" column and add the time (s) by dividing Duration
 # (the duration of the stripe motion) / Actual_frame (the number of frames during each stripe motion)
 dplyr::mutate(seconds = Duration/Actual_frame) %>%
 # Add "Time" column summing up "seconds"
 dplyr::mutate(Time = cumsum(seconds))

#p <- Frame %>%
#  ggplot +
#  geom_point(aes(x= Frame, y = Time), size = 0.1)


```


Load detection.csv file
Clean up detections.csv file
  * for dish setup with 5x5
  * for videos with 20FPS
  * change colnames to contining only dish numbers
  * adjust time column to 20 FPS
  * add column in case columns for specific dish was missing
    + move "time" column to the rightmost
    + "time" column has to be the right most for the calculation of detection rate

```{r Clean up detections.csv file}

if (Single_camera == TRUE) {
DISH <- as.character(0:29)
} else {
DISH <- as.character(0:24)}

#Load detection.csv file
my_csv <- read.csv(file = Path_detection_file, na.strings = c("", "NA"))

#Clean the column name only to dish number
names(my_csv) <- str_replace(names(my_csv), "dishes.", "")
names(my_csv) <- str_replace(names(my_csv), ".relative_theta", "")

######CHANGED from ver 0.8
#time in detection.csv is set as 10FPS but as the video was recorded as 20 FPS
#add column with true time,

if (camera_count == 80000) {
  } else {
my_csv <- my_csv[1:camera_count,] %>%
  dplyr::select(!time) %>%
  dplyr::mutate(time = Frame$Time[1:camera_count]) #column "X" has row number from 0
######
   }

my_csv <- my_csv[1:camera_count,]

######Changed from 20230904
#Check if column for all the dish exists and if not add mock column that is missing
#For dish setup for 5x5
```

```
for (i in DISH) {
  col_name <- as.character(i)
  if (!(col_name %in% colnames(my_csv))) {
    # If the column is missing, add a mock column with the same name
    my_csv[[col_name]] <- c(rep(1, 5), rep(NA, nrow(my_csv) - 5))
  }
}

######Changed from 2023110
# Check if all of the column have more than 2 rows if not replace with mock values
for (i in DISH ) {
col_name <- as.character(i)
if (length(unique(my_csv[,i])) <= 2) {

  # If the column contains less than 2 value, add a mock column with the same name
    my_csv[[col_name]] <- c(rep(1, 5), rep(NA, nrow(my_csv) - 5))
  }
}

#Modify the .csv file by moving "time" column to the right most edge
# Identify the column index of "time"
time_col_index <- which(colnames(my_csv) == "time")

# Move the "time" column to the rightmost edge
##seq_along(df) generates a sequence of integers from 1 to the number of columns in the data frame df.
##*In other words, it provides a sequence of indices that correspond to the columns in df.
##*For example, if df has 4 columns, seq_along(df) will produce the sequence 1, 2, 3, 4.
##setdiff(seq_along(df), time_col_index):setdiff() is a function that computes the set difference between two vectors.
##*It takes two vectors as arguments and returns a vector containing all elements that are in the first vector but not in the second vector.

my_csv <- my_csv[, c(setdiff(seq_along(my_csv), time_col_index), time_col_index)]

##########


```

Calculate detection rate
  * for those videos with any frames in total
  * calculate the detection rate for each dish by counting NA in each column
  * generate dataframe with columns of Position of dish and detection rate

```{r Calculate detection rate}

#Calculate the detection rate for each dish by counting NA in each column:
##Counting na: sapply(z, function(x)sum(is.na(x)))
##sapply to only where the dish is: my_csv[, c(2:(ncol(my_csv)-1))]
##*first column is "X" and the last column is "time"
##Calculate the detection rate: 170000-sum(is.na(x)))/170000*100
Detection_rate <- as.data.frame(sapply(my_csv[, c(2:(ncol(my_csv)-1))],
                function(x)(clean_metadata$Cam_count[clean_metadata$Exp_ID ==
                        paste("RS_", Exp_n, sep ="")][1] -
                        sum(is.na(x)))/clean_metadata$Cam_count[clean_metadata$Exp_ID ==
                        paste("RS_", Exp_n, sep ="")][1]*100)) ######CHANGED from ver 0.8
#Change the row name to column
Detection_rate <- tibble::rownames_to_column(Detection_rate, "Position")
#Set the column name
colnames(Detection_rate) <- c("Position", "Det_rate")

```

Calculate fish position difference in angle
  * for dish setup with 5x5 (setting DISH as vector from 0 to 24)
  * loop for each dish/fish and make a list with data frame with difference in angular position in each row
    + omit NA
    + calculate difference in angular position in each frame saved in column "angle_dif"
    ++ (= when fish are not detected, its position is filled by estimated trajectory as difference are calculated with NA omitted rows)
    + when difference was >300 or <-300, it was estimated as fish crossed the border between 0/360
    ++ therefore the difference is calculated as (360 - angle_dif), (-360 - angle_dif) and saved in column "angl_dif_mod"

```{r Calculate fish position difference in angle}

#Extract only one dish and time, omit rows with NA
##DISH <- as.character(c(1:as.numeric(metadata_csv$Dish[metadata_csv$Exp_ID %in% paste("RS_", Exp_n, sep ="")][1]))-1)
if (Single_camera == TRUE) {
DISH <- as.character(0:29)
```

```
} else {
DISH <- as.character(0:24)} #has to be character to select as column name not index

#Make a list of "frame-wise (na.omitted) difference in fish position in angle" for each fish

##Make a data frame "my_csv_dish" with column of dish and time with NA omitted
list_1 <- list()

for (i in DISH) {
  my_csv_dish <- my_csv %>%
  select(i, time) %>%
  na.omit(my_csv_dish)

##Make a data frame "my_csv_dish_dif" with calculated difference between each row
my_csv_dish_dif <- data.frame(diff(as.matrix(my_csv_dish))) %>%
  select(!time) %>%
  mutate(Time = my_csv_dish$time[2:nrow(my_csv_dish)]) %>%
  setNames(c("angle_dif", "Time"))

#Modify those rows where it exceeds +-300
my_csv_dish_dif <- my_csv_dish_dif %>%
  mutate(angl_dif_mod = case_when(angle_dif > 300 ~ 360 - angle_dif,
                                  angle_dif < -300 ~ -360 - angle_dif,
                                  angle_dif >=-300 ~ angle_dif,
                                  angle_dif <= 300 ~ angle_dif))

list_1[[i]]<- as.data.frame(my_csv_dish_dif)

}


```

```
## Calculate the response value
  * only in these case Parameter.csv has 7 column,
    + >>if not change "parameter_csv[i,l+7]"  "as.character(DISH+8)" "cols = DISH+8"
  * use function to calculate the response value
  * add new column to parameter.csv with calculated response values every 5 seconds
  * modify parameter.csv and add column with "Response_value" & "Position" (dish position) save as "parameter_tidy_r"
```{r Calculate the response value}

#Function to calculate "the sum of the angle fish moved within the x-y time period, in dish z" divided by "the sum of the angle stripe moved within x-y time period"
##ex: list_1[[1]] = data frame of dish position "0"

fn <- function(x,y,z,a) {
  fish_m <- sum(list_1[[z]]$angl_dif_mod[which(list_1[[z]]$Time >= x &
                                 list_1[[z]]$Time <= y)], na.rm = TRUE) #Sum of the angle which fish moved
if (parameter_csv$Speed_deg.s[a] != 0){ #If stripe is in motion
  stripe_m <- parameter_csv$Speed_deg.s[a]*(y-x) #Sum of the angle which stripe moved in (y-x) seconds

} else if (a < 7 & parameter_csv$Speed_deg.s[a] == 0 ) { # 0-6th row and whenever speed is 0 and not right after the stripe motion
   stripe_m <- 20.6*(y-x) #If stripe was stationary & not right after the stripe motion, assign 20.6 deg/s as stripe speed and calculate the sum of the angle which stripe moved in (y-x) seconds
} else if (parameter_csv$Speed_deg.s[a] == 0 & parameter_csv$Speed_deg.s[a-6] != 0){
   stripe_m <- parameter_csv$Speed_deg.s[a-6]*(y-x) #If stripe was stationary, assign parameter_csv$Speed_deg.s[a-6] (= speed of previous stripe movement) as stripe speed and calculate the sum of the angle which stripe moved in (y-x) seconds
} else {
    stripe_m <- 20.6*(y-x)
  }
    value <- fish_m / stripe_m *-1 #Calculate the ratio

    return(value)}

#Calculate the response value for each rows as in parameter.csv, adding the each response value of dish from 0:(DISH-1)
DISH <- as.numeric(DISH)
for (l in DISH+1) {
  for(i in c(1:nrow(parameter_csv))) {
  parameter_csv[i,l+ncol_parameter] = fn(parameter_csv$Start_s[i], parameter_csv$End_s[i], l,i) #calculate response value of fish in dish "l" and assign in parameter_csv[i, l+ncol_parameter]
}
}


#Delete "V" in the colnames of parameter_csv and change them to dish numbers
```
```

```
names(parameter_csv) <- str_replace(names(parameter_csv), "V", "")
colnames(parameter_csv)[which(names(parameter_csv) %in% as.character(DISH+ncol_parameter+1))] <- DISH #response value

#Add new column, assigning rows to "Response_value" and colnames to  "Position"
parameter_tidy_r<- pivot_longer(parameter_csv, cols = DISH+ncol_parameter+1, names_to = "Position",
                    values_to = "Response_value", values_drop_na = FALSE)


```


Calculate the difference in angular position in every frame and transform list to dataframe
  * calculate the difference in angular position in each row for each fish and store the dataframes to list
  * change list to dataframe
  * add column with Exp_ID and L_R (left or right information)
  * adjust the position of dish depending if the video was from Left camera of Right camera
```{r Transform to dataframe}

#Extract only one dish and time, without omitting rows with NA
##DISH <- as.character(c(1:as.numeric(metadata_csv$Dish[metadata_csv$Exp_ID %in% paste("RS_", Exp_n, sep ="")][1]))-1)
if (Single_camera == TRUE) {
DISH <- as.character(0:29)
} else {
DISH <- as.character(0:24)}#has to be character to select as column name not index

#Make a list of "frame-wise difference in fish position in angle" for each fish


##Make a data frame "my_csv_dish" with column of dish and time without NA omitted
list_1 <- list()

for (i in DISH) {
  my_csv_dish <- my_csv[seq(2, nrow(my_csv), by =2),] %>% #Make it about 10FPS
  select(i, time)

##Make a data frame "my_csv_dish_dif" with calculated difference between each row
my_csv_dish_dif <- data.frame(diff(as.matrix(my_csv_dish))) %>%
  select(!time) %>%
  mutate(Time = my_csv_dish$time[2:nrow(my_csv_dish)]) %>%
  setNames(c("angle_dif", "Time"))

#Modify those rows where it exceeds +-300
my_csv_dish_dif <- my_csv_dish_dif %>%
  mutate(angl_dif_mod = case_when(angle_dif > 300 ~ 360 - angle_dif,
                                  angle_dif < -300 ~ -360 - angle_dif,
                                  angle_dif >=-300 ~ angle_dif,
                                  angle_dif <= 300 ~ angle_dif))

list_1[[i]]<- as.data.frame(my_csv_dish_dif)

}

# Transform list_1 to dataframe
Diff_angle <- ldply (list_1, data.frame)

# Add Exp_ID and L_R
Diff_angle <- Diff_angle %>%
  mutate(Exp_ID = paste("RS_", Exp_n, sep = "")) %>%
  mutate(L_R = L_R)

#Change the colname .id to Position
names(Diff_angle)[names(Diff_angle) == ".id"] <- "Position"

# Adjust position number in final_df if the video was from camera mounted on the right side
Diff_angle$Position <- as.numeric(Diff_angle$Position)
if(L_R == "R"){
  Diff_angle$Position <- Diff_angle$Position + 25
}

# Change the Position column to character vector
Diff_angle$Position <- as.character(Diff_angle$Position)
```

Calculate the angular velocity per seconds
```{r}


```

```
# Calculate the velocity
# Make a copy of Diff_angle to calculate the speed
Speed <- Diff_angle

# Round down the "Time" values to the nearest second
Speed$Time <- floor(Speed$Time)

# Aggregate by "Time" and "Position" and calculate the sum of angl_dif_mod every second
Speed <- aggregate(angl_dif_mod ~ Time + Position, data = Speed, sum)

# Add Exp_ID and L_R
Speed <- Speed %>%
  mutate(Exp_ID = paste("RS_", Exp_n, sep = "")) %>%
  mutate(L_R = L_R)

# Change col name
names(Speed)[names(Speed) == "angl_dif_mod"] <- "Fish_deg.s"

#p <- Speed %>%
#  group_by(Position) %>%
#  ggplot(aes(x = Time, y = Fish_deg.s, color = Position)) +
#  geom_point(size = 0.1, shape = ".")

```

```
## Combine "parameter_tidy_r" and "Detection_rate" to generate data frame "final_df"
  * only in 2 camera setup,
  * position of the dish is adjusted in case the video was from the camera mounted on the right side
    + >>if it was not the video from 2 camera setup, change below
    + >>>final_df$Position <- as.numeric(final_df$Position)
    + >>>if(L_R == "R"){final_df$Position <- final_df$Position + 25}
  * combine metadata sheet to response value result

``` {r Combine dataframe}
#Combine detection rate and response value
final_df <- full_join(parameter_tidy_r, Detection_rate,  by = "Position")

#Adjust position number in final_df if the video was from camera mounted on the right side
final_df$Position <- as.numeric(final_df$Position)
if(L_R == "R"){
  final_df$Position <- final_df$Position + 25
}

#Select only the information of this specific experiment from Metadata sheet
metadata_csv_clean <- metadata_csv %>%
select(Strain, Generation, StockID, Eggs_collected, Hatched, Exp_ID,
  Exp_date, Stripe_parameter, Record_start, Position, Dish, dpf, dph) %>%
  filter(Exp_ID %in% paste("RS_", Exp_n, sep =""))

#Set the class of "Position" column to character to merge
final_df$Position <- as.character(final_df$Position)
metadata_csv_clean$Position <- as.character(metadata_csv_clean$Position)

#add if it was recorded with the camera mounted on the right side or the left side
metadata_csv_clean <- metadata_csv_clean %>%
  mutate(L_R = L_R)

#Combine all
final_df_m <- full_join(final_df, metadata_csv_clean,  by = "Position")

```
```

- **Distance and position analysis (cOMR_distance_v2.Rmd)**: Calculates traveled distances, XY coordinates, and spatial metrics from fish trajectories.

``` {r setup, eval=FALSE}
# Load necessary libraries for data manipulation and parallel processing
library(dplyr)  # Provides functions for data manipulation such as select and mutate
library(tidyr)  # For reshaping data from wide to long format and vice versa
library(plyr)   # For combining lists into data frames
library(zoo)    # For handling missing values with na.locf()
library(furrr) # For parallel processing to improve efficiency
library(future)
```

```

# Recalculating Correct Timestamp
## This step adjusts the timestamps to align with the actual frame durations.

```{r recalculating-timestamp}
recalculate_timestamps <- function(metadata, time_adj) {
  frame_data <- data.frame(Frame = c(1:metadata$Cam_count[metadata$Exp_ID == paste("RS_", Exp_n, sep = "")][1]))

  frame_data <- frame_data %>%
    full_join(time_adj, by = c("Frame" = "Start_f")) %>%
    dplyr::mutate(
    # fill the NA values with the closest non-NA value.
      End_f = na.locf(End_f),
      Actual_frame = na.locf(Actual_frame),
      Duration = na.locf(Duration),
      seconds = Duration / Actual_frame, # Calculate seconds per frame
      Time = cumsum(seconds) # Compute cumulative time
    )
  return(frame_data)
}

Path_save_time_csv <- paste(Path_folder_3, "/", File_n, "_timestamp_full.csv", sep = "")
Time_adj <- read.csv(file = Path_save_time_csv) # Load the timestamp adjustment data
Frame <- recalculate_timestamps(clean_metadata, Time_adj) # Recalculate timestamps


```

Cleaning Up `detections.csv`
Clean and restructure the detection data to prepare for further analysis.
  * for dish setup with 5x5 or 5x6
  * make column name to dish number
  * add column in case columns for specific dish was missing
    + move "time" column to the rightmost
    + "time" column has to be the right most for the calculation of detection rate

```{r clean-detections}
clean_detections <- function(detection_file, frame_data, single_camera) {
  detection_data <- read.csv(file = detection_file, na.strings = c("", "NA"))[1:camera_count,] %>% # Load detection data
    dplyr::select(dplyr::contains("centroid"), time, X) %>% # Select relevant columns
    dplyr::rename_with(~ str_replace_all(., c("dishes." = "", ".centroid" = ""))) %>% # Clean column names
    dplyr::mutate(time = frame_data$Time[1:nrow(.)]) # Align time using recalculated time stamp

  # Generate all possible dish columns
  DISH <- if (single_camera) as.character(0:29) else as.character(0:24) # Determine dish range based on camera setup
  all_cols <- c(paste0(DISH, "_x"), paste0(DISH, "_y")) # Generate all dish column names

  # Add missing columns with placeholders
  for (col in setdiff(all_cols, names(detection_data))) {
    detection_data[[col]] <- c(rep(1, 5), rep(NA, nrow(detection_data) - 5)) # Add mock data for missing columns
  }

  # Ensure valid data in existing columns, handling NA-only or low-unique-value columns (that only has one value in)
  detection_data <- detection_data %>%
    dplyr::mutate(across(ends_with("_x"), ~ {
      if (all(is.na(.)) || length(unique(stats::na.omit(.))) <= 2) {  # Check for all NA or <= 2 unique values
      c(rep(1, 5), rep(NA, nrow(detection_data) - 5)) # Replacing with mock data
      } else {
        .
      }
      })) %>%
    dplyr::mutate(across(ends_with("_y"), ~ {
      if (all(is.na(.)) || length(unique(stats::na.omit(.))) <= 2) {  # Check for all NA or <= 2 unique values
      c(rep(1, 5), rep(NA, nrow(detection_data) - 5)) # Replacing with mock data
      } else {
        .
      }
    }))

  return(detection_data %>% relocate(time, .after = dplyr::last_col())) # Ensure time column is at the last position
}

my_csv <- clean_detections(Path_detection_file_2, Frame, Single_camera) # Clean detection data
```
```

```

# Recording XY Positions and Distances from the Center
## This step calculates the center coordinates and distances.

```{r calculate-distance-center}

calculate_distances <- function(simpler_csv, config_path) {
  simpler_csv %>%
    dplyr::select(-X) %>%
    tidyr:::pivot_longer(cols = -time, names_to = "variable", values_to = "value") %>%  # Convert to long format
    dplyr::mutate(
      Position = sub("^(\\d+)_.*", "\\1", variable),  # Extract dish numbers from column names
      coord = sub("^\\d+_(.*)", "\\1", variable)  # Extract coordinate types (x or y)
    ) %>%
    dplyr::select(-variable) %>%
    tidyr::pivot_wider(names_from = coord, values_from = value) %>%  # Convert back to wide format
    dplyr::mutate(Dish_numeric = suppressWarnings(as.numeric(Position))) %>%  # Add numeric Dish column
    dplyr::arrange(time, Dish_numeric) %>%  # Arrange using numeric Dish
    dplyr::select(-Dish_numeric) %>%  # Remove the temporary numeric column
    dplyr::left_join(
      read.csv(file = config_path, na.strings = c("", "NA")) %>%  # Load configuration data for dish centers
        dplyr::rename(center_x = X, center_y = Y) %>%  # Rename columns for clarity
        dplyr::mutate(Position = as.character(Position)) %>%  # Ensure Position is a character
        dplyr::select(center_x, center_y, Position, Exp_ID)  # Select relevant columns
    ) %>%
    dplyr::rename(Time = time) %>%#
    dplyr::group_by(Position) %>%  # Group by dish for center calculations
    dplyr::mutate(
      dist_cent = sqrt((x - center_x)^2 + (y - center_y)^2), # Calculate distance to center
      Position = as.numeric(Position) + ifelse(L_R == "R", 25, 0), #Adjust Position for right camera
              Position = as.character(Position)) %>%
    dplyr::select(-c(x,y)) %>%
    dplyr::mutate(Time_group = floor(Time/5) * 5) %>% # Make group of every 5 seconds
    dplyr::group_by(Position, Exp_ID, Time_group, center_x, center_y) %>%
    dplyr::summarise(med_dist_cent = median(dist_cent, na.rm = TRUE),
                     .groups = "drop") %>%
    dplyr::rename(Time = Time_group)

}

if (Single_camera == TRUE) {
  Path_config_csv <- paste(Path_config, "/RS_", File_n, "_config.csv", sep = "")} # Construct path to configuration file}
if (Single_camera != TRUE) {
  Path_config_csv <- paste(Path_config, "/RS_", File_n, "_", L_R, "_config.csv", sep = "")}  # Construct path to configuration file}


simpler_my_csv <- calculate_distances(my_csv, Path_config_csv)  # Calculate distances

#simpler_my_csv <- simpler_my_csv %>%
#  dplyr::mutate(Position = as.numeric(Position) + ifelse(L_R == "R", 25, 0),
#                Position = as.character(Position))

```

Calculating Distance Swam
This step calculates the cumulative distance swam by each fish.
  * for dish setup with 5x5 or 5x6
  * loop for each dish/fish and make a list with data frame with difference in x,y position in each row
    + omit NA
    + calculate difference in x,y position in each frame saved in column "dist_dif"
    ++ (= when fish are not detected, its position is filled by estimated trajectory as difference are calculated with NA omitted rows)

```{r calculate-swim-distance}

future::plan(future::multisession)  # Enable parallel processing

calculate_swim_distance <- function(csv_data) {
  # Define DISH based on single camera configuration
  DISH <- if (Single_camera) as.character(0:29) else as.character(0:24)  # Ensure character type for column access

  furrr::future_map(DISH, function(pos) {  # Parallel processing for each dish
    x_col <- paste(pos, "_x", sep = "")  # Construct x-coordinate column name
    y_col <- paste(pos, "_y", sep = "")  # Construct y-coordinate column name

```
  # Select relevant columns for processing
  my_csv_dish <- csv_data %>%
    dplyr::select(dplyr::all_of(c(x_col, y_col, "time")))

  # Check unique values and replace NAs in first 5 rows if necessary
  if (length(unique(my_csv_dish[[x_col]])) <= 6 | length(unique(my_csv_dish[[y_col]])) <= 6) {
    my_csv_dish[1:5,] <- replace(my_csv_dish[1:5,], is.na(my_csv_dish[1:5,]), 0)
  }

  # Omit NA rows
  my_csv_dish <- my_csv_dish %>% tidyr::drop_na()

  # Calculate differences and distances
  my_csv_dish_dif <- data.frame(diff(as.matrix(my_csv_dish))) %>%
    dplyr::rename(x_dif = 1, y_dif = 2) %>%
    dplyr::mutate(
      Time = my_csv_dish$time[2:nrow(my_csv_dish)],
      dist_dif = sqrt(x_dif^2 + y_dif^2), # Compute distance between consecutive points
      dist = cumsum(dist_dif) # Compute cumulative distance
) %>%
  dplyr::select(-time)

  as.data.frame(my_csv_dish_dif) # Return the processed data frame
 }) %>%
 setNames(DISH) # Set list names to match dish IDs
}

list_1 <- calculate_swim_distance(my_csv) # Calculate distances for all dishes

```

# Transforming to Final Dataframe
## Transform the list of distances into a consolidated dataframe.
 * change list to dataframe
 * add column with Exp_ID and L_R (left or right information)
 * adjust the position of dish depending if the video was from Left camera of Right camera
```{r transform-dataframe}
transform_to_dataframe <- function(dist_list, exp_id, lr_flag) {
  plyr::ldply(dist_list, data.frame, .id = "Position") %>%  # Combine list into a single dataframe with .id as Position
    dplyr::mutate(
      Exp_ID = paste("RS_", exp_id, sep = ""),  # Add experiment ID
      L_R = lr_flag,  # Add camera position (left or right)
    ) %>%
    dplyr::mutate(Position = as.character(Position)) %>%  # Convert Position to character
    dplyr::group_by(Position) %>%  # Group by adjusted Position
    dplyr::mutate(row_n = dplyr::row_number()) %>%  # Add row number within each position
    dplyr::filter(row_n %% FPS == 0)  # Retain every second row for further analysis
}

Distance <- transform_to_dataframe(list_1, Exp_n, L_R)  # Transform distances into final dataframe

Distance <- Distance %>%
  dplyr::mutate(Position = as.numeric(Position) + ifelse(L_R == "R", 25, 0),
                Position = as.character(Position)) # Adjust Position for right camera


```

- Calculate visual acuity and other phenotypes (test_parameters.Rmd):

```{r}

selected_condition <- 11
distance_parameter = 0.25

```

Group by the time of starting of experiment
   * Add column "Time_gr" grouping by the time where the experiment was started
```{r Group by the time of starting of experiment}


#Clean up results_full
```

```
results_full  <- results_full %>%
  dplyr::filter(!is.na(Det_rate)) #%>%
#  dplyr::filter(Strain == "HdrR")
#  dplyr::filter(Exp_ID %in% c("RS_306", "RS_312", "RS_310", "RS_325"))

#Add the time group information and add strain_dph
#Because there were cases the time were not properly added in the metadata sheet...
results_full$Record_start <- str_sub(results_full$Record_start, 1, 2)
results_full <- results_full %>%
 dplyr::mutate(Time_gr = case_when(Record_start %in% c("9:", "10") ~ "09:XX-10:XX",
                          Record_start %in% c("11", "12") ~ "11:XX-12:XX",
                          Record_start %in% c("13", "14") ~ "13:XX-14:XX",
                          Record_start %in% c("17", "18") ~ "17:XX-18:XX",
                          Record_start %in% c("15", "16") ~ "15:XX-16:XX",
                          Record_start %in% c("20", "21") ~ "20:XX-21:XX"),
        st_dph = paste(Strain, dph, "dph", sep = "_")) %>%
 dplyr::mutate(Exp_ID_pos = paste(Exp_ID, Position, sep = "_"))

# Add column of position+Strain+dph to facet
results_full <- results_full %>%
 filter(Strain %!in% c(NaN, NA)) %>%
 dplyr::mutate(pos_st_dph = paste(Position, Strain, dph, "dph", sep = "_")) %>%
 dplyr::mutate(str_width_sp = case_when(abs(Speed_deg.s) == 61.88 ~ paste(Stripe_width, "_faster",sep =""),
                             abs(Speed_deg.s) == 20.63 ~ as.character(Stripe_width),
                             abs(Speed_deg.s) == 41.26 ~ as.character(Stripe_width))) %>%
  dplyr::mutate(Strain_expid_pos = paste(Strain, Exp_ID_pos, sep ="_"))

# Set the levels to the str_width_sp to adjust the order in y axis
str_width_sp_pos <- unique(results_full$str_width_sp)
results_full$str_width_sp <- factor(results_full$str_width_sp, levels = str_width_sp_pos)

```

Select strains for plotting
```{r select strains}

results_full$Strain <- sub("@.*", "", results_full$Strain)
results_full$Strain <- sub(")/.*", ")", results_full$Strain)
results_xy_pos_full$Strain <- sub("@.*", "", results_xy_pos_full$Strain)
results_xy_pos_full$Strain <- sub(")/.*", ")", results_xy_pos_full$Strain)
if(exists("F0_analysis") == TRUE){
results_full$Strain <- sub("/.*", "", results_full$Strain)
results_xy_pos_full$Strain <- sub("/.*", "", results_xy_pos_full$Strain)
}

#results_full$Strain <- sub("/.*", "", results_full$Strain)

#results_full$Strain <- sub(").*", "", results_full$Strain)
selected_strain <- unique(results_full$Strain)
#selected_strain <- setdiff(unique(results_full$Strain), "HdrR")
#selected_strain <-  c("IP75-1", "IP127-2", "IP79-2", "IP130-2", "IP47-1", "IP23-1-1",  "IP139-4")
Exclute_fish  <-  c("ObstacleNotCenterIP(139-4x79-2)",  "ObstacleNotCenter",  "IP(23-1/1x139-4)-mixed?",  "2-fish-IP(139-4x130-2)",
"2-fish?-IP(139-4x127-2)")

selected_strain <- selected_strain[!selected_strain %in% Exclute_fish]


```


Threshold calculation

```{r Threshold calculation}

results_full <- results_full %>%
  dplyr::filter(Strain %in% selected_strain) %>%
  dplyr::mutate(str_width_dir = paste(Stripe_width, Direction, Speed_deg.s, sep = "_")) %>%
  dplyr::group_by(Strain_expid_pos, str_width_dir) %>%
  dplyr::mutate(phase_n = row_number(),
                str_width_dir_n = paste(str_width_dir, phase_n, sep = "_"),
                str_abssp = paste(Stripe_width, abs(Speed_deg.s), sep ="_")) %>%
  distinct()

# make a new dataframe and add columns counting how many time fish was at
```

```
# Strain_expid_pos: pause
# count_pass_p: following stripe/swam in the same direction
# count_pass_n: following stripe/swam in the opposite direction
# in each phase (CW/CCW/Pause x Stripe width x speed)

Threshold <- results_full %>%
 ungroup() %>%
 select(Direction:Speed_deg.s, End_s, Position:StockID, Exp_ID:Stripe_parameter, Dish:str_width_sp, Exp_ID_pos, str_width_dir) %>%
 dplyr::group_by(Exp_ID, Position, str_width_dir) %>%
 dplyr::mutate(
   count_pass_p = sum(Response_value >= 0.5, na.rm = TRUE),
   count_pass_n = sum(Response_value <= -0.5, na.rm = TRUE),
   count_paused = sum(Response_value < 0.3 & Response_value > -0.3, na.rm = TRUE),
   Strain_expid_pos = paste(Strain, Exp_ID, Position, sep ="_"),
   str_abssp = paste(Stripe_width, abs(Speed_deg.s), sep ="_")
 ) #%>%

# add column indicating if the fish responded in the possitive or negative direction more than half of the phase
# add column indicating if the fish did one directional movement
Threshold <- Threshold %>%
 distinct() %>%
#  ungroup() %>%
#  dplyr::mutate(str_cols_dir = paste(str_cols, Direction, Speed_deg.s, sep = "_")) %>%
#  dplyr::group_by(Strain_expid_pos, str_cols_dir) %>%
#  dplyr::mutate(phase_n = row_number(),
#                str_cols_dir_n = paste(str_cols_dir, phase_n, sep = "_"))
 ungroup() %>%
 dplyr::group_by(Exp_ID_pos, str_abssp) %>%
 dplyr::reframe(
#    Responded = if_else(
#      (count_pass_p[Direction == "CW"] > 7 &
#      count_pass_p[Direction == "CCW"] > 7 &
#      count_paused[Direction == "P_2"] > 2) |
#      (count_pass_n[Direction == "CW"] > 7 &
#      count_pass_n[Direction == "CCW"] > 7 &
#      count_paused[Direction == "P_2"] > 2),
#      first(str_cols),
#      factor(NA, levels = levels(str_cols))),
   Responded = if_else(
     (!is.na(count_pass_p[Direction == "CW"]) & count_pass_p[Direction == "CW"] > 15 &
     !is.na(count_pass_p[Direction == "CCW"]) & count_pass_p[Direction == "CCW"] > 15 ) |
     (!is.na(count_pass_n[Direction == "CW"]) & count_pass_n[Direction == "CW"] > 15 &
     !is.na(count_pass_n[Direction == "CCW"]) & count_pass_n[Direction == "CCW"] > 15 ),
     first(str_abssp, na.rm = TRUE),
     factor(NA, levels = levels(str_abssp))),
   One_direction = if_else(
#      (count_pass_p[Direction == "CW"] > 20 &
##       count_pass_n[Direction == "CCW"] > 20) |
#      (count_pass_n[Direction == "CW"] > 20 &
#      count_pass_p[Direction == "CCW"] > 20) ,
     (!is.na(count_pass_p[Direction == "CW"]) & count_pass_p[Direction == "CW"] > 20 &
     !is.na(count_pass_n[Direction == "CCW"]) & count_pass_n[Direction == "CCW"] > 20) |
     (!is.na(count_pass_n[Direction == "CW"]) & count_pass_n[Direction == "CW"] > 20 &
     !is.na(count_pass_p[Direction == "CCW"]) & count_pass_p[Direction == "CCW"] > 20),
     first(str_abssp, na.rm = TRUE),
     factor(NA, levels = levels(str_abssp))
 )) %>%
 distinct()

Threshold <- Threshold %>%
 dplyr::mutate(One_direction = if_else(!is.na(One_direction), TRUE, NA)) %>%
#  full_join(Contrast) %>%
   dplyr::mutate(Responded = if_else(!is.na(Responded), TRUE, NA))

#View(Threshold %>%dplyr::filter(Exp_ID_pos == "RS_336_18"))

results_full <- results_full %>%
 distinct() %>%
 full_join(Threshold)


##### If we want to exclude one directional swimmer
#Exclude <- results_full %>%
#  select(Exp_ID_pos, str_sp, One_direction) %>%
#  distinct() %>%
```

```
#  dplyr::group_by(Exp_ID_pos) %>%
#  dplyr::mutate(
#    Exclude = ifelse(
#      lead(One_direction) == TRUE,
#      TRUE,
#      FALSE
#    )
#  )


#results_full <- results_full %>%
#  full_join(Exclude)

####

Threshold_SF <- results_full %>%
  ungroup()%>%
  dplyr::select(Strain, Exp_ID_pos, Responded, Stripe_width) %>%
  dplyr::filter(!is.na(Responded)
#                & is.na(Exclude)
                ) %>%
  distinct() %>%
  dplyr::group_by(Exp_ID_pos) %>%
  dplyr::summarise(Thinnest_st_wid = min(Stripe_width)) %>%
#  dplyr::select(-c(Responded,
#                   str_abssp,
#                   Exclude,
#                   One_direction,
#                   Stripe_width
#)) %>%
  distinct()

results_full <- results_full %>%
  full_join(Threshold_SF)


Threshold_SF_full <- results_full %>%
  ungroup() %>%
  dplyr::select(Strain, Exp_ID_pos, Responded, Stripe_width) %>%
  dplyr::filter(!is.na(Responded)
#                & is.na(Exclude)
                ) %>%
  distinct() %>%
  dplyr::group_by(Exp_ID_pos) %>%
  dplyr::mutate(Responded_st_wid = Stripe_width) %>%
#  dplyr::select(-c(Responded,
#                   str_abssp,
#                   Exclude,
#                   One_direction,
#                   Stripe_width
#)) %>%
  distinct()

results_full <- results_full %>%
  full_join(Threshold_SF_full)


```

Calculating distances from the center for each stripe width
```{r}

# Combine dataframe VA-calc with xy (Start_t = Time)
## - Dir_res: Positive or Negative dir = Median response value + or - ?
## - keep only these rows with RV > 0.5 or RV < -0.5 depending on Dir_res
## - keep only Responded = TRUE
# - Filter dist_cent with Responded = TRUE
# - make new column top_25p_cent_dist to keep only 25 % of most inner position from dist_cent
# - make new column for med_top_25p_cent_dist
# - keep only this and stripe_width & Exp_ID_pos
# - make a new column with dist_from_rim_mm
# - calculate visual acuity for each individual fish
# - make a line plot of VA vs stripe_width for individual fish to see if the line slopes are the same
# check seconds and frames are the same between distance df and response value df
  # dist_from_rim_mm: keep only longer distance between CW or CCW
```

```
VA_calc <- results_full %>%
  select(Direction:Start_s, Position, Response_value,
         dpf, Time_gr, Exp_ID_pos, Exp_ID,
         str_width_dir, Strain_expid_pos, str_abssp:Responded_st_wid) %>%
  dplyr::group_by(Exp_ID, Position, str_width_dir) %>%
  dplyr::mutate(med_Response_value = median(Response_value, na.rm = TRUE))

# Check if median is fetched correctly
#p <- VA_calc %>%
#  dplyr::filter(Exp_ID == "RS_201", Position == 2) %>%
#  ggplot() +
#  geom_histogram(aes(x = Response_value)) +
#  geom_histogram(aes(x = med_Response_value), color= "red") +
#  facet_wrap(~str_width_dir)

#ggsave(filename = "test.png",plot = p, path = "/Users/Risa/Desktop",
#       width = 20, height = 7)

temp <- VA_calc %>%
  dplyr::select(Position, Direction, Exp_ID_pos,
                str_width_dir, str_abssp, Response_value) %>%
  distinct() %>%
  dplyr::filter(Direction %in% c("CW", "CCW")) %>%
  dplyr::group_by(Exp_ID_pos, str_abssp) %>%
  dplyr::mutate(res_dir = median(Response_value, na.rm = TRUE))

# Check if median is fetched correctly
#p <- temp %>%
#  dplyr::filter(Exp_ID == "RS_201", Position == 2) %>%
#  ggplot() +
#  geom_histogram(aes(x = Response_value)) +
#  geom_histogram(aes(x = res_dir), color= "red") +
#  facet_wrap(~str_abssp)

#ggsave(filename = "test.png",plot = p, path = "/Users/Risa/Desktop",
#       width = 20, height = 7)


VA_calc <- VA_calc %>%
  full_join(temp)

VA_calc$str_abssp <- factor(VA_calc$str_abssp, levels = unique(VA_calc$str_abssp))

#p <- VA_calc %>%
#  dplyr::filter(Direction %in% c("CW", "CCW")) %>%
#  ggplot() +
#  geom_tile(aes(x = str_abssp, y = Exp_ID_pos, fill = res_dir)) +
#scale_fill_gradientn(
#  colors = c("black", "black", "gray", "gray", "gray", "yellow", "yellow"),
#  values = scales::rescale(c(-1.75, -0.5, -0.2, 0.2, 0.5, 1.75)),
#  breaks = c(-1.75, -0.5, -0.2, 0.2, 0.5, 1.75),
#  labels = c("~ -1.75", "-1.75 ~ -0.5", "-0.5 ~ -0.2", "-0.2 ~ 0.2", "0.2 ~ 0.5", "0.5 ~ 1.75"))


#p <- VA_calc %>%
#  dplyr::filter(Direction %in% c("CW", "CCW")) %>%
#  ggplot() +
#  geom_point(aes(x = Start_s, y = res_dir)) +
#  geom_hline(aes(yintercept = -1)) +
#  facet_wrap(~Exp_ID_pos)


#ggsave(filename = "test.png",plot = p, path = "/Users/Risa/Desktop",
#       width = 20, height = 7)


VA_calc <- results_xy_pos_full %>%
  dplyr::rename(Start_s = Time) %>%
  full_join(VA_calc)

VA_calc <- VA_calc %>%
  dplyr::filter(Direction %in% c("CW", "CCW"),
                Responded == TRUE,
                (res_dir >= 0.5 & Response_value >= 0.5) | (res_dir <= -0.5 & Response_value <= -0.5)) %>%
```

```
 ungroup() %>%
 dplyr::group_by(Position, Exp_ID, str_width_dir) %>%
dplyr::mutate(
   top_25p_cent_dist = ifelse(med_dist_cent <= quantile(med_dist_cent, distance_parameter, na.rm = TRUE), med_dist_cent, NA),
   med_top_25p_cent_dist = median(top_25p_cent_dist, na.rm = TRUE),
   dist_from_rim_mm = mm_per_pix * (Radius - med_top_25p_cent_dist)
 )

Distance_rim <- VA_calc %>%
 ungroup() %>%
 dplyr::select(Direction, Stripe_width, Speed_deg.s,
           dist_from_rim_mm, Exp_ID_pos,str_abssp) %>%
 dplyr::group_by(Exp_ID_pos, str_abssp) %>%
 # take the median
 dplyr::mutate(dist_from_rim_mm_far = mean(dist_from_rim_mm, na.rm = TRUE)) %>%
 distinct()


temp <- results_full %>%
 dplyr::filter(Direction %in% c("CW", "CCW"))


```

VA calculation

```{r, Visual Acuity calculation}

# Function to count continuous rows with values > 0.5
count_continuous_streaks <- function(values) { rle_result <- rle(values >= 0.5) # Run Length Encoding
streak_lengths <- rle_result$lengths[rle_result$values] # Extract streaks where value > 0.5
if (length(streak_lengths) == 0) {
 return(0) # No streaks found
} else {
 return(max(streak_lengths)) # Return the longest streak
}
}

count_continuous_streaks_n <- function(values) { rle_result <- rle(values <= -0.5) # Run Length Encoding
streak_lengths <- rle_result$lengths[rle_result$values] # Extract streaks where value > 0.5
if (length(streak_lengths) == 0) {
 return(0) # No streaks found
} else {
 return(max(streak_lengths)) # Return the longest streak
}
}


Summary_RV <- results_full %>%
 ungroup() %>%
 dplyr::select(Exp_ID_pos, Direction, Response_value, str_width_dir_n, str_width_dir, One_direction) %>%
 distinct() %>%
 dplyr::filter(Direction %in% c("CW", "CCW")) %>%
 dplyr::group_by(Exp_ID_pos, str_width_dir) %>%
 dplyr::mutate(longest_response = count_continuous_streaks(Response_value),
           longest_response_n = count_continuous_streaks_n(Response_value)) %>%
 ungroup() %>%
 dplyr::select(-str_width_dir) %>%
 dplyr::group_by(Exp_ID_pos, Direction) %>%
 dplyr::mutate(longest_response =  max(longest_response),
           longest_response_n =  max(longest_response_n),
           Sum_res = sum(Response_value),
           Sum_res_abs = sum(abs(Response_value))) %>%
 dplyr::mutate(longest_response = pmax(longest_response, longest_response_n, na.rm = TRUE)) %>%
 select(-longest_response_n) %>%
#  dplyr::filter(One_direction != TRUE) %>%
 dplyr::mutate(blind = case_when(
      # Condition 1
   selected_condition == 1 & abs(Sum_res) < 150 & Sum_res_abs < 300 ~ TRUE,

   # Condition 2
   selected_condition == 2 & abs(Sum_res) < 250 & Sum_res_abs < 300 ~ TRUE,

   # Condition 3
   selected_condition == 3 & longest_response < 10 ~ TRUE,
```

```
    # Condition 4
    selected_condition == 4 & longest_response < 15 ~ TRUE,

    # Condition 5 (Condition 1 & Condition 3)
    selected_condition == 5 & abs(Sum_res) < 150 & longest_response < 10 & Sum_res_abs < 300 ~ TRUE,

    # Condition 6 (Condition 1 & Condition 4)
    selected_condition == 6 & abs(Sum_res) < 150 & longest_response < 15 & Sum_res_abs < 300 ~ TRUE,

    # Condition 7 (Condition 2 & Condition 3)
    selected_condition == 7 & abs(Sum_res) < 250 & longest_response < 10 & Sum_res_abs < 300 ~ TRUE,

    # Condition 8 (Condition 2 & Condition 4)
    selected_condition == 8 & abs(Sum_res) < 250 & longest_response < 15 & Sum_res_abs < 300 ~ TRUE,
    selected_condition == 9 & longest_response < 9 & Sum_res_abs < 300 ~ TRUE,
    selected_condition == 10 & longest_response < 10 & Sum_res_abs < 300 ~ TRUE,
    selected_condition == 11 & longest_response < 11 & Sum_res_abs < 300 ~ TRUE,
    selected_condition == 12 & longest_response < 8 & Sum_res_abs < 300 ~ TRUE)) %>%
  ungroup() %>%
  dplyr::select(Exp_ID_pos, longest_response, Sum_res, Sum_res_abs, blind) %>%
  dplyr::group_by(Exp_ID_pos) %>%
   distinct() %>%
  slice_min(longest_response) %>%
  slice_min(abs(Sum_res)) %>%
  slice_min(Sum_res_abs)



#results_full$Strain <- sub("@.*", "", results_full$Strain)
#results_full$Strain <- sub("/.*", "", results_full$Strain)

Total_n <- results_full %>%
  ungroup() %>%
  dplyr::select(Strain, Exp_ID_pos) %>%
  distinct() %>%
  group_by(Strain) %>%
  dplyr::summarise(total_n = n())

results_full <- results_full %>%
  left_join(Total_n) %>%
  left_join(Summary_RV) %>%
          dplyr::mutate(Strain_n = paste(Strain, " (n = ",
                                total_n, ")", sep = ""))

Stripe_thr <- results_full %>%
  dplyr::filter(Direction %in% c("CW", "CCW")) %>%
  full_join(Distance_rim) %>%
  ungroup() %>%
  dplyr::filter(Responded_st_wid == Thinnest_st_wid |
              blind == TRUE & is.na(Thinnest_st_wid)) %>%
#  dplyr::filter(Strain == "HdrR") %>%
  dplyr::select(Strain, Strain_expid_pos, Thinnest_st_wid,
              Sum_res_abs, blind, longest_response,
              total_n, Strain_n, dist_from_rim_mm_far,
#               Responded_st_wid
              ) %>%
  dplyr::group_by(Strain_expid_pos) %>%
#  dplyr::filter(Responded_st_wid == min(Responded_st_wid, na.rm = TRUE)) %>%
  distinct() %>%
  dplyr::mutate(
     # * (180 / pi) is used to convert the angular size from radians (default output of atan() in most programming languages, including R and
Python) to degrees.
    Stripe_Width_deg = 2 * atan(Thinnest_st_wid / (2 * dist_from_rim_mm_far)) * (180 / pi),
    # Calculate Spatial Frequency (cycles/degree)
    Spatial_Frequency_cpd = 1 / Stripe_Width_deg,
    # Calculate Visual Acuity (dimensionless)
    Visual_Acuity = case_when(!is.na(blind) & is.na(Thinnest_st_wid) ~ 0,
                          TRUE ~ 1 / Stripe_Width_deg)
) %>%
  distinct()



VA_all_pos <- results_full %>%
```

```
  dplyr::filter(Direction %in% c("CW", "CCW")) %>%
  full_join(Distance_rim) %>%
  ungroup() %>%
#  dplyr::filter(Strain == "HdrR") %>%
  dplyr::select(Strain, Strain_expid_pos,
                Sum_res_abs, blind, longest_response,
                total_n, Strain_n, dist_from_rim_mm_far,
                Responded_st_wid, One_direction
                ) %>%
  distinct() %>%
  dplyr::mutate(
     # * (180 / pi) is used to convert the angular size from radians (default output of atan() in most programming languages, including R and
Python) to degrees.
    Stripe_Width_deg_all = 2 * atan(Responded_st_wid / (2 * dist_from_rim_mm_far)) * (180 / pi),
    # Calculate Spatial Frequency (cycles/degree)
    Spatial_Frequency_cpd_all = 1 / Stripe_Width_deg_all,
    # Calculate Visual Acuity (dimensionless)
    Visual_Acuity_all = case_when(is.na(One_direction) & !is.na(blind) & is.na(Responded_st_wid) ~ 0,
                          TRUE ~ 1 / Stripe_Width_deg_all)
) %>%
  distinct()


temp <- VA_all_pos %>%
  ungroup() %>%
  dplyr::group_by(Strain_expid_pos) %>%
  dplyr::select(- c(Responded_st_wid,
                    #Stripe_Width_deg_all,
                  #longest_response, Sum_res_abs
                  )) %>%
  dplyr::slice_max(order_by = Visual_Acuity_all, na_rm = TRUE) %>%
  dplyr::rename(Visual_Acuity = Visual_Acuity_all) %>%
  distinct()


VA <- temp

Total_n_VA <- VA %>%
  distinct() %>%
  group_by(Strain) %>%
  dplyr::summarise(total_n_VA = n()) %>%
  dplyr::mutate(Strain_n_VA = paste(Strain, " (n = ",
                                    total_n_VA, ")", sep = ""))

temp <- Stripe_thr %>%
  ungroup() %>%
  dplyr::select(Strain_expid_pos, Thinnest_st_wid) %>%
  distinct()

VA <- VA %>%
  full_join(Total_n_VA) %>%
  full_join(temp)

```

Cleaning up high false positive detection & missing detection
```{r cleaning up}
#--------------------
# Filter out those
  # False positive detection: detecting bubble near obstacle in the beginning
  # False positive detection: detecting stripes
  # Failure in detection: Beginning
  # Failure in detection: Entire video
#--------------------

# False positive detection: detecting bubble near obstacle in the beginning
# Filter out these rows that have med_dist_cent lower than XX during the time of 0 to 450 s
# Failure in detection: Beginning 18 for roi53

Filtering_fp <- results_xy_pos_full %>%
  dplyr::group_by(Exp_ID, Position) %>%
  dplyr::filter(sum(med_dist_cent[Time >= 300 & Time <= 450] < 18, na.rm = TRUE) /
        sum(Time >= 300 & Time <= 450, na.rm = TRUE) > 0.5)
```

```
False_pos <- Filtering_fp %>%
  select(Exp_ID, Position,Strain) %>%
  distinct() %>%
  dplyr::mutate(False_detection = TRUE,
                Exp_ID_pos = paste(Exp_ID, Position, sep ="_"),
                Strain_expid_pos = paste(Strain, Exp_ID, Position, sep ="_"))

Path_save_csv_fp <- paste(Save_path, "/False_positive_detection.csv", sep = "")
write.csv(False_pos, Path_save_csv_fp, row.names=TRUE)

  # False positive detection: detecting stripes - ignore for now

  # Failure in detection: Entire video - not important?


```


Removing false positive

```{r}

VA <- VA %>%
  dplyr::filter(Strain_expid_pos %!in% unique(False_pos$Strain_expid_pos))

results_full <- results_full %>%
    dplyr::filter(Strain_expid_pos %!in% unique(False_pos$Strain_expid_pos))

results_xy_pos_full <- results_xy_pos_full %>%
  dplyr::mutate(Strain_expid_pos = paste(Strain, Exp_ID, Position, sep ="_")) %>%
  dplyr::filter(Strain_expid_pos %!in% unique(False_pos$Strain_expid_pos),
                !is.na(Strain))
temp <- VA %>%
  dplyr::select(Strain_expid_pos, Visual_Acuity)

results_full <- results_full %>%
  full_join(temp)

Exp_info <- results_full %>%
  ungroup() %>%
  dplyr::select(Exp_ID_pos, Strain_expid_pos)

One_direction_times <- Threshold %>%
  dplyr::filter(One_direction == TRUE) %>%
  dplyr::group_by(Exp_ID_pos) %>%
  dplyr::mutate(One_direction_count = n()) %>%
  dplyr::select(Exp_ID_pos, One_direction_count) %>%
  distinct() %>%
  left_join(Exp_info) %>%
  distinct()


VA <- VA %>%
  left_join(One_direction_times) %>%
  mutate(One_direction_count = replace_na(One_direction_count, 0))


Path_save_csv <- paste(Save_path, "/VA.csv", sep = "")
write.csv(VA, Path_save_csv, row.names=TRUE)


Path_save_csv <- paste(Save_path, "/results_full_VA.csv", sep = "")
write.csv(results_full, Path_save_csv, row.names=TRUE)

```

5.3.3 Stripe Motion Automation Setting

The following setting was used for the automated stripe motion generation.

Duration (s), number of stripes, rotation speed (rad/s), stripe color 1, stripe color 2
Standard stripe motion with varying stripe width, speed: 20.6°/s (Parameter 1-6):
60,2,0, #000000 , #000000
240,94,0, #FFFFFF , #000000
150,94,0.36, #FFFFFF , #000000
30,94,0, #FFFFFF , #000000
150,94,-0.36, #FFFFFF , #000000
30,71,0, #FFFFFF , #000000
150,71,0.36, #FFFFFF , #000000
30,71,0, #FFFFFF , #000000
150,71,-0.36, #FFFFFF , #000000
30,57,0, #FFFFFF , #000000
150,57,0.36, #FFFFFF , #000000
30,57,0, #FFFFFF , #000000
150,57,-0.36, #FFFFFF , #000000
30,47,0, #FFFFFF , #000000
150,47,0.36, #FFFFFF , #000000
30,47,0, #FFFFFF , #000000
150,47,-0.36, #FFFFFF , #000000
30,40,0, #FFFFFF , #000000
150,40,0.36, #FFFFFF , #000000
30,40,0, #FFFFFF , #000000
150,40,-0.36, #FFFFFF , #000000
30,35,0, #FFFFFF , #000000
150,35,0.36, #FFFFFF , #000000
30,35,0, #FFFFFF , #000000
150,35,-0.36, #FFFFFF , #000000
30,31,0, #FFFFFF , #000000
150,31,0.36, #FFFFFF , #000000
30,31,0, #FFFFFF , #000000
150,31,-0.36, #FFFFFF , #000000
30,28,0, #FFFFFF , #000000
150,28,0.36, #FFFFFF , #000000
30,28,0, #FFFFFF , #000000
150,28,-0.36, #FFFFFF , #000000
30,26,0, #FFFFFF , #000000
150,26,0.36, #FFFFFF , #000000
30,26,0, #FFFFFF , #000000
150,26,-0.36, #FFFFFF , #000000
30,24,0, #FFFFFF , #000000
150,24,0.36, #FFFFFF , #000000
30,24,0, #FFFFFF , #000000
150,24,-0.36, #FFFFFF , #000000
30,22,0, #FFFFFF , #000000
150,22,0.36, #FFFFFF , #000000
30,22,0, #FFFFFF , #000000
150,22,-0.36, #FFFFFF , #000000
30,20,0, #FFFFFF , #000000
150,20,0.36, #FFFFFF , #000000
30,20,0, #FFFFFF , #000000
150,20,-0.36, #FFFFFF , #000000
30,19,0, #FFFFFF , #000000
150,19,0.36, #FFFFFF , #000000

30,19,0, #FFFFFF , #000000
150,19,-0.36, #FFFFFF , #000000
30,18,0, #FFFFFF , #000000
150,18,0.36, #FFFFFF , #000000
30,18,0, #FFFFFF , #000000
150,18,-0.36, #FFFFFF , #000000
30,17,0, #FFFFFF , #000000
150,17,0.36, #FFFFFF , #000000
30,17,0, #FFFFFF , #000000
150,17,-0.36, #FFFFFF , #000000
30,16,0, #FFFFFF , #000000
150,16,0.36, #FFFFFF , #000000
30,16,0, #FFFFFF , #000000
150,16,-0.36, #FFFFFF , #000000
30,15,0, #FFFFFF , #000000
150,15,0.36, #FFFFFF , #000000
30,15,0, #FFFFFF , #000000
150,15,-0.36, #FFFFFF , #000000
30,14,0, #FFFFFF , #000000
150,14,0.36, #FFFFFF , #000000
30,14,0, #FFFFFF , #000000
150,14,-0.36, #FFFFFF , #000000
30,14,0, #FFFFFF , #000000
150,14,1.08, #FFFFFF , #000000
30,14,0, #FFFFFF , #000000
150,14,-1.08, #FFFFFF , #000000
30,7,0, #FFFFFF , #000000
150,7,0.36, #FFFFFF , #000000
30,7,0, #FFFFFF , #000000
150,7,-0.36, #FFFFFF , #000000
30,7,0, #FFFFFF , #000000
150,7,1.08, #FFFFFF , #000000
30,7,0, #FFFFFF , #000000
150,7,-1.08, #FFFFFF , #000000
150,7,0, #FFFFFF , #000000

Stripe motion with varying stripe width, speed: 41.2°/s, (Parameter 1-7):
60,2,0, #000000 , #000000
240,94,0, #FFFFFF , #000000
150,94,0.72, #FFFFFF , #000000
30,94,0, #FFFFFF , #000000
150,94,-0.72, #FFFFFF , #000000
30,71,0, #FFFFFF , #000000
150,71,0.72, #FFFFFF , #000000
30,71,0, #FFFFFF , #000000
150,71,-0.72, #FFFFFF , #000000
30,57,0, #FFFFFF , #000000
150,57,0.72, #FFFFFF , #000000
30,57,0, #FFFFFF , #000000
150,57,-0.72, #FFFFFF , #000000
30,47,0, #FFFFFF , #000000
150,47,0.72, #FFFFFF , #000000
30,47,0, #FFFFFF , #000000
150,47,-0.72, #FFFFFF , #000000
30,40,0, #FFFFFF , #000000
150,40,0.72, #FFFFFF , #000000
30,40,0, #FFFFFF , #000000
150,40,-0.72, #FFFFFF , #000000

30,35,0, #FFFFFF , #000000
150,35,0.72, #FFFFFF , #000000
30,35,0, #FFFFFF , #000000
150,35,-0.72, #FFFFFF , #000000
30,31,0, #FFFFFF , #000000
150,31,0.72, #FFFFFF , #000000
30,31,0, #FFFFFF , #000000
150,31,-0.72, #FFFFFF , #000000
30,28,0, #FFFFFF , #000000
150,28,0.72, #FFFFFF , #000000
30,28,0, #FFFFFF , #000000
150,28,-0.72, #FFFFFF , #000000
30,26,0, #FFFFFF , #000000
150,26,0.72, #FFFFFF , #000000
30,26,0, #FFFFFF , #000000
150,26,-0.72, #FFFFFF , #000000
30,24,0, #FFFFFF , #000000
150,24,0.72, #FFFFFF , #000000
30,24,0, #FFFFFF , #000000
150,24,-0.72, #FFFFFF , #000000
30,22,0, #FFFFFF , #000000
150,22,0.72, #FFFFFF , #000000
30,22,0, #FFFFFF , #000000
150,22,-0.72, #FFFFFF , #000000
30,20,0, #FFFFFF , #000000
150,20,0.72, #FFFFFF , #000000
30,20,0, #FFFFFF , #000000
150,20,-0.72, #FFFFFF , #000000
30,19,0, #FFFFFF , #000000
150,19,0.72, #FFFFFF , #000000
30,19,0, #FFFFFF , #000000
150,19,-0.72, #FFFFFF , #000000
30,18,0, #FFFFFF , #000000
150,18,0.72, #FFFFFF , #000000
30,18,0, #FFFFFF , #000000
150,18,-0.72, #FFFFFF , #000000
30,17,0, #FFFFFF , #000000
150,17,0.72, #FFFFFF , #000000
30,17,0, #FFFFFF , #000000
150,17,-0.72, #FFFFFF , #000000
30,16,0, #FFFFFF , #000000
150,16,0.72, #FFFFFF , #000000
30,16,0, #FFFFFF , #000000
150,16,-0.72, #FFFFFF , #000000
30,15,0, #FFFFFF , #000000
150,15,0.72, #FFFFFF , #000000
30,15,0, #FFFFFF , #000000
150,15,-0.72, #FFFFFF , #000000
30,14,0, #FFFFFF , #000000
150,14,0.72, #FFFFFF , #000000
30,14,0, #FFFFFF , #000000
150,14,-0.72, #FFFFFF , #000000
30,14,0, #FFFFFF , #000000
150,14,1.08, #FFFFFF , #000000
30,14,0, #FFFFFF , #000000
150,14,-1.08, #FFFFFF , #000000
30,7,0, #FFFFFF , #000000
150,7,0.72, #FFFFFF , #000000

30,7,0, #FFFFFF , #000000
150,7,-0.72, #FFFFFF , #000000
30,7,0, #FFFFFF , #000000
150,7,1.08, #FFFFFF , #000000
30,7,0, #FFFFFF , #000000
150,7,-1.08, #FFFFFF , #000000
150,7,0, #FFFFFF , #000000

Stripe motion with varying contrast and color (Parameter CvGr2):
60,2,0,#000000,#000000
240,14,0,#838383,#7C7C7C
90,14,0.36,#838383,#7C7C7C
30,14,0,#838383,#7C7C7C
90,14,-0.36,#838383,#7C7C7C
30,14,0,#878787,#787878
90,14,0.36,#878787,#787878
30,14,0,#878787,#787878
90,14,-0.36,#878787,#787878
30,14,0,#8B8B8B,#747474
90,14,0.36,#8B8B8B,#747474
30,14,0,#8B8B8B,#747474
90,14,-0.36,#8B8B8B,#747474
30,14,0,#999999,#666666
90,14,0.36,#999999,#666666
30,14,0,#999999,#666666
90,14,-0.36,#999999,#666666
30,14,0,#B3B3B3,#4D4D4D
90,14,0.36,#B3B3B3,#4D4D4D
30,14,0,#B3B3B3,#4D4D4D
90,14,-0.36,#B3B3B3,#4D4D4D
30,14,0,#D9D9D9,#262626
90,14,0.36,#D9D9D9,#262626
30,14,0,#D9D9D9,#262626
90,14,-0.36,#D9D9D9,#262626
30,14,0,#FFFFFF,#000000
90,14,0.36,#FFFFFF,#000000
30,14,0,#FFFFFF,#000000
90,14,-0.36,#FFFFFF,#000000
180,14,0,#000000,#000020
90,14,0.36,#000000,#000020
30,14,0,#000000,#000020
90,14,-0.36,#000000,#000020
30,14,0,#000000,#000040
90,14,0.36,#000000,#000040
30,14,0,#000000,#000040
90,14,-0.36,#000000,#000040
30,14,0,#000000,#00007F
90,14,0.36,#000000,#00007F
30,14,0,#000000,#00007F
90,14,-0.36,#000000,#00007F
30,14,0,#000000,#0000FF
90,14,0.36,#000000,#0000FF
30,14,0,#000000,#0000FF
90,14,-0.36,#000000,#0000FF
180,14,0,#000000,#200000
90,14,0.36,#000000,#200000
30,14,0,#000000,#200000
90,14,-0.36,#000000,#200000

30,14,0,#000000,#400000
90,14,0.36,#000000,#400000
30,14,0,#000000,#400000
90,14,-0.36,#000000,#400000
30,14,0,#000000,#7F0000
90,14,0.36,#000000,#7F0000
30,14,0,#000000,#7F0000
90,14,-0.36,#000000,#7F0000
30,14,0,#000000,#FF0000
90,14,0.36,#000000,#FF0000
30,14,0,#000000,#FF0000
90,14,-0.36,#000000,#FF0000
180,14,0,#000000,#002000
90,14,0.36,#000000,#002000
30,14,0,#000000,#002000
90,14,-0.36,#000000,#002000
30,14,0,#000000,#004000
90,14,0.36,#000000,#004000
30,14,0,#000000,#004000
90,14,-0.36,#000000,#004000
30,14,0,#000000,#007F00
90,14,0.36,#000000,#007F00
30,14,0,#000000,#007F00
90,14,-0.36,#000000,#007F00
30,14,0,#000000,#00FF00
90,14,0.36,#000000,#00FF00
30,14,0,#000000,#00FF00
90,14,-0.36,#000000,#00FF00
600,14,0,#ff7b00,#ff00b2
90,14,0.36,#ff7b00,#ff00b2
30,14,0,#ff7b00,#ff00b2
90,14,-0.36,#ff7b00,#ff00b2
30,14,0,#FF76D5,#BABABA
90,14,0.36,#FF76D5,#BABABA
30,14,0,#FF76D5,#BABABA
90,14,-0.36,#FF76D5,#BABABA
30,14,0,#FF76D5,#77D6FD
90,14,0.36,#FF76D5,#77D6FD
30,14,0,#FF76D5,#77D6FD
90,14,-0.36,#FF76D5,#77D6FD
30,14,0,#ccc200,#29cc00
90,14,0.36,#ccc200,#29cc00
30,14,0,#ccc200,#29cc00
90,14,-0.36,#ccc200,#29cc00
30,14,0,#33ff00,#ff7b00
90,14,0.36,#33ff00,#ff7b00
30,14,0,#33ff00,#ff7b00
90,14,-0.36,#33ff00,#ff7b00
30,14,0,#666666,#ccc200
90,14,0.36,#666666,#ccc200
30,14,0,#666666,#ccc200
90,14,-0.36,#666666,#ccc200
30,14,0,#05b4ff,#05ff05
90,14,0.36,#05b4ff,#05ff05
30,14,0,#05b4ff,#05ff05
90,14,-0.36,#05b4ff,#05ff05
30,14,0,#3B3BFF,#a13dff
90,14,0.36,#3B3BFF,#a13dff

30,14,0,#3B3BFF,#a13dff
90,14,-0.36,#3B3BFF,#a13dff
30,14,0,#3B3BFF,#a13df

6
Appendix

6.1 Evaluation of Retinal Regeneration Diversity in Medaka Strains

One of the key advantages of studying model organisms such as fish lies in their remarkable ability to regenerate their own tissues. This regenerative capacity can vary significantly across species. The previous study showed that in contrast to zebrafish, where Muller glia (MG) cells can regenerate all retinal cell types by generating proliferating neurogenic clusters and eventually differentiate into retinal neurons (Nagashima et al., 2013), MG cells of Cab strain medaka exhibit a restricted regenerative capacity. After the injury, MG cells of Cab strain medaka reenter the cell cycle, proliferate but do not generate neurogenic clusters and ultimately restore only photoreceptors (Lust & Wittbrodt, 2018).

Katharina Lust in the lab previously reported that the regenerative capacity between Cab and HdrR differs and HdrR showed neurogenic cluster-like phenotype. These findings raise important questions about the diversity in regenerative ability across strains. The phenotypic diversity present within the MIKK panel further offers a unique opportunity to explore variation in regenerative capacity following retinal degeneration.

To evaluate the regenerative capacity within the MIKK panel, Cab and HdrR, I performed retinal injury experiments, using a glass needle as an instrument to perform those injuries.
0 dph medaka hatchlings were injured on the right retina using a fine glass needle; maintaining the left retina uninjured to use as control. Following this procedure, hatchlings were incubated in 2.5 mM BrdU solution for 3 days, in the case of Cab and HdrR, and for 4 days for MIKK strains IP10-1 and IP15-1; followed by fixation in 4% PFA.

Cryosectioned samples were immunostained for DAPI and BrdU and imaged with a confocal microscope (Figure 6.1, B). Confirming the previous study, BrdU-positive cells were

observed in mainly outer nuclear layer (ONL) and inner nuclear layer (INL) but no cluster of proliferating cells were observed in Cab and IP10-1 strain of medaka (Figure 6.1, D). Strikingly, a cluster of proliferating cells across ONL and INL was observed in HdrR and IP15-1 strain of medaka, indicating that retinal regenerative ability varies among medaka strains.

Given that *sox2* overexpression in the Cab strain has been shown to promote retinal regeneration (Lust & Wittbrodt, 2018), I compared the *sox2* loci across strains to investigate whether genetic variation at these regions could be associated with regenerative capacity (Figure 6.2). As Sox3 has been shown in some models to compensate for Sox2 function (Adikusuma et al., 2017), the *sox3* locus was also examined for potential associations with regenerative capacity. In the *sox3* locus, two distinct genotype groups were observed: one resembling the Cab-like *sox3* promoter and the other more similar to the HdrR-like s*ox3* promoter (Figure 6.2).

As the occurrence of neurogenic clusters which is defined as clustering of more than 4 proliferating cells, was sometimes rare or inconsistent across sections, I quantified regeneration by counting the number of retinal sections in which such clusters were observed. In addition, I evaluated two other features that were observed after the needle injury: the presence of BrdU-positive cells across all retinal layers (ganglion cell layer, inner nuclear layer, and outer nuclear layer), which indicates active regeneration, and the expansion of proliferative cells in the ciliary marginal zone (CMZ).

These regenerative phenotypes were compared across two selected MIKK panel strains (IP15-1 and IP10-1), along with HdrR and Cab. Since HdrR and Cab were intercrossed several generations ago, producing an HdrR-based strain with mixed ancestry, I genotyped this strain for the s*ox3* promoter region to distinguish individuals carrying either Cab-like or HdrR-like alleles. Regenerative features were then evaluated in relation to these genotypes to assess the potential association between *sox2* locus variation and the likelihood of regeneration.

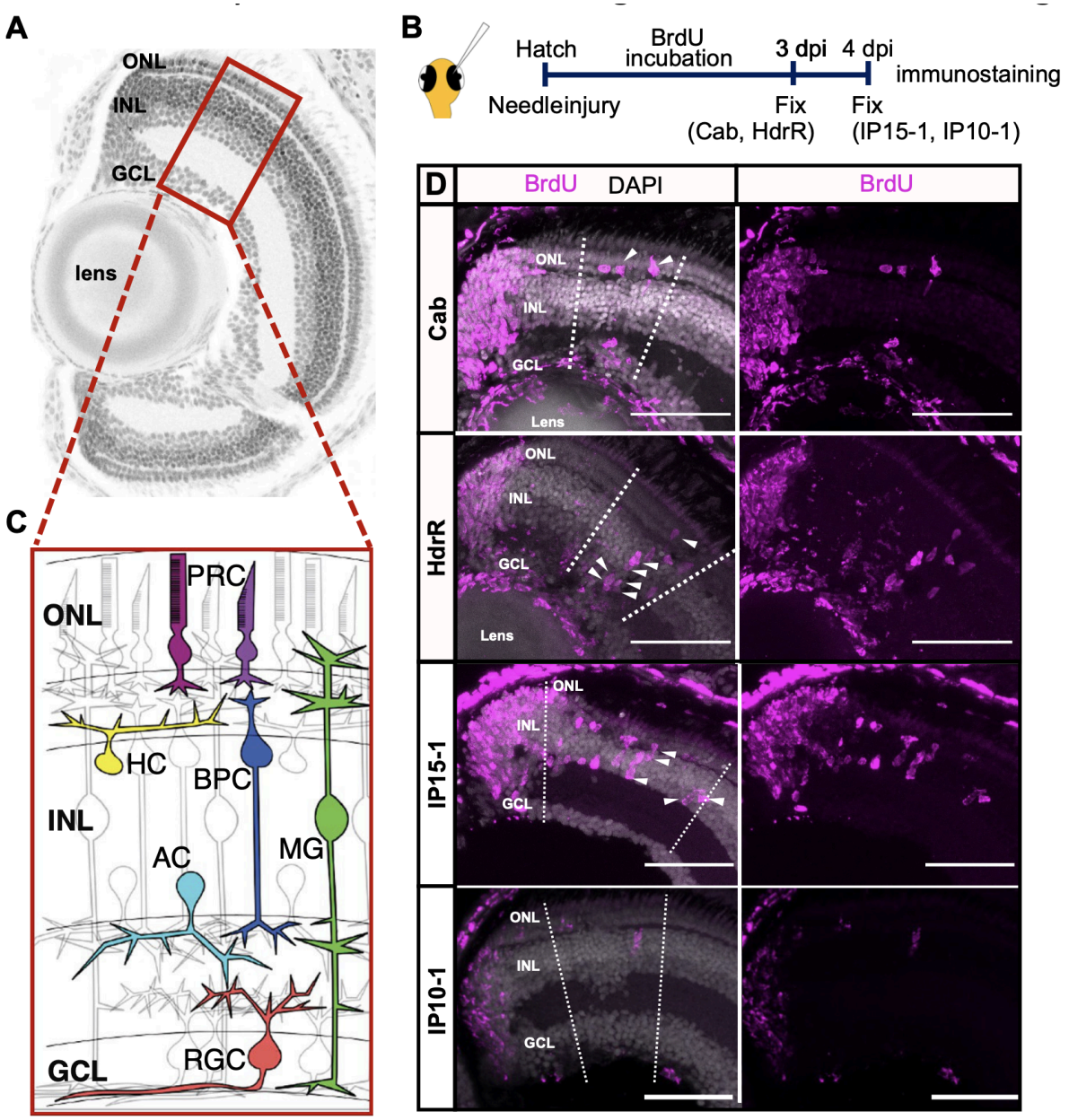


Figure 6.1 The Retinal Regenerative Response Differs Among MIKK Strains.
(A) Cryosection of a medaka retina at hatchling stage. The nuclei are stained with DAPI. Three distinct nuclear layers, the outer nuclear layer (ONL), the inner nuclear layer (INL), and the ganglion cell layer (GCL) are present. (C) Schematic drawing of the cell type composition of the neural retina with six neuronal and one glial cell type. Rod and cone photoreceptors (PRCs) (purple) reside in the ONL, whereas nuclei of Horizontal cells (HCs) (yellow), Amacrine cells (AC) (light blue), Bipolar cells (BCs) (dark blue), and Müller glia (MG) cells (green) are located in the INL. Retinal ganglion cells (RGCs) (red) which are located in the GCL form synapses with ACs and BCs. MG cells span the entire thickness of the retinal column. (modified from Lust, PhD thesis, 2017) (B) Schematic representation of the protocol followed to perform retinal injury. Right retina of the hatchling was stabbed with a fine needle at the hatch. Injured hatchlings were incubated in BrdU solution for 3 to 4 days, then fixed in 4% PFA. Cryosectioned samples were immunostained for DAPI and BrdU and imaged with a confocal microscope. (D) A cluster of proliferating cells across ONL and INL was observed in HdrR and IP15-1 strain of medaka. BrdU-positive cells were observed in ONL and INL but no cluster was observed in Cab and IP10-1 strain of medaka. Scale bar: 50 µm

Neurogenic clusters, defined as localized groups of proliferating cells within the inner nuclear layer (INL), were observed in only 1-2 sections in HdrR and IP15-1, and were absent in IP10-1 as well as in HdrR individuals carrying the Cab-like sox3 locus (Figure 6.3 B). Interestingly, a single hatchling from the Cab strain also exhibited a neurogenic cluster, which was unexpected based on prior report regarding its regenerative profile (Figure 6.3 B).

BrdU-positive cells were detected across all three retinal layers, that are the ganglion cell layer (GCL), inner nuclear layer (INL), and outer nuclear layer (ONL), in all strains examined, suggesting that regeneration occurs broadly following injury (Figure 6.3 C).
An increase in proliferative cells in the ciliary marginal zone (CMZ) was observed in most strains, with the exception of IP10-1, and was particularly prominent in IP15-1 with an HdrR-like genetic background (Figure 6.3 D). This enhanced proliferative response was consistently observed in both HdrR and IP15-1.

Although these findings suggest a potential association between the HdrR background and increased regenerative capacity, this link is not fully explained by the sox3 locus, as both IP10-1 and IP15-1 share the HdrR-like sox3 allele yet differ in their regenerative phenotypes. This data implies that factors such as the speed or origin of proliferative responses may also differ across the strains. Additionally, sampling variability and limitations in injury consistency may have affected detection of some regenerative events.

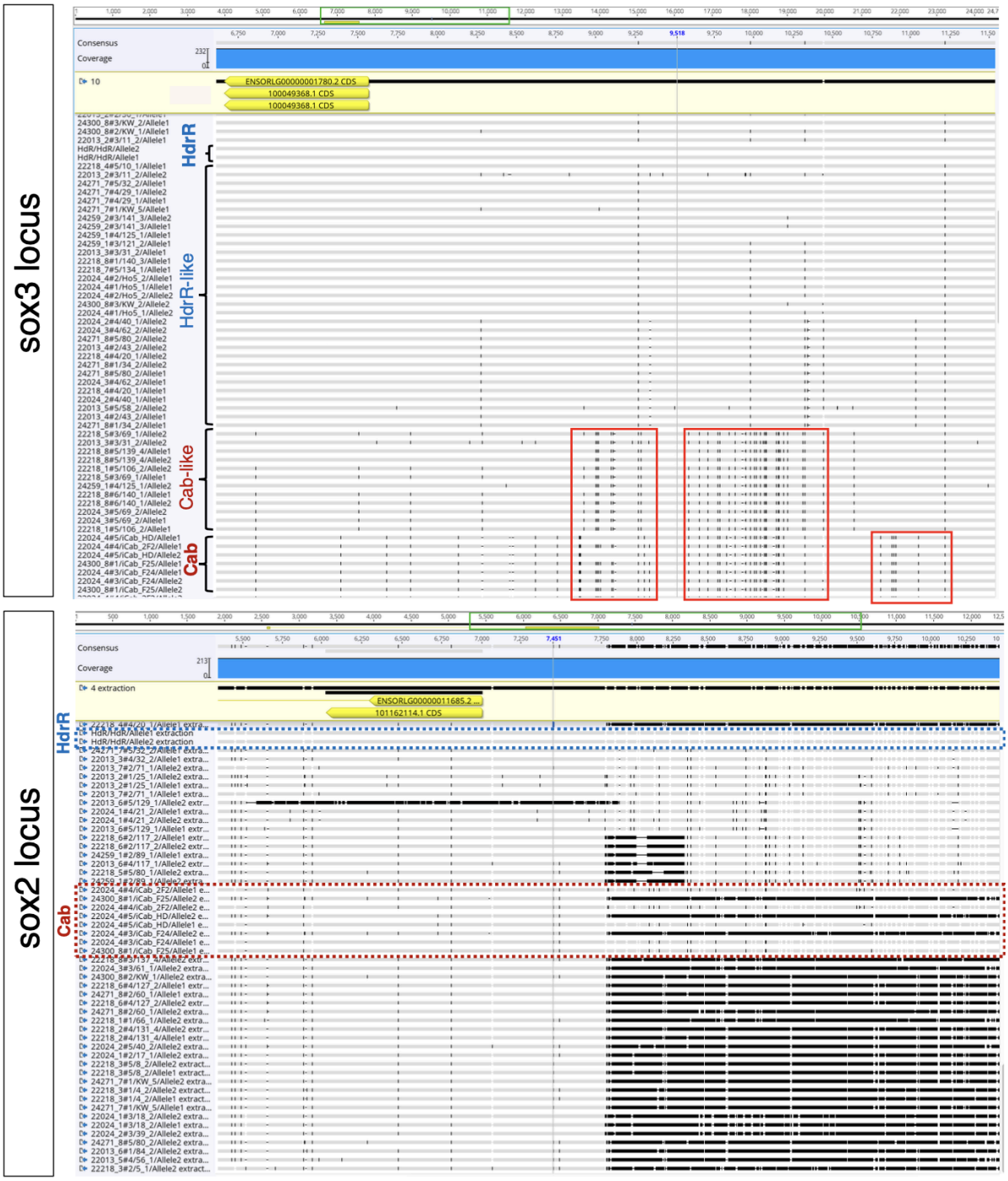


Figure 6.2 Comparison of sox2 and sox3 Locus across MIKK Panel, HdrR and Cab.
Genomic regions corresponding to sox2 (ENSORLG00000011685) and sox3 (ENSORLG00000001780) loci were extracted from the MIKK panel and Cab lines and aligned to the HdrR reference genome. Sequence extraction was performed using a custom script provided by Fanny Defranoux (EBI).

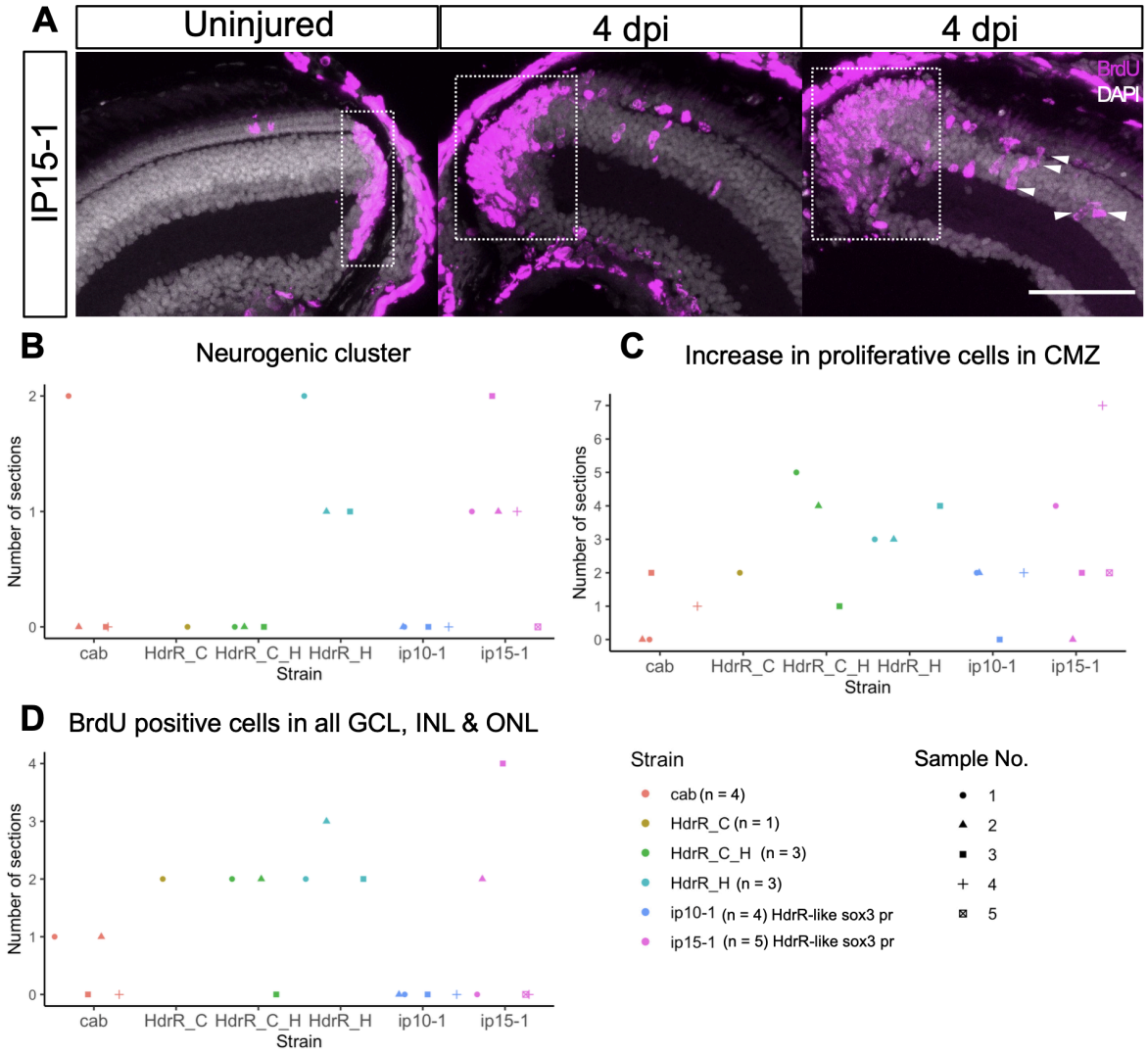


Figure 6.3 Strain Dependent Variation in Retinal Regeneration and Proliferation.
Point plots show the number of retinal sections in which specific regenerative or proliferative cells were observed across different strains. Each point represents a number of sections where the event was observed, with 24-40 sections imaged per strain and 2-10 sections per individual hatchlings.
(A) Representative immunostaining of retinal sections from 4 day post hatch (dph) IP15-1 hatchlings, comparing injured and uninjured eyes. DAPI is shown in gray, and BrdU-positive cells are visualized in magenta. Arrowheads indicate clusters of proliferating cells. The dotted box highlights proliferative cells near the ciliary marginal zone (CMZ). Scale bar: 50µm.(B) Number of sections showing neurogenic clusters, defined as more than four BrdU-positive cells clustered in the inner nuclear layer (INL). (C) Number of sections at which an increase in the size of proliferative cells in the ciliary marginal zone (CMZ) was observed. (D) Number of sections with BrdU-positive cells detected in the ganglion cell layer (GCL), INL, and outer nuclear layer (ONL). HdrR_C: Homozygous for the Cab-like sox3 locus, HdrR_H: Homozygous for the HdrR-like sox3 locus, and HdrR_C_H: Heterozygous, carrying both HdrR-like and Cab-like sox3 loci.

Perspective on Regenerative Ability Differences between Zebrafish and Medaka

The variability in retinal regenerative capacity observed across different strains may be attributed to differences in the expression of key transcription factors such as sox2 and sox3, potentially driven by variation in their promoter regions. These findings of specific genomic sequence patterns existing in certain groups of strains, raise the possibility that regenerative capacity is influenced by divergence in regulatory elements. Mutations accumulated in non-coding regions may lead to differences in chromatin accessibility, potentially driving the strain-specific transcriptional responses to retinal injury. This may help explain the inconsistent regenerative responses observed within the Cab strain.

To better understand the molecular mechanisms behind these differences, multi-omics approaches can be employed. Integrating single-cell RNA sequencing with single-cell ATAC sequencing or spatial transcriptomics, alongside whole-genome sequencing at multiple time points post-injury, would allow for detailed comparisons of gene expression, chromatin accessibility, and genetic variation across strains. These datasets could be analyzed to identify strain-specific transcription factor usage, map open chromatin regions, and determine which cell types or retinal regions exhibit differential regulatory activity. Moreover, given the potential importance of non-coding regions, applying deep learning models to predict changes in chromatin accessibility or to identify functional cis-regulatory elements could be a powerful strategy. Such models could be trained on scATAC-seq and transcriptomic data to infer causative non-coding sequences that drive strain-specific transcriptional programs. Ultimately, this approach may help pinpoint regulatory variants or enhancers responsible for differences in regenerative capacity, thereby providing insight into the genetic architecture underlying retinal repair mechanisms (Sarropoulos et al., 2025).

6.2 Supplementary Figures

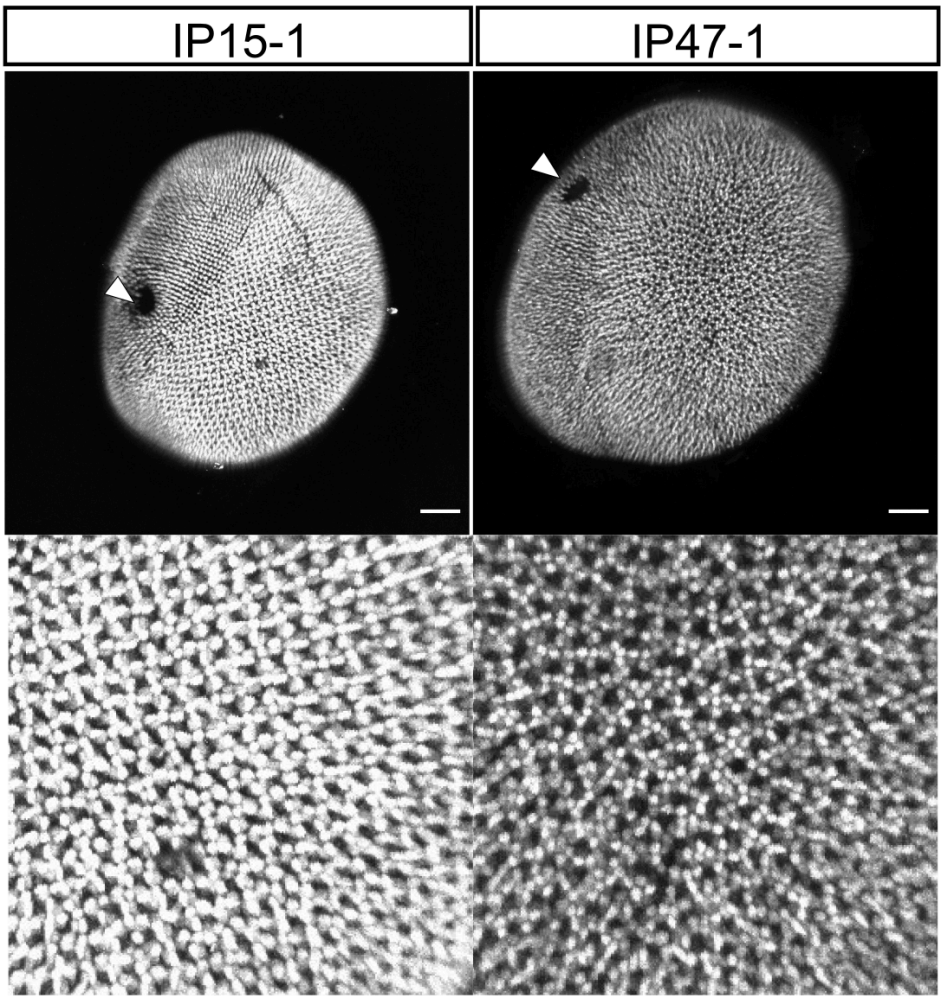


Figure 6.4 Differential Photoreceptor Patterns between MIKK Strains.
A whole retina of IP15-1 and IP47-1 medaka, immunostained with anti-Zpr-1 antibody (cone PRC marker) at 2 dph. Scale bar: 50 μm

Figure 6.5 Minimum Stripe Width Responded and Count of One-directional Swimming Behaviour Across 74 MIKK Panel Strains.

Box plots representing visual and behavioural performance metrics across 74 MIKK panel strains: (A) Minimum stripe width that triggered a measurable response, representing sensitivity to spatial detail, and (B) One-directional swimming trait, reflecting the number of stripe motion phases in which individuals swam consistently in a single direction. Strains are ranked by the median value for each trait. Data were aggregated from multiple technical replicates per strain.

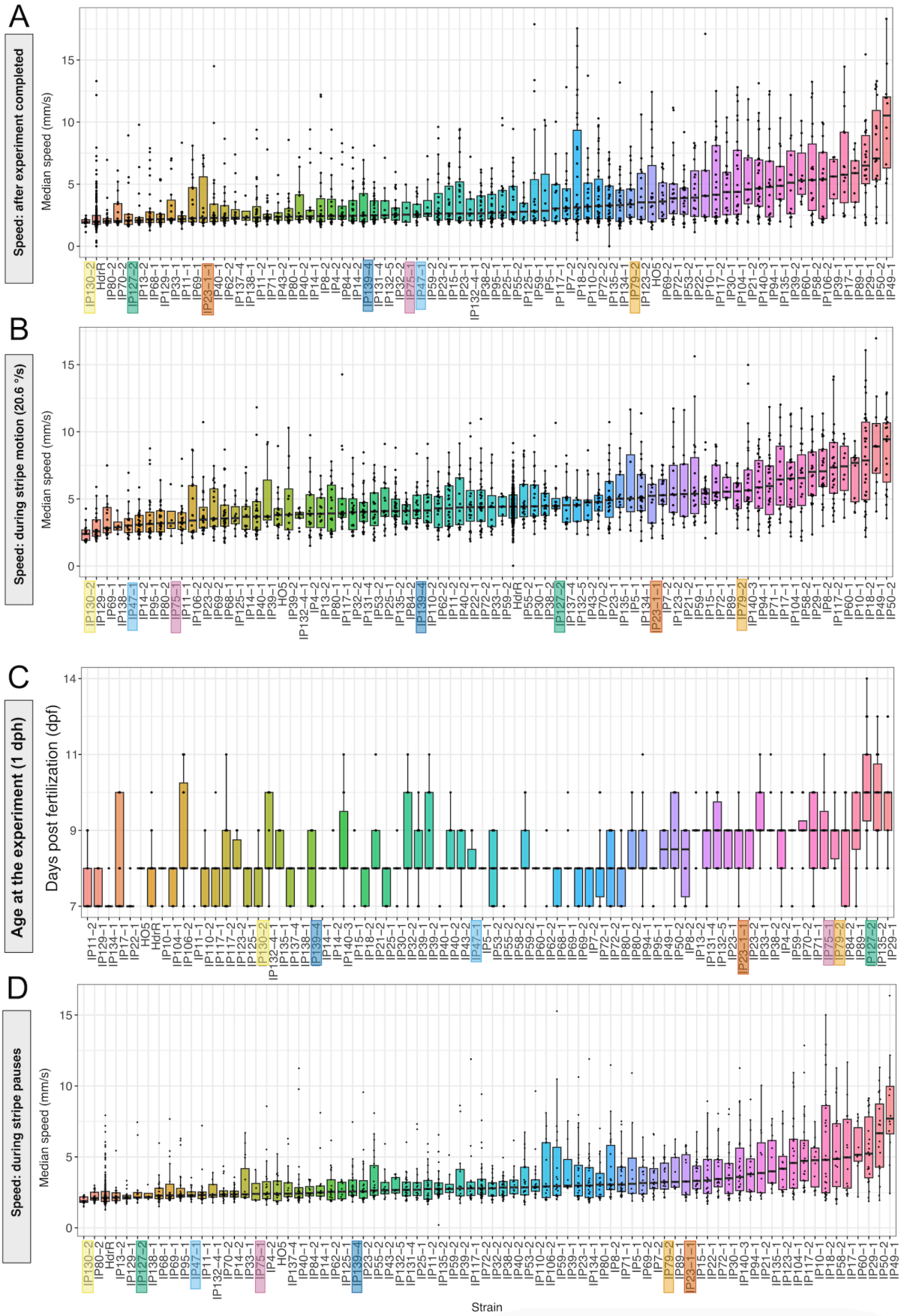
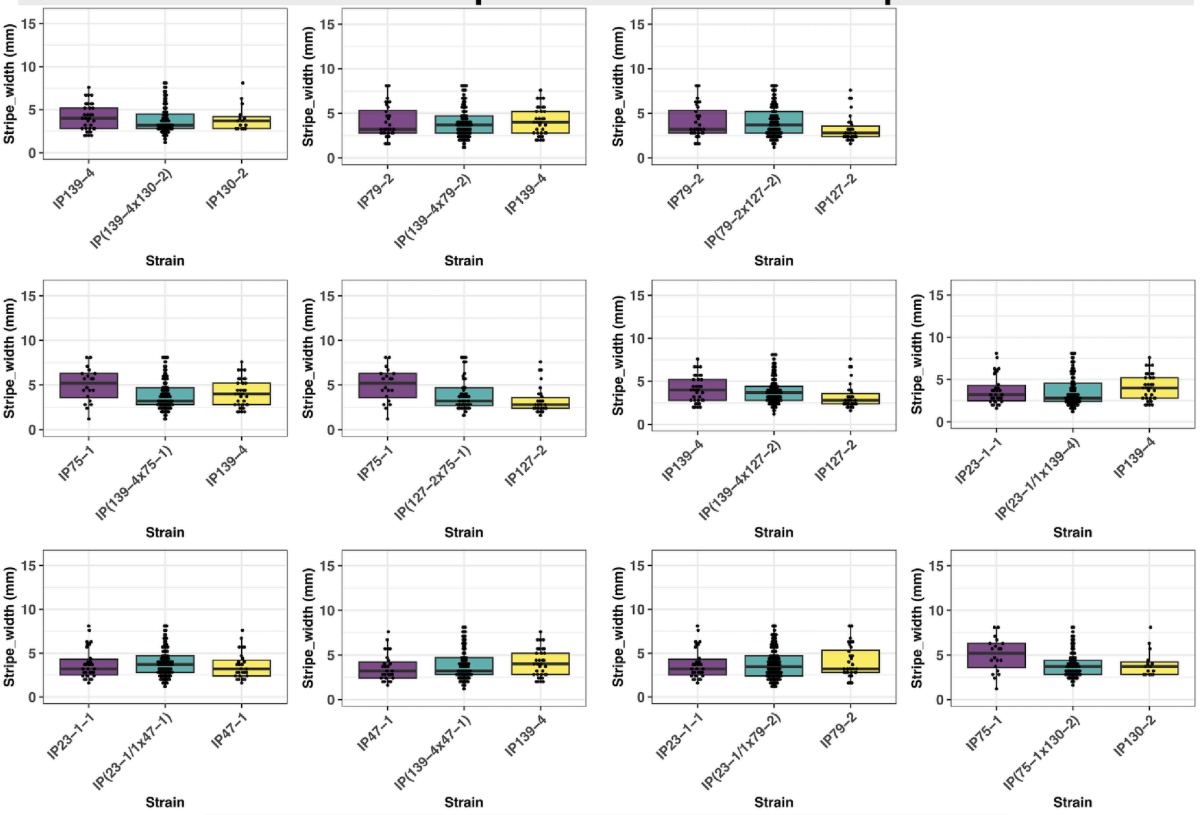

A
Speed: after experiment completed
Median speed (mm/s)
B
Speed: during stripe motion (20.6 °/s)
Median speed (mm/s)
C
Age at the experiment (1 dph)
Days post fertilization (dpf)
D
Speed: during stripe pauses
Median speed (mm/s)
Strain

Figure 6.6 Swimming Speed and Developmental Age of MIKK Panel Strains During and After OMR Assays.

Box plots representing swimming speed across 74 MIKK panel strains under four conditions: during stripe motion at a speed of 20.6°/s (~6.28 mm/s) (B), during pause phases within the experiment (D), and 15 minutes post-experiment (A). (C) Developmental age, expressed as days post-fertilization (dpf) at 1 day post-hatching (dph). Strains are ranked by the median swimming speed for each condition. Data were compiled from multiple technical replicates per strain.

Minimum Stripe Width Elicited a Response



Total distance swam

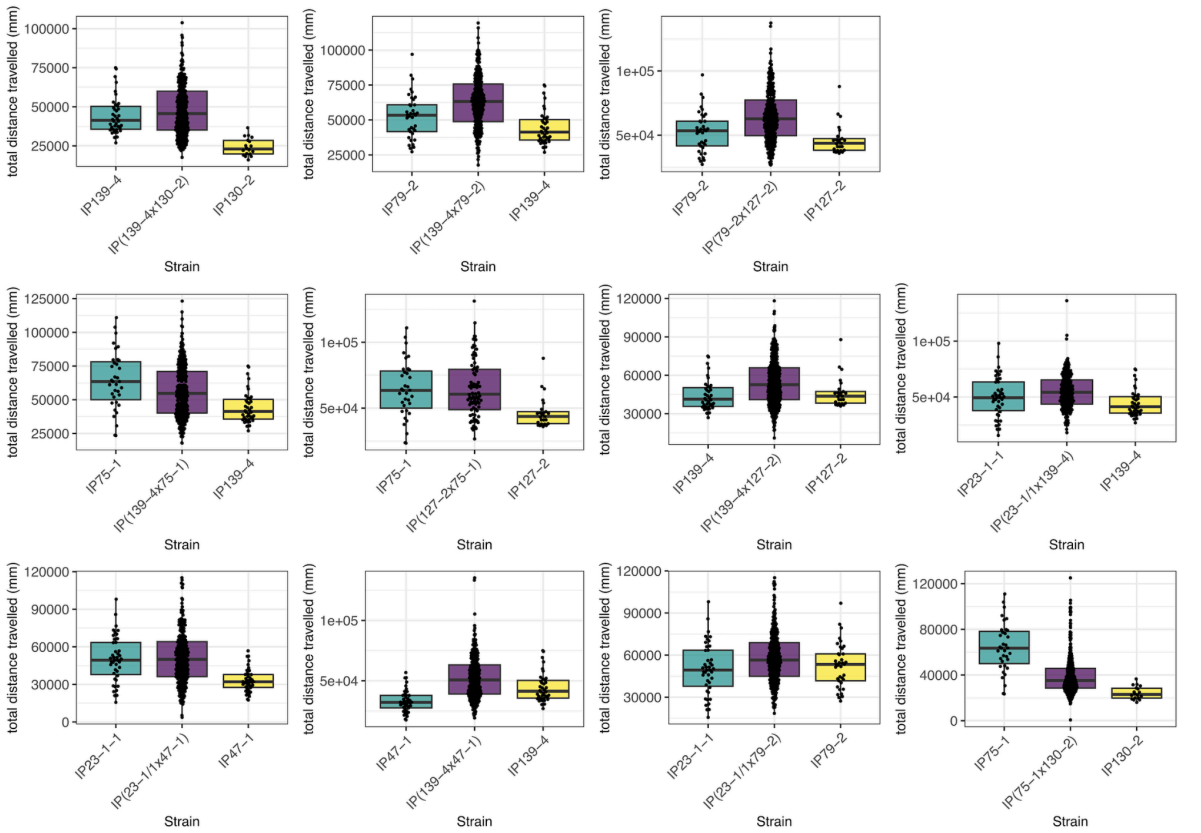


Figure 6.7 Cross-Wise Comparison of Minimum Stripe Width Responded and Total Distance Travelled in F2 Offspring.

Box plots showing minimum stripe width triggering response (upper panel) and total distance swam (lower panel) of F2 hatchlings (purple) and MIKK panel strains (F0, turquoise and yellow) selected for strategic crossing. Every dot corresponds to one individual.

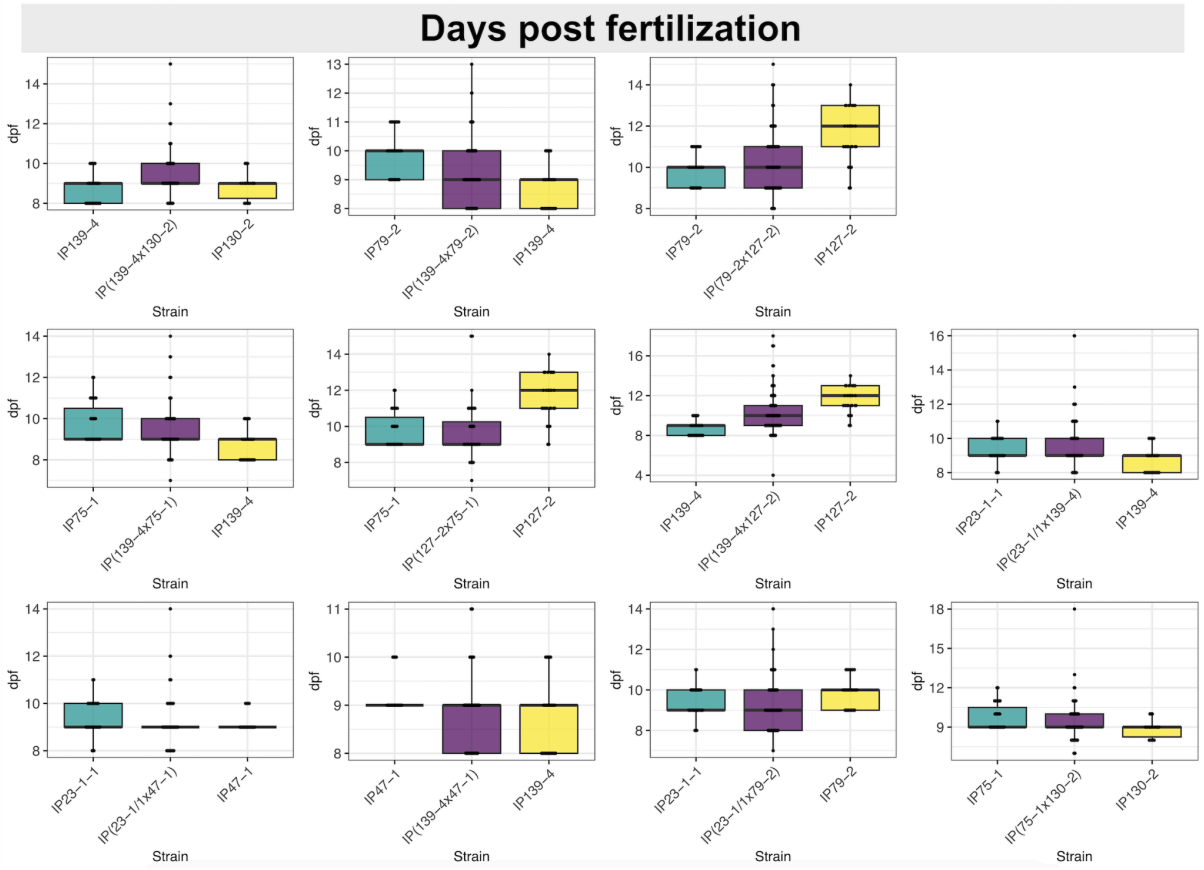


Figure 6.8 Cross-Dependent Comparison of Time Took To Hatch in F2 Offspring.

Box plots showing age in days post fertilization at 1 day post hatch stage of F2 hatchlings (purple) and the founder inbred strains (F0, yellow and turquoise) MIKK panel strains selected for strategic crossing. Every dot corresponds to one individual.

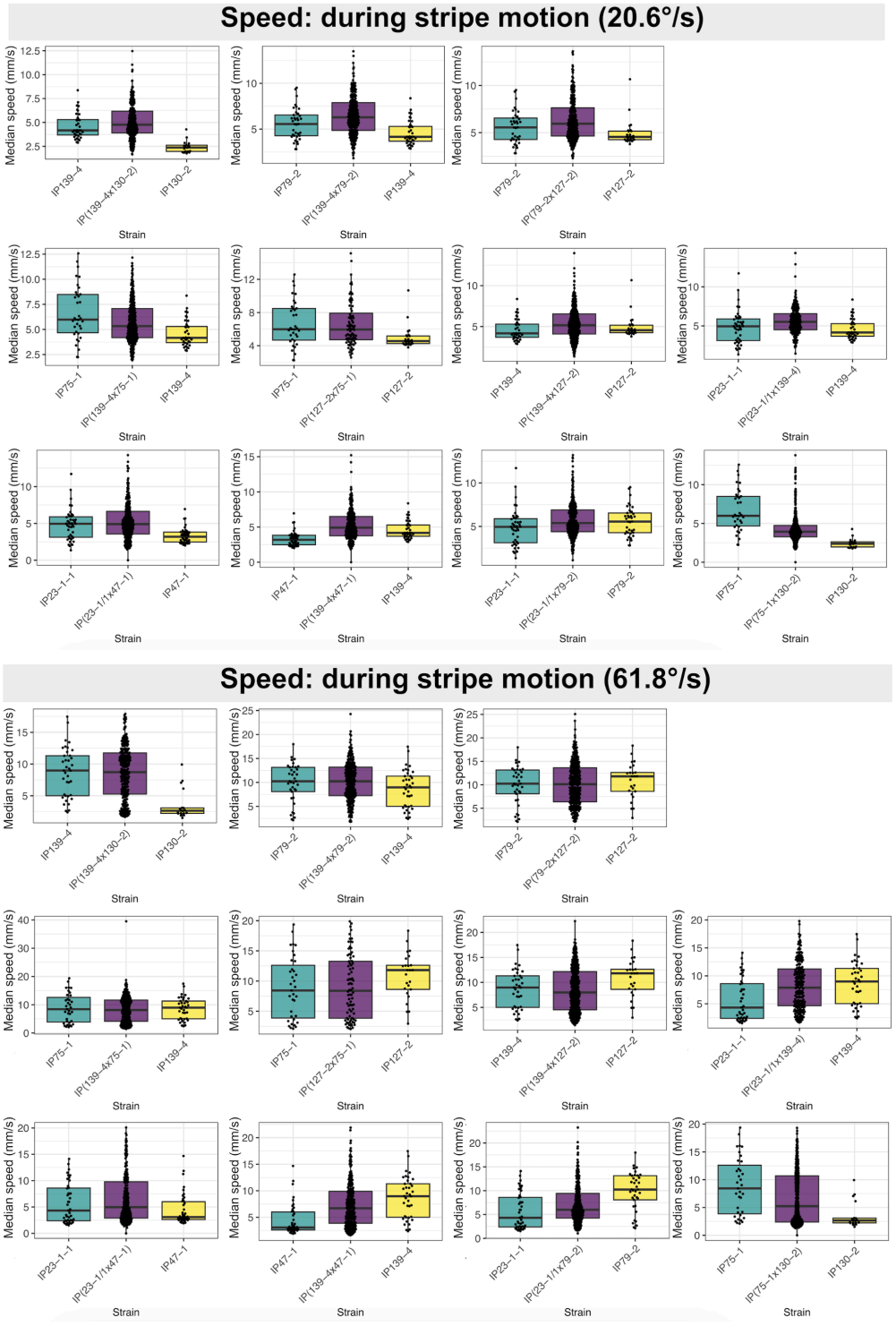


Figure 6.9 Swimming Speed (Stripe in Motion) in F2 Offspring at Two Different Speeds.
Box plots showing swimming speed during stripe in motion (stripe speed at 20.6°/s (approximately 6.28 mm/s) (upper panel) and 61.8°/s (approximately 18.87 mm/s) (lower panel)) of F2 offsprings (purple) and MIKK panel strains (F0, yellow and turquoise) selected for strategic crossing. Every dot corresponds to one individual.

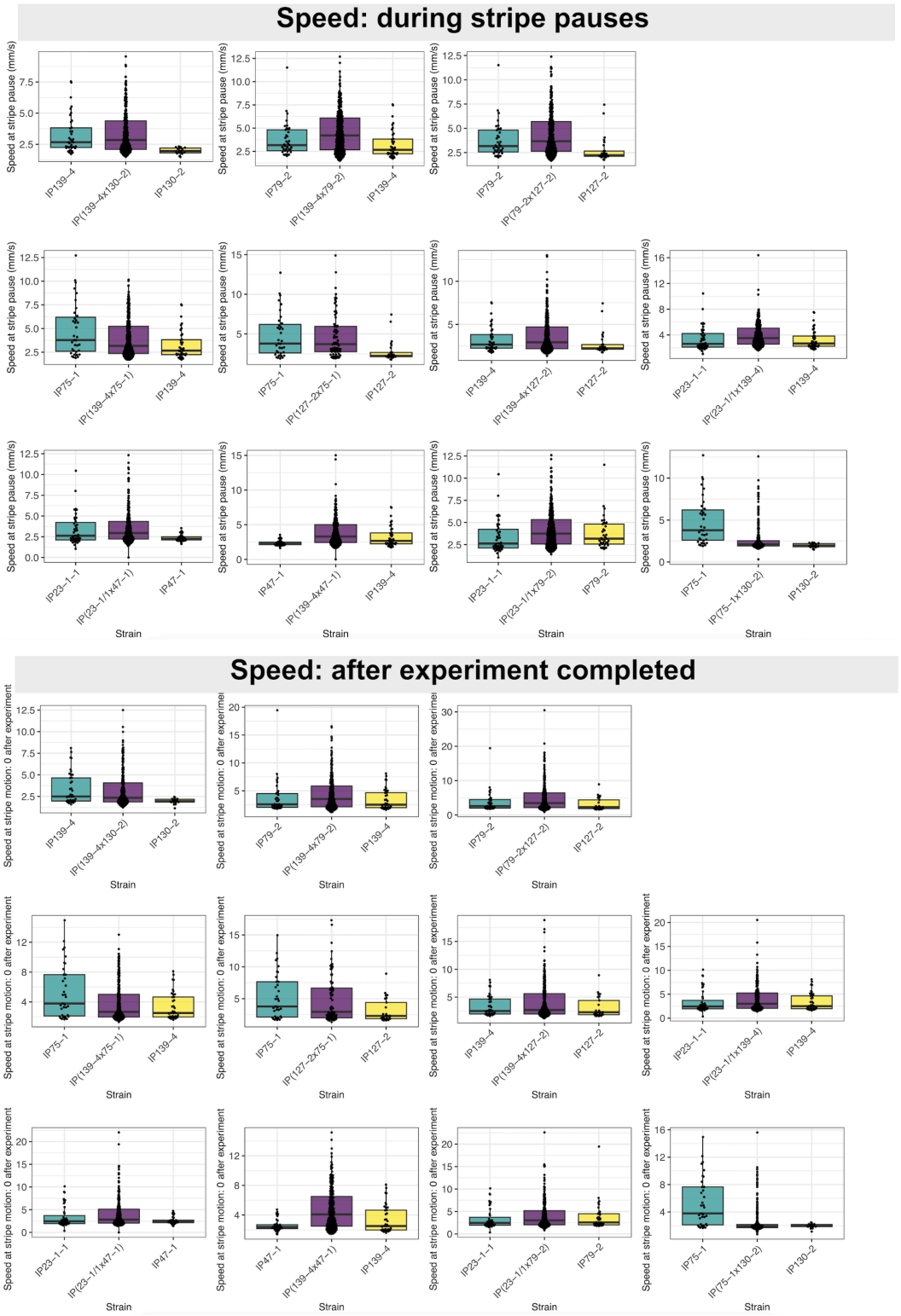


Figure 6.10 Swimming Speed (Stripe at Pause) in F2 Offspring.

Box plots showing swimming speed during stripe at pause (during experiments (upper panel) and during 10 minutes after experiments (lower panel)) of F2 offsprings (purple) and F0 MIKK panel strains (yellow, and turquoise) selected for strategic crossing. Every dot corresponds to one individual.

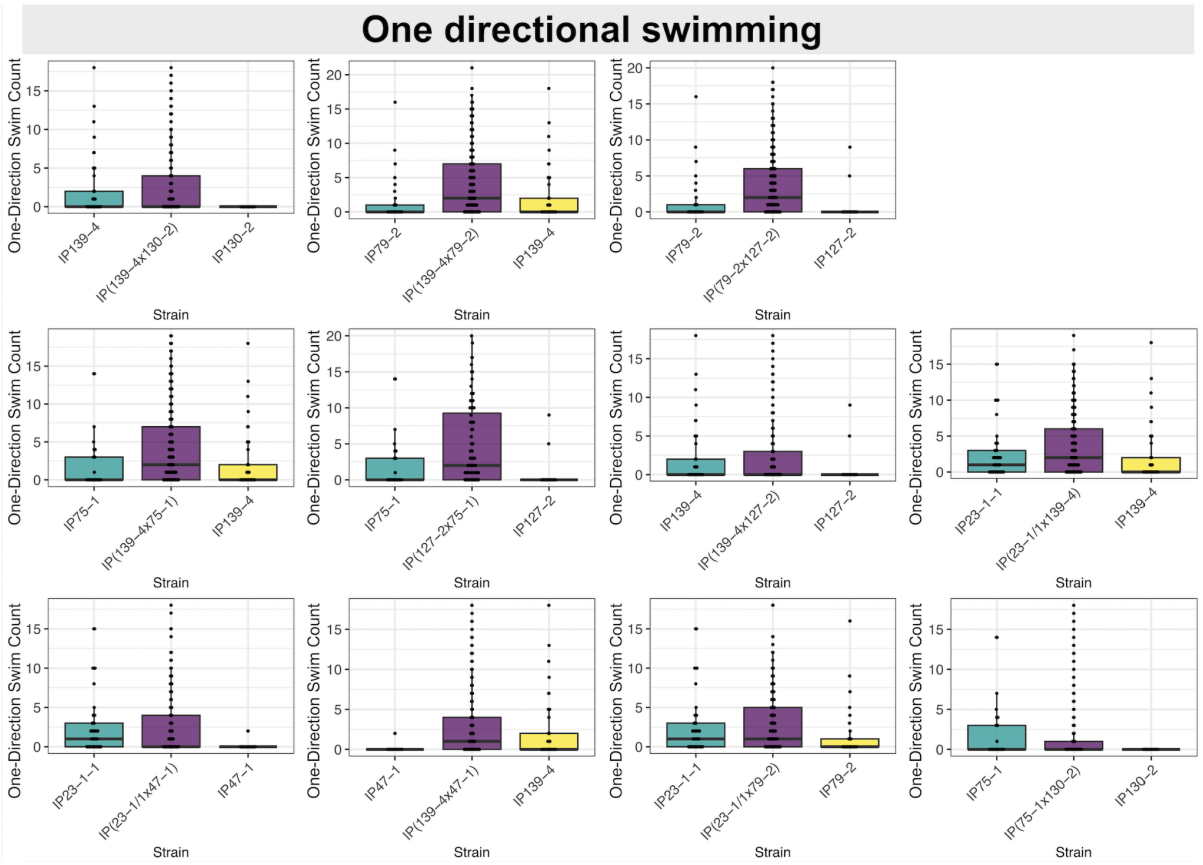


Figure 6.11 One-Directional Swimming Trait in F2 Offspring.

Box plots showing one directional swimming behavior of MIKK panel strains selected for strategic crossing. The occurrence of one directional swimming behavior in F2 hatchlings (purple) and that of the founder inbred strains (F0, yellow and turquoise). Every dot corresponds to one individual.

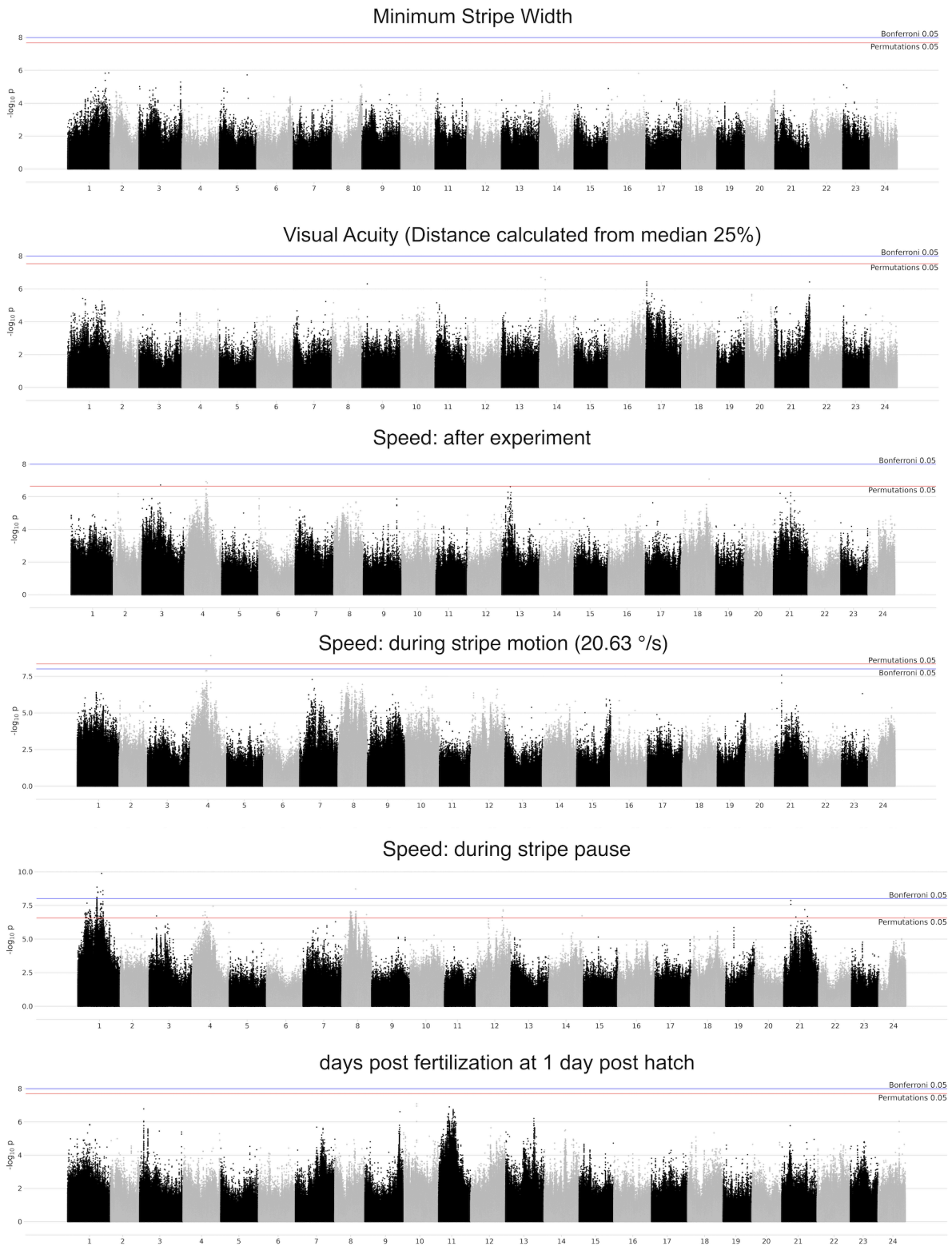


Figure 6.12 Manhattan Plots of QTLs Linked to Behavioral and Developmental Traits.
A statistical association analysis was performed to link genetic variants with visual function phenotypes in the F2 population using a linear mixed model, with the strain of the parent cross and plate ID set as covariates. The phenotypes analyzed included: from the top, the minimum stripe width that elicited a response, visual acuity estimated using the median distance of the top 25% of data points closest to the center, swimming speed during stripe motion (speed of 20.63 °/s), swimming speed during stripe pause, and developmental stage (days post fertilization) at 1 day post hatch. The

significance threshold is indicated in red and was set by 10 permutations. Data produced together with Esther Yoo (EMBL-EBI). Figure adapted from plots generated by Esther Yoo (EMBL-EBI). Significance threshold indicated as red line (α = 0.05), and with the Bonferroni-corrected significance threshold as blue line.

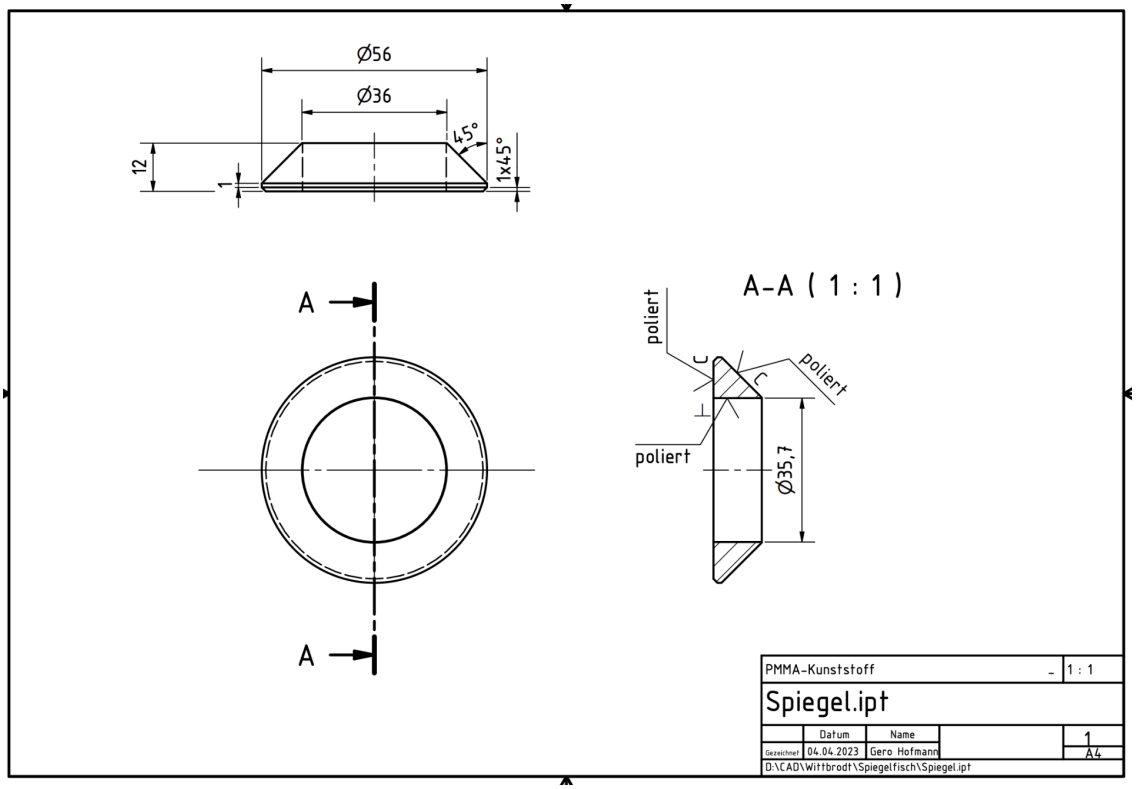


Figure 6.13 Technical drawing of the prisms in the infinity-pool-style OMR assay.

References

Acevedo-Arozena, A., Wells, S., Potter, P., Kelly, M., Cox, R. D., & Brown, S. D. M. (2008). ENU mutagenesis, a way forward to understand gene function. *Annual Review of Genomics and Human Genetics*, *9*(1), 49–69.

Adikusuma, F., Pederick, D., McAninch, D., Hughes, J., & Thomas, P. (2017). Functional equivalence of the SOX2 and SOX3 transcription factors in the developing mouse brain and testes. *Genetics*, *206*(3), 1495–1503.

Agarwal, R., Benjaminsen, J., Lust, K., Becker, C., Fuchs, N., Hasel de Carvalho, E., Eggeler, F., El Said Ibrahim, O., Aghaallaei, N., Bajoghli, B., & Wittbrodt, J. (2025). Immune surveillance and pruning of neuronal stem cells in the medaka retina. In *bioRxiv*. https://doi.org/10.1101/2025.04.01.646524

Ahsanuddin, S., & Wu, A. Y. (2023). Single-cell transcriptomics of the ocular anterior segment: a comprehensive review. *Eye*, *37*(16), 3334–3350.

Asai, T., Senou, H., & Hosoya, K. (2011). Oryzias sakaizumii, a new ricefish from northern Japan (Teleostei: Adrianichthyidae). *Ichthyological Exploration of Freshwaters*, *22*, 289–299.

Ashbrook, D. G., Arends, D., Prins, P., Mulligan, M. K., Roy, S., Williams, E. G., Lutz, C. M., Valenzuela, A., Bohl, C. J., Ingels, J. F., McCarty, M. S., Centeno, A. G., Hager, R., Auwerx, J., Lu, L., & Williams, R. W. (2021). A platform for experimental precision medicine: The extended BXD mouse family. *Cell Systems*, *12*(3), 235–247.e9.

Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., Tarone, A. M., Hu, T. T., Jiang, R., Muliyati, N. W., Zhang, X., Amer, M. A., Baxter, I., Brachi, B., Chory, J., Dean, C., Debieu, M., … Nordborg, M. (2010). Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. *Nature*, *465*(7298), 627–631.

Auer, F., Nardone, K., Matsuda, K., Hibi, M., & Schoppik, D. (2024). Cerebellar Purkinje Cells Control Posture in Larval Zebrafish (Danio rerio). In *eLife* (p. RP97614). https://doi.org/10.7554/elife.97614

Baier, H., & Scott, E. K. (2024). The visual systems of zebrafish. *Annual Review of Neuroscience*, *47*(1), 255–276.

Bert, B., Chmielewska, J., Bergmann, S., Busch, M., Driever, W., Finger-Baier, K., Hößler, J., Köhler, A., Leich, N., Misgeld, T., Nöldner, T., Reiher, A., Schartl, M.,

Seebach-Sproedt, A., Thumberger, T., Schönfelder, G., & Grune, B. (2016). Considerations for a European animal welfare standard to evaluate adverse phenotypes in teleost fish. *The EMBO Journal*, *35*(11), 1151–1154.

Bilotta, J., & Saszik, S. (2001). The zebrafish as a model visual system. *International Journal of Developmental Neuroscience: The Official Journal of the International Society for Developmental Neuroscience*, *19*(7), 621–629.

Bower, M. A., Schimmenti, L. A., & Eccles, M. R. (1993). PAX2-related disorder. In *GeneReviews(®)*. University of Washington, Seattle.

Brastrom, L. K., Scott, C. A., Dawson, D. V., & Slusarski, D. C. (2019). A high-throughput assay for congenital and age-related eye diseases in zebrafish. *Biomedicines*, *7*(2), 28.

Brockerhoff, S. E., Hurley, J. B., Janssen-Bienhold, U., Neuhauss, S. C., Driever, W., & Dowling, J. E. (1995). A behavioral screen for isolating zebrafish mutants with visual system defects. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(23), 10545–10549.

Carl, M., & Wittbrodt, J. (1999). Graded interference with FGF signalling reveals its dorsoventral asymmetry at the mid-hindbrain boundary. *Development (Cambridge, England)*, *126*(24), 5659–5667.

Carvalho, P. S. M., Noltie, D. B., & Tillitt, D. E. (2002). Ontogenetic improvement of visual function in the medaka Oryzias latipes based on an optomotor testing system for larval and adult fish. *Animal Behaviour*, *64*(1), 1–10.

Caso, F., & Davies, B. (2022). Base editing and prime editing in laboratory animals. *Laboratory Animals*, *56*(1), 35–49.

Caves, E. M., Sutton, T. T., & Johnsen, S. (2017). Visual acuity in ray-finned fishes correlates with eye size and habitat. *The Journal of Experimental Biology*, *220*(Pt 9), 1586–1596.

Centanin, L., & Wittbrodt, J. (2014). Retinal neurogenesis. *Development* , *141*(2), 241–244.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), 7.

Chen, T.-C., Huang, D.-S., Lin, C.-W., Yang, C.-H., Yang, C.-M., Wang, V. Y., Lin, J.-W., Luo, A. C., Hu, F.-R., & Chen, P.-L. (2021). Genetic characteristics and epidemiology of inherited retinal degeneration in Taiwan. *Npj Genomic Medicine*, *6*(1), 16.

Chien, S.-E., Yeh, S.-L., Yamashita, W., & Tsujimura, S.-I. (2023). Enhanced human contrast sensitivity with increased stimulation of melanopsin in intrinsically photosensitive retinal ganglion cells. *Vision Research*, *209*(108271), 108271.

Chowdhury, K., Lin, S., & Lai, S.-L. (2022). Comparative study in zebrafish and medaka unravels the mechanisms of tissue regeneration. *Frontiers in Ecology and Evolution*, *10*. https://doi.org/10.3389/fevo.2022.783818

Choy, S., Thakur, S., Polyakov, E., Abdelaziz, J., Lloyd, E., Enriquez, M., Jayan, N., Mensinger, A., Fily, Y., McGaugh, S., Keene, A. C., & Kowalko, J. E. (2025). Mutations in the albinism gene oca2 alter vision-dependent prey capture behavior in the Mexican tetra. *The Journal of Experimental Biology*, *228*(7). https://doi.org/10.1242/jeb.249881

Chrispell, J. D., Rebrik, T. I., & Weiss, E. R. (2015). Electroretinogram analysis of the visual response in zebrafish larvae. *Journal of Visualized Experiments: JoVE*, *97*. https://doi.org/10.3791/52662

Cleymaet, A. M., de Linde Henriksen, M., Pederson, S. L., Sadar, M. J., Johnston, M., Tovar-Lopez, G., Daniels, J., Sharkey, L., & Teixeira, L. B. C. (2022). Pathology in practice. *Journal of the American Veterinary Medical Association*, *259*(S2), 1–4.

Cornean, A., Gierten, J., Welz, B., Mateo, J. L., Thumberger, T., & Wittbrodt, J. (2022). Precise in vivo functional analysis of DNA variants with base editing using ACEofBASEs target prediction. *eLife*, *11*. https://doi.org/10.7554/eLife.72124

Cote, R. H., Gupta, R., Irwin, M. J., & Wang, X. (2022). Photoreceptor phosphodiesterase (PDE6): Structure, regulatory mechanisms, and implications for treatment of retinal diseases. *Advances in Experimental Medicine and Biology*, *1371*, 33–59.

Crouzier, L., Richard, E. M., Diez, C., Denus, M., Peyrel, A., Alzaeem, H., Cubedo, N., Delaunay, T., Maurice, T., & Delprat, B. (2022). NCS1 overexpression restored mitochondrial activity and behavioral alterations in a zebrafish model of Wolfram syndrome. *Molecular Therapy. Methods & Clinical Development*, *27*, 295–308.

Delpero, M., Korkuć, P., Arends, D., Brockmann, G. A., & Hesse, D. (2024). Identification of additional body weight QTLs in the Berlin Fat Mouse BFMI861 lines using time series data. *Scientific Reports*, *14*(1), 6159.

Devoldere, J., Peynshaert, K., De Smedt, S. C., & Remaut, K. (2019). Müller cells as a target for retinal therapy. *Drug Discovery Today*, *24*(8), 1483–1498.

Diacou, R., Nandigrami, P., Fiser, A., Liu, W., Ashery-Padan, R., & Cvekl, A. (2022). Cell fate decisions, transcription factors and signaling during early retinal development. *Progress in Retinal and Eye Research*, *91*(101093), 101093.

Doering, L., Cornean, A., Thumberger, T., Benjaminsen, J., Wittbrodt, B., Kellner, T., Hammouda, O. T., Gorenflo, M., Wittbrodt, J., & Gierten, J. (2023). CRISPR-based knockout and base editing confirm the role of MYRF in heart development and

congenital heart disease. *Disease Models & Mechanisms*, *16*(8). https://doi.org/10.1242/dmm.049811

Dubucs, C., Plaisancié, J., Courtade-Saidi, M., & Damase-Michel, C. (2024). The first review on prenatal drug exposure and ocular malformation occurrence. *Frontiers in Pediatrics*, *12*, 1379875.

Dumont, B. L., Gatti, D. M., Ballinger, M. A., Lin, D., Phifer-Rixey, M., Sheehan, M. J., Suzuki, T. A., Wooldridge, L. K., Frempong, H. O., Lawal, R. A., Churchill, G. A., Lutz, C., Rosenthal, N., White, J. K., & Nachman, M. W. (2024). Into the Wild: A novel wild-derived inbred strain resource expands the genomic and phenotypic diversity of laboratory mouse models. *PLoS Genetics*, *20*(4), e1011228.

Edelstein, A. D., Tsuchida, M. A., Amodaj, N., Pinkard, H., Vale, R. D., & Stuurman, N. (2014). Advanced methods of microscope control using Î¼Manager software. *Journal of Biological Methods*, *1*(2), 1.

Espino-Saldaña, A. E., Rodríguez-Ortiz, R., Pereida-Jaramillo, E., & Martínez-Torres, A. (2020). Modeling neuronal diseases in zebrafish in the era of CRISPR. *Current Neuropharmacology*, *18*(2), 136–152.

Figueroa, D., Ríos, J., Araneda, O. F., Contreras, H. R., Concha, M. L., & García, C. (2022). Oxidative stress parameters and morphological changes in Japanese medaka (Oryzias latipes) after acute exposure to OA-group toxins. *Life (Basel, Switzerland)*, *13*(1), 15.

Fitzgerald, T., Brettell, I., Leger, A., Wolf, N., Kusminski, N., Monahan, J., Barton, C., Herder, C., Aadepu, N., Gierten, J., Becker, C., Hammouda, O. T., Hasel, E., Lischik, C., Lust, K., Sokolova, N., Suzuki, R., Tsingos, E., Tavhelidse, T., … Loosli, F. (2022). The Medaka Inbred Kiyosu-Karlsruhe (MIKK) panel. *Genome Biology*, *23*(1), 59.

Fleisch, V. C., Jametti, T., & Neuhauss, S. C. F. (2008). Electroretinogram (ERG) measurements in larval zebrafish. *CSH Protocols*, *2008*(3), db.prot4973.

Ford, L. M., & Petersen-Jones, S. M. (2025). Modifiers and their impact on inherited retinal diseases: a review. *Ophthalmic Genetics*, *46*(1), 1–14.

Forrester, J. V., Dick, A. D., McMenamin, P. G., Roberts, F., & Pearlman, E. (2020). *The eye* (5th ed.). Elsevier Health Sciences.

Furukawa, R., & Ijiri, K. (2002). Swimming behavior of larval Medaka fish under microgravity. *Advances in Space Research: The Official Journal of the Committee on Space Research* , *30*(4), 733–738.

Furutani-Seiki, M., Sasado, T., Morinaga, C., Suwa, H., Niwa, K., Yoda, H., Deguchi, T., Hirose, Y., Yasuoka, A., Henrich, T., Watanabe, T., Iwanami, N., Kitagawa, D., Saito,

K., Asaka, S., Osakada, M., Kunimatsu, S., Momoi, A., Elmasri, H., … Kondoh, H. (2004). A systematic genome-wide screen for mutations affecting organogenesis in Medaka, Oryzias latipes. *Mechanisms of Development*, *121*(7-8), 647–658.

Gierten, J., Welz, B., Fitzgerald, T., Thumberger, T., Agarwal, R., Hummel, O., Leger, A., Weber, P., Naruse, K., Hassel, D., Hübner, N., Birney, E., & Wittbrodt, J. (2025). Natural genetic variation quantitatively regulates heart rate and dimension. *Nature Communications*, *16*(1), 4062.

Gollisch, T., & Meister, M. (2010). Eye smarter than scientists believed: neural computations in circuits of the retina. *Neuron*, *65*(2), 150–164.

Gücüm, S., Sakson, R., Hoffmann, M., Grote, V., Becker, C., Pakari, K., Beedgen, L., Thiel, C., Rapp, E., Ruppert, T., Thumberger, T., & Wittbrodt, J. (2021). A patient-based medaka alg2 mutant as a model for hypo-N-glycosylation. *Development* , *148*(11). https://doi.org/10.1242/dev.199385

Gutierrez-Triana, J. A., Tavhelidse, T., Thumberger, T., Thomas, I., Wittbrodt, B., Kellner, T., Anlas, K., Tsingos, E., & Wittbrodt, J. (2018). Efficient single-copy HDR by 5’ modified long dsDNA donors. *eLife*, *7*. https://doi.org/10.7554/eLife.39468

Haffter, P., Granato, M., Brand, M., Mullins, M. C., Hammerschmidt, M., Kane, D. A., Odenthal, J., van Eeden, F. J., Jiang, Y. J., Heisenberg, C. P., Kelsh, R. N., Furutani-Seiki, M., Vogelsang, E., Beuchle, D., Schach, U., Fabian, C., & Nüsslein-Volhard, C. (1996). The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. *Development (Cambridge, England)*, *123*, 1–36.

Hanany, M., Rivolta, C., & Sharon, D. (2020). Worldwide carrier frequency and genetic prevalence of autosomal recessive inherited retinal diseases. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(5), 2710–2716.

Hanssen, F., Garcia, M. U., Folkersen, L., Pedersen, A. S., Lescai, F., Jodoin, S., Miller, E., Seybold, M., Wacker, O., Smith, N., Gabernet, G., & Nahnsen, S. (2024). Scalable and efficient DNA sequencing analysis on different compute infrastructures aiding variant discovery. *NAR Genomics and Bioinformatics*, *6*(2), lqae031.

Harding, P., & Moosajee, M. (2019). The molecular basis of human anophthalmia and microphthalmia. *Journal of Developmental Biology*, *7*(3), 16.

Haug, P., Koller, S., Maggi, J., Lang, E., Feil, S., Wlodarczyk, A., Bähr, L., Steindl, K., Rohrbach, M., Gerth-Kahlert, C., & Berger, W. (2021). Whole Exome Sequencing in Coloboma/Microphthalmia: Identification of Novel and Recurrent Variants in Seven

Genes. *Genes*, *12*(1). https://doi.org/10.3390/genes12010065

Heap, L. A., Vanwalleghem, G. C., Thompson, A. W., Favre-Bulle, I., Rubinsztein-Dunlop, H., & Scott, E. K. (2017). Hypothalamic projections to the optic tectum in larval zebrafish. *Frontiers in Neuroanatomy*, *11*, 135.

Heermann, S., Schütz, L., Lemke, S., Krieglstein, K., & Wittbrodt, J. (2015). Eye morphogenesis driven by epithelial flow into the optic cup facilitated by modulation of bone morphogenetic protein. *eLife*, *4*. https://doi.org/10.7554/eLife.05216

Hilgers, L., & Schwarzer, J. (2019). The untapped potential of medaka and its wild relatives. *eLife*, *8*. https://doi.org/10.7554/eLife.46994

Howe, K. L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Azov, A. G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., Davidson, C., Dodiya, K., El Houdaigui, B., Fatima, R., … Flicek, P. (2021). Ensembl 2021. *Nucleic Acids Research*, *49*(D1), D884–D891.

Hyodo-Taguchi, Y., & Egami, N. (1985). Establishment of Inbred Strains of the Medaka Oryzias latipes and the Usefulness of the Strains for Biomedical Research. *Zoological Science*, *2*, 305–316.

Imada, H., Hoki, M., Suehiro, Y., Okuyama, T., Kurabayashi, D., Shimada, A., Naruse, K., Takeda, H., Kubo, T., & Takeuchi, H. (2010). Coordinated and cohesive movement of two small conspecific fish induced by eliciting a simultaneous optomotor response. *PloS One*, *5*(6), e11248.

Ishihara, M. (1916). Medaka no taishoku no iden ni tsuite [On the inheritance of body color of medaka] (translated by author). *Fukuoka Ikadaigaku Zashi*, *9*, 259–267.

Ishikawa, C. (1912). Genshu kairyo ron [The theory of the improvement of original species] (title translated by author). *Suisan Koshujo Publication*, 104.

Iwamatsu, T. (2004). Stages of normal development in the medaka Oryzias latipes. *Mechanisms of Development*, *121*(7-8), 605–618.

Iwamatsu, T. (2018). *Medakagaku zensho (The integrated book for the biology of the medaka), Completely revised and enlarged version*. Daigaku Kyoiku Shuppan.

Kajiwara, K., Berson, E. L., & Dryja, T. P. (1994). Digenic retinitis pigmentosa due to mutations at the unlinked peripherin/RDS and ROM1 loci. *Science (New York, N.Y.)*, *264*(5165), 1604–1608.

Kasahara, M., Naruse, K., Sasaki, S., Nakatani, Y., Qu, W., Ahsan, B., Yamada, T., Nagayasu, Y., Doi, K., Kasai, Y., Jindo, T., Kobayashi, D., Shimada, A., Toyoda, A., Kuroki, Y., Fujiyama, A., Sasaki, T., Shimizu, A., Asakawa, S., … Kohara, Y. (2007). The medaka

draft genome and insights into vertebrate genome evolution. *Nature*, *447*(7145), 714–719.

Kirchmaier, S., Naruse, K., Wittbrodt, J., & Loosli, F. (2015). The genomic and genetic toolbox of the teleost medaka (Oryzias latipes). *Genetics*, *199*(4), 905–918.

Knoll, A. T., Jiang, K., & Levitt, P. (2018). Quantitative trait locus mapping and analysis of heritable variation in affiliative social behavior and co-occurring traits. *Genes, Brain, and Behavior*, *17*(5), e12431.

Kohno, H., Chen, Y., Kevany, B. M., Pearlman, E., Miyagi, M., Maeda, T., Palczewski, K., & Maeda, A. (2013). Photoreceptor proteins initiate microglial activation via Toll-like receptor 4 in retinal degeneration mediated by all-trans-retinal. *The Journal of Biological Chemistry*, *288*(21), 15326–15341.

Koornneef, M., & Meinke, D. (2010). The development of Arabidopsis as a model plant. *The Plant Journal: For Cell and Molecular Biology*, *61*(6), 909–921.

Larhammar, D., Nordström, K., & Larsson, T. A. (2009). Evolution of vertebrate rod and cone phototransduction genes. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1531), 2867–2880.

Lee, K. S., Choi, Y. J., Cho, J., Lee, H., Lee, H., Park, S. J., Park, J. S., & Hong, Y. C. (2021). Environmental and genetic risk factors of congenital anomalies: An umbrella review of systematic reviews and meta-analyses. *Journal of Korean Medical Science*, *36*(28), e183.

Leger, A., Brettell, I., Monahan, J., Barton, C., Wolf, N., Kusminski, N., Herder, C., Aadepu, N., Becker, C., Gierten, J., Hammouda, O. T., Hasel, E., Lischik, C., Lust, K., Sokolova, N., Suzuki, R., Tavhelidse, T., Thumberger, T., Tsingos, E., … Fitzgerald, T. (2022). Genomic variations and epigenomic landscape of the Medaka Inbred Kiyosu-Karlsruhe (MIKK) panel. *Genome Biology*, *23*(1), 58.

Le, N.-Q., He, W., & MacGregor, S. (2024). Polygenic risk scores and genetically complex eye disease. *Annual Review of Vision Science*, *10*(1), 403–423.

Le, Q. V., Isbell, L. A., Matsumoto, J., Nguyen, M., Hori, E., Maior, R. S., Tomaz, C., Tran, A. H., Ono, T., & Nishijo, H. (2013). Pulvinar neurons reveal neurobiological evidence of past selection for rapid detection of snakes. *Proc Natl Acad Sci U S A*, *110*, 19000–19005.

Lewis, E. B., & Bacher, F. (1968). Methods feeding ethyl methane sulphonate (EMS) Drosophila males. *Drosoph. Inf. Serv*, *43*.

Lewis, T. R., Makia, M. S., Castillo, C. M., Hao, Y., Al-Ubaidi, M. R., Skiba, N. P., Conley,

S. M., Arshavsky, V. Y., & Naash, M. I. (2023). ROM1 is redundant to PRPH2 as a molecular building block of photoreceptor disc rims. In *bioRxivorg*. https://doi.org/10.1101/2023.07.02.547380

Life of my teacher tatsuo AIDA. (2020).

Loosli, F., Del Bene, F., Quiring, R., Rembold, M., Martinez-Morales, J.-R., Carl, M., Grabher, C., Iquel, C., Krone, A., Wittbrodt, B., Winkler, S., Sasado, T., Morinaga, C., Suwa, H., Niwa, K., Henrich, T., Deguchi, T., Hirose, Y., Iwanami, N., … Wittbrodt, J. (2004). Mutations affecting retina development in Medaka. *Mechanisms of Development*, *121*(7-8), 703–714.

Loosli, F., Staub, W., Finger-Baier, K. C., Ober, E. A., Verkade, H., Wittbrodt, J., & Baier, H. (2003). Loss of eyes in zebrafish caused by mutation of chokh/rx3. *EMBO Reports*, *4*(9), 894–899.

Lust, K., & Wittbrodt, J. (2018). Activating the regenerative potential of Müller glia cells in a regeneration-deficient retina. *eLife*, *7*, e32319.

Mackay, T. F. C. (2004). The genetic architecture of quantitative traits: lessons from Drosophila. *Current Opinion in Genetics & Development*, *14*(3), 253–257.

Mackay, T. F. C., & Anholt, R. R. H. (2022). Gregor Mendel's legacy in quantitative genetics. *PLoS Biology*, *20*(7), e3001692.

Mackay, T. F. C., & Anholt, R. R. H. (2024). Pleiotropy, epistasis and the genetic architecture of quantitative traits. *Nature Reviews. Genetics*, *25*(9), 639–657.

Mackay, T. F. C., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Zhu, D., Casillas, S., Han, Y., Magwire, M. M., Cridland, J. M., Richardson, M. F., Anholt, R. R. H., Barrón, M., Bess, C., Blankenburg, K. P., Carbone, M. A., Castellano, D., Chaboub, L., Duncan, L., … Gibbs, R. A. (2012). The Drosophila melanogaster Genetic Reference Panel. *Nature*, *482*(7384), 173–178.

Maillet, C., Guilbaud, L., Monier, I., Khoshnood, B., Quoc, E. B., Dugas, A., Lelong, N., & Jouannic, J.-M. (2024). Prevalence and prenatal diagnosis of congenital eye anomalies: A population-based study. *BJOG: An International Journal of Obstetrics and Gynaecology*, *131*(10), 1385–1391.

Makhankov, Y. V., Rinner, O., & Neuhauss, S. C. F. (2004). An inexpensive device for non-invasive electroretinography in small aquatic vertebrates. *Journal of Neuroscience Methods*, *135*(1-2), 205–210.

Mata, N. L., Weng, J., & Travis, G. H. (2000). Biosynthesis of a major lipofuscin fluorophore in mice and humans with ABCR-mediated retinal and macular degeneration.

Proceedings of the National Academy of Sciences of the United States of America, *97*(13), 7154–7159.

Mathers, P. H., Grinberg, A., Mahon, K. A., & Jamrich, M. (1997). The Rx homeobox gene is essential for vertebrate eye development. *Nature*, *387*(6633), 603–607.

Matsuda, K., & Kubo, F. (2021). Circuit Organization Underlying Optic Flow Processing in Zebrafish. *Frontiers in Neural Circuits*, *15*, 709048.

Matsumoto, Y., Fukamachi, S., Mitani, H., & Kawamura, S. (2006). Functional characterization of visual opsin repertoire in Medaka (Oryzias latipes). *Gene*, *371*(2), 268–278.

Matsuo, M., Ando, Y., Kamei, Y., & Fukamachi, S. (2018). A semi-automatic and quantitative method to evaluate behavioral photosensitivity in animals based on the optomotor response (OMR). *Biology Open*, *7*(6), bio033175.

Mikula Mrstakova, S., & Kozmik, Z. (2024). Genetic analysis of medaka fish illuminates conserved and divergent roles of Pax6 in vertebrate eye development. *Frontiers in Cell and Developmental Biology*, *12*, 1448773.

Milner, S. G., Maccaferri, M., Huang, B. E., Mantovani, P., Massi, A., Frascaroli, E., Tuberosa, R., & Salvi, S. (2016). A multiparental cross population for mapping QTL for agronomic traits in durum wheat (Triticum turgidum ssp. durum). *Plant Biotechnology Journal*, *14*(2), 735–748.

Mizoguchi, K., Sato, M., Saito, R., Koshikuni, M., Sakakibara, M., Manabe, R., Harada, Y., Uchikawa, T., Ansai, S., Kamei, Y., Naruse, K., & Fukamachi, S. (2023). Behavioral photosensitivity of multi-color-blind medaka: enhanced response under ultraviolet light in the absence of short-wavelength-sensitive opsins. *BMC Neuroscience*, *24*(1), 67.

Moshkovitz, A., Lev, M., & Polat, U. (2024). Crowding under scotopic and photopic vision in albino and normal-sighted participants. *Scientific Reports*, *14*(1), 8234.

Mueller, K. P., & Neuhauss, S. C. F. (2010). Quantitative measurements of the optokinetic response in adult fish. *Journal of Neuroscience Methods*, *186*(1), 29–34.

Mueller, K. P., Schnaedelbach, O. D. R., Russig, H. D., & Neuhauss, S. C. F. (2011). VisioTracker, an innovative automated approach to oculomotor analysis. *Journal of Visualized Experiments: JoVE*, *56*. https://doi.org/10.3791/3556

Müller, K. (2011). *Visual Behavior of Zebrafish*. University of Zurich.

Nagashima, M., Barthel, L. K., & Raymond, P. A. (2013). A self-renewing division of zebrafish Müller glial cells generates neuronal progenitors that require N-cadherin to regenerate retinal neurons. *Development (Cambridge, England)*, *140*(22), 4510–4521.

Neuhauss, S. C., Biehlmaier, O., Seeliger, M. W., Das, T., Kohler, K., Harris, W. A., & Baier, H. (1999). Genetic disorders of vision revealed by a behavioral screen of 400 essential loci in zebrafish. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *19*(19), 8603–8615.

Neuhauss, S. C. F. (2010). Zebrafish vision. In *Fish Physiology* (Vol. 29, pp. 81–122). Elsevier.

Nevin, L. M., Robles, E., Baier, H., & Scott, E. K. (2010). Focusing on optic tectum circuitry through the lens of genetics. *BMC Biology*, *8*(1), 126.

Niklaus, S., Glasauer, S. M. K., Kovermann, P., Farshori, K. F., Cadetti, L., Früh, S., Rieser, N. N., Gesemann, M., Zang, J., Fahlke, C., & Neuhauss, S. C. F. (2024). Glutamate transporters are involved in direct inhibitory synaptic transmission in the vertebrate retina. *Open Biology*, *14*(7), 240140.

Nonarath, H. J. T., Simpson, S. L., Slobodianuk, T. L., Tran, H., Collery, R. F., Dinculescu, A., & Link, B. A. (2025). The USH3A causative gene clarin1 functions in Müller glia to maintain retinal photoreceptors. *PLoS Genetics*, *21*(3), e1011205.

Nüsslein-Volhard, C., & Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. *Nature*, *287*(5785), 795–801.

Pakari, K., Wittbrodt, J., & Thumberger, T. (2022). The inceptionist's guide to base editing – *de novo* PAM generation to reach initially inaccessible target-sites. In *bioRxiv*. https://doi.org/10.1101/2022.07.07.499158

Pakari, K., Wittbrodt, J., & Thumberger, T. (2023). De novo PAM generation to reach initially inaccessible target sites for base editing. *Development (Cambridge, England)*, *150*(2). https://doi.org/10.1242/dev.201115

Parker, M. O., Gaviria, J., Haigh, A., Millington, M. E., Brown, V. J., Combe, F. J., & Brennan, C. H. (2012). Discrimination reversal and attentional sets in zebrafish (Danio rerio). *Behavioural Brain Research*, *232*(1), 264–268.

Pierotti, S., Fitzgerald, T., & Birney, E. (2024). FlexLMM: a Nextflow linear mixed model framework for GWAS. In *arXiv [q-bio.GN]*. arXiv. http://arxiv.org/abs/2410.01533

Pierotti, S., Welz, B., Osuna-López, M., Fitzgerald, T., Wittbrodt, J., & Birney, E. (2024). Genotype imputation in F2 crosses of inbred lines. *Bioinformatics Advances*, *4*(1), vbae107.

Poort, J., Wilmes, K. A., Blot, A., Chadwick, A., Sahani, M., Clopath, C., Mrsic-Flogel, T. D., Hofer, S. B., & Khan, A. G. (2022). Learning and attention increase visual response selectivity through distinct mechanisms. *Neuron*, *110*(4), 686–697.e6.

Querubin, A., Lee, H. R., Provis, J. M., & O'Brien, K. M. B. (2009). Photoreceptor and ganglion cell topographies correlate with information convergence and high acuity regions in the adult pigeon (Columba livia) retina. *The Journal of Comparative Neurology*, *517*(5), 711–722.

Radu, R. A., Han, Y., Bui, T. V., Nusinowitz, S., Bok, D., Lichter, J., Widder, K., Travis, G. H., & Mata, N. L. (2005). Reductions in serum vitamin A arrest accumulation of toxic retinal fluorophores: a potential therapy for treatment of lipofuscin-based retinal diseases. *Investigative Ophthalmology & Visual Science*, *46*(12), 4393–4401.

Reis, L. M., Seese, S. E., Costakos, D., & Semina, E. V. (2024). Congenital anterior segment ocular disorders: Genotype-phenotype correlations and emerging novel mechanisms. *Progress in Retinal and Eye Research*, *102*(101288), 101288.

Ren, X., & Léveillard, T. (2022). Modulating antioxidant systems as a therapeutic approach to retinal degeneration. *Redox Biology*, *57*(102510), 102510.

Richardson, R., Tracey-White, D., Webster, A., & Moosajee, M. (2017). The zebrafish eye-a paradigm for investigating human ocular genetics. *Eye*, *31*(1), 68–86.

Rodwell, V., Patil, M., Kuht, H. J., Neuhauss, S. C. F., Norton, W. H. J., & Thomas, M. G. (2023). Zebrafish optokinetic reflex: Minimal reporting guidelines and recommendations. *Biology*, *13*(1). https://doi.org/10.3390/biology13010004

Rojas-Muñoz, A., Dahm, R., & Nüsslein-Volhard, C. (2005). Chokh/rx3 specifies the retinal pigment epithelium fate independently of eye morphogenesis. *Developmental Biology*, *288*(2), 348–362.

Sakamaki, K., Nozaki, M., Kominami, K., & Satou, Y. (2007). The evolutionary conservation of the core components necessary for the extrinsic apoptotic signaling pathway, in Medaka fish. *BMC Genomics*, *8*, 141.

Salomè, S., Ciampa, N., Giordano, M., Raimondi, R., Capone, E., Grieco, C., Coppola, C., Capasso, L., & Raimondi, F. (2023). Ophthalmological impairment in patients with congenital cytomegalovirus infection. *Frontiers in Pediatrics*, *11*, 1251893.

Sampath, K., Rubinstein, A. L., Cheng, A. M., Liang, J. O., Fekany, K., Solnica-Krezel, L., Korzh, V., Halpern, M. E., & Wright, C. V. (1998). Induction of the zebrafish ventral brain and floorplate requires cyclops/nodal signalling. *Nature*, *395*(6698), 185–189.

Santiago-Colón, A., Rocheleau, C. M., Chen, I.-C., Sanderson, W., Waters, M. A., Lawson, C. C., Langlois, P. H., Cragan, J. D., Reefhuis, J., & National Birth Defects Prevention Study. (2020). Association between maternal occupational exposure to polycyclic aromatic hydrocarbons and rare birth defects of the face and central nervous system.

Birth Defects Research, *112*(5), 404–417.

Sarropoulos, I., Sepp, M., Yamada, T., Schäfer, P. S. L., Trost, N., Schmidt, J., Schneider, C., Drummer, C., Mißbach, S., Taskiran, I. I., Hecker, N., González-Blas, C. B., Kempynck, N., Frömel, R., Joshi, P., Leushkin, E., Arnskötter, F., Leiss, K., Okonechnikov, K., … Kaessmann, H. (2025). The evolution of gene regulation in mammalian cerebellum development. In *Evolutionary Biology* (No. biorxiv;2025.03.14.643248v1). bioRxiv. https://www.biorxiv.org/content/10.1101/2025.03.14.643248v1

Sayed, A. E.-D. H., Soliman, H. A. M., & Mitani, H. (2019). UVA-induced neurotoxicity in Japanese medaka (Oryzias latipes). *Photochemical & Photobiological Sciences: Official Journal of the European Photochemistry Association and the European Society for Photobiology*, *18*(1), 71–79.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682.

Schlegel, D. K., & Neuhauss, S. C. F. (2020). Chapter 3 - The larval visual system and behavioral responses to visual stimuli. In R. T. Gerlai (Ed.), *Behavioral and Neural Genetics of Zebrafish* (pp. 35–48). Academic Press.

Schneider, G. E. (1969). Two visual systems. *Science (New York, N.Y.)*, *163*(3870), 895–902.

Scott, M. F., Ladejobi, O., Amer, S., Bentley, A. R., Biernaskie, J., Boden, S. A., Clark, M., Dell'Acqua, M., Dixon, L. E., Filippi, C. V., Fradgley, N., Gardner, K. A., Mackay, I. J., O'Sullivan, D., Percival-Alwyn, L., Roorkiwal, M., Singh, R. K., Thudi, M., Varshney, R. K., … Mott, R. (2020). Multi-parent populations in crops: a toolbox integrating genomics and genetic mapping with breeding. *Heredity*, *125*(6), 396–416.

Sheeladevi, S., Lawrenson, J. G., Fielder, A. R., & Suttle, C. M. (2016). Global prevalence of childhood cataract: a systematic review. *Eye*, *30*(9), 1160–1169.

Shi, C., Yuan, X., Chang, K., Cho, K.-S., Xie, X. S., Chen, D. F., & Luo, G. (2018). Optimization of optomotor response-based visual function assessment in mice. *Scientific Reports*, *8*(1), 9708.

Shima, A., & Shimada, A. (1988). Induction of mutations in males of the fish Oryzias latipes at a specific locus after gamma-irradiation. *Mutation Research*, *198*(1), 93–98.

Singh, R. K., Zhao, Y., Elze, T., Fingert, J., Gordon, M., Kass, M. A., Luo, Y., Pasquale, L. R., Scheetz, T., Segrè, A. V., Wiggs, J. L., & Zebardast, N. (2023). Polygenic risk score improves prediction of primary open angle glaucoma onset in the Ocular Hypertension

Treatment Study. *medRxiv: The Preprint Server for Health Sciences*. https://doi.org/10.1101/2023.08.15.23294141

Sinn, R., Peravali, R., Heermann, S., & Wittbrodt, J. (2014). Differential responsiveness of distinct retinal domains to Atoh7. *Mechanisms of Development*, *133*, 218–229.

Sinn, R., & Wittbrodt, J. (2013). An eye on eye development. *Mechanisms of Development*, *130*(6-8), 347–358.

Sirisi, S., Folgueira, M., López-Hernández, T., Minieri, L., Pérez-Rius, C., Gaitán-Peñas, H., Zang, J., Martínez, A., Capdevila-Nortes, X., De La Villa, P., Roy, U., Alia, A., Neuhauss, S., Ferroni, S., Nunes, V., Estévez, R., & Barrallo-Gimeno, A. (2014). Megalencephalic leukoencephalopathy with subcortical cysts protein 1 regulates glial surface localization of GLIALCAM from fish to humans. *Human Molecular Genetics*, *23*(19), 5069–5086.

Sparrow, J. R., Gregory-Roberts, E., Yamamoto, K., Blonska, A., Ghosh, S. K., Ueda, K., & Zhou, J. (2012). The bisretinoids of retinal pigment epithelium. *Progress in Retinal and Eye Research*, *31*(2), 121–135.

Spence, R., Gerlach, G., Lawrence, C., & Smith, C. (2007). The behaviour and ecology of the zebrafish, Danio rerio. *Biological Reviews of the Cambridge Philosophical Society*, *83*(1), 13–34.

Spivakov, M., Auer, T. O., Peravali, R., Dunham, I., Dolle, D., Fujiyama, A., Toyoda, A., Aizu, T., Minakuchi, Y., Loosli, F., Naruse, K., Birney, E., & Wittbrodt, J. (2014). Genomic and phenotypic characterization of a wild medaka population: towards the establishment of an isogenic population genetic resource in fish. *G3* , *4*(3), 433–445.

Stepniewska, I., Ql, H. X., & Kaas, J. H. (2000). Projections of the superior colliculus to subdivisions of the inferior pulvinar in New World and Old World monkeys. *Visual Neuroscience*, *17*(4), 529–549.

St Johnston, D. (2002). The art and design of genetic screens: Drosophila melanogaster. *Nature Reviews. Genetics*, *3*(3), 176–188.

Strauss, O. (1995). The retinal pigment epithelium. In *Webvision: The Organization of the Retina and Visual System*. University of Utah Health Sciences Center.

Strickler, A. G., Yamamoto, Y., & Jeffery, W. R. (2007). The lens controls cell survival in the retina: Evidence from the blind cavefish Astyanax. *Developmental Biology*, *311*(2), 512–523.

Suzuki, R., Woo, J. Z., Thumberger, T., Hofmann, G., Wittbrodt, J., & Tavhelidse-Suck, T. (2024). Characterizing medaka visual features using a high-throughput optomotor

response assay. *PloS One*, *19*(6), e0302092.

Takamiya, M., Weger, B. D., Schindler, S., Beil, T., Yang, L., Armant, O., Ferg, M., Schlunck, G., Reinhard, T., Dickmeis, T., Rastegar, S., & Strähle, U. (2015). Molecular description of eye defects in the zebrafish Pax6b mutant, sunrise, reveals a Pax6b-dependent genetic network in the developing anterior chamber. *PloS One*, *10*(2), e0117645.

Tandon, A., & Mulvihill, A. (2009). Ocular teratogens: old acquaintances and new dangers. *Eye*, *23*(6), 1269–1274.

Taniguchi, Y., Takeda, S., Furutani-Seiki, M., Kamei, Y., Todo, T., Sasado, T., Deguchi, T., Kondoh, H., Mudde, J., Yamazoe, M., Hidaka, M., Mitani, H., Toyoda, A., Sakaki, Y., Plasterk, R. H. A., & Cuppen, E. (2006). Generation of medaka gene knockout models by target-selected mutagenesis. *Genome Biology*, *7*(12), R116.

Team R. (2020). *RStudio: Integrated Development Environment for R*. https://doi.org/10.1109/ICAC3N53548.2021.9725647

Thomas, J. L., & Thummel, R. (2013). A novel light damage paradigm for use in retinal regeneration studies in adult zebrafish. *Journal of Visualized Experiments: JoVE*, *80*, e51017.

Toyama, K. (1916). Ichi ni no MENDEL seishitsui ni tsuite [On Some Mendelian Traits] (title translated by author). *Nihon Ikushugaku Kaiho*, *1*, 1–9.

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., & Posthuma, D. (2021). Genome-wide association studies. *Nature Reviews. Methods Primers*, *1*(1). https://doi.org/10.1038/s43586-021-00056-9

Verma, A. S., & Fitzpatrick, D. R. (2007). Anophthalmia and microphthalmia. *Orphanet Journal of Rare Diseases*, *2*, 47.

Vinberg, F., Chen, J., & Kefalov, V. J. (2018). Regulation of calcium homeostasis in the outer segments of rod and cone photoreceptors. *Progress in Retinal and Eye Research*, *67*, 87–101.

Vlasiuk, A., & Asari, H. (2021). Feedback from retinal ganglion cells to the inner retina. *PloS One*, *16*(7), e0254611.

Weber, A., Hochmann, S., Cimalla, P., Gärtner, M., Kuscha, V., Hans, S., Geffarth, M., Kaslin, J., Koch, E., & Brand, M. (2013). Characterization of light lesion paradigms and optical coherence tomography as tools to study adult retina regeneration in zebrafish. *PloS One*, *8*(11), e80483.

Welz, B., Pierotti, S., Fitzgerald, T. W., Thumberger, T., Suzuki, R., Watson, P., Fuss, J.,

Cordeiro da Trindade, T., Defranoux, F., Ferreira, M., Naruse, K., Gierten, J., Loosli, F., Wittbrodt, J., & Birney, E. (2025). Discovery and characterisation of gene by environment and epistatic genetic effects in a vertebrate model. In *bioRxiv*. https://doi.org/10.1101/2025.04.24.650462

Wilson, S. W., & Houart, C. (2004). Early steps in the development of the forebrain. *Developmental Cell*, *6*(2), 167–181.

Winkler, S., Loosli, F., Henrich, T., Wakamatsu, Y., & Wittbrodt, J. (2000). The conditional medaka mutation eyeless uncouples patterning and morphogenesis of the eye. *Development (Cambridge, England)*, *127*(9), 1911–1919.

Wittbrodt, J., Shima, A., & Schartl, M. (2002). Medaka--a model organism from the far East. *Nature Reviews. Genetics*, *3*(1), 53–64.

Wu, J., Mao, L., Tao, J., Wang, X., Zhang, H., Xin, M., Shang, Y., Zhang, Y., Zhang, G., Zhao, Z., Wang, Y., Cui, M., Wei, L., Song, X., & Sun, X. (2022). Dynamic quantitative trait loci mapping for plant height in recombinant inbred line population of upland cotton. *Frontiers in Plant Science*, *13*, 914140.

Xue, Y., Sato, S., Razafsky, D., Sahu, B., Shen, S. Q., Potter, C., Sandell, L. L., Corbo, J. C., Palczewski, K., Maeda, A., Hodzic, D., & Kefalov, V. J. (2017). The role of retinol dehydrogenase 10 in the cone visual cycle. *Scientific Reports*, *7*(1), 2390.

Yang, S., Zhou, J., & Li, D. (2021). Functions and diseases of the retinal pigment epithelium. *Frontiers in Pharmacology*, *12*, 727870.

Yan, J., Wang, L., Yang, Q.-L., Yang, Q.-X., He, X., Dong, Y., Hu, Z., Seeliger, M. W., Jiao, K., & Paquet-Durand, F. (2024). T-type voltage-gated channels, Na+/Ca2+-exchanger, and calpain-2 promote photoreceptor cell death in inherited retinal degeneration. *Cell Communication and Signaling: CCS*, *22*(1), 92.

Yan, W., Deng, X. W., Yang, C., & Tang, X. (2021). The genome-wide EMS Mutagenesis bias correlates with sequence context and chromatin structure in rice. *Frontiers in Plant Science*, *12*, 579675.

Yates, A. D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Azov, A. G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Marugán, J. C., Cummins, C., Davidson, C., Dodiya, K., Fatima, R., Gall, A., … Flicek, P. (2020). Ensembl 2020. *Nucleic Acids Research*, *48*(D1), D682–D688.

Yau, K.-W., & Hardie, R. C. (2009). Phototransduction motifs and variations. *Cell*, *139*(2), 246–264.

Yedutenko, M., Howlett, M. H. C., & Kamermans, M. (2020). High Contrast Allows the

Retina to Compute More Than Just Contrast. *Frontiers in Cellular Neuroscience*, *14*, 595193.

Yeritsyan, N., Lehmann, K., Puk, O., Graw, J., & Löwel, S. (2012). Visual capabilities and cortical maps in BALB/c mice. *The European Journal of Neuroscience*, *36*(6), 2801–2811.

Zhu, X., Huang, L., Zheng, Y., Song, Y., Xu, Q., Wang, J., Si, K., Duan, S., & Gong, W. (2019). Ultrafast optical clearing method for three-dimensional imaging with cellular resolution. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(23), 11480–11489.

Zuber, M. E., Gestri, G., Viczian, A. S., Barsacchi, G., & Harris, W. A. (2003). Specification of the vertebrate eye by a network of eye field transcription factors. *Development (Cambridge, England)*, *130*(21), 5155–5167.

Publications

Publications during my PhD

Suzuki, R., Woo, J.Z., Thumberger, T., Hofmann, G., Wittbrodt, J., Tavhelidse-Suck, T., 2024. Characterizing medaka visual features using a high-throughput optomotor response assay. PLoS One 19, e0302092.

Fitzgerald, T., Brettell, I., Leger, A., Wolf, N., Kusminski, N., Monahan, J., Barton, C., Herder, C., Aadepu, N., Gierten, J., Becker, C., Hammouda, O.T., Hasel, E., Lischik, C., Lust, K., Sokolova, N., **Suzuki, R.,** Tsingos, E., Tavhelidse, T., Thumberger, T., Watson, P., Welz, B., Khouja, N., Naruse, K., Birney, E., Wittbrodt, J., Loosli, F., 2022. The Medaka Inbred Kiyosu-Karlsruhe (MIKK) panel. Genome Biol. 23, 59.

Leger, A., Brettell, I., Monahan, J., Barton, C., Wolf, N., Kusminski, N., Herder, C., Aadepu, N., Becker, C., Gierten, J., Hammouda, O.T., Hasel, E., Lischik, C., Lust, K., Sokolova, N., **Suzuki, R.,** Tavhelidse, T., Thumberger, T., Tsingos, E., Watson, P., Welz, B., Naruse, K., Loosli, F., Wittbrodt, J., Birney, E., Fitzgerald, T., 2022. Genomic variations and epigenomic landscape of the Medaka Inbred Kiyosu-Karlsruhe (MIKK) panel. Genome Biol. 23, 58.

Submitted manuscript

Welz, B., Pierotti, S., Fitzgerald, T.W., Thumberger, T., **Suzuki, R**., Watson, P., Fuss, J., Cordeiro da Trindade, T., Defranoux, F., Ferreira, M., Naruse, K., Gierten, J., Loosli, F., Wittbrodt, J., Birney, E., 2025. Discovery and characterisation of gene by environment and epistatic genetic effects in a vertebrate model. bioRxiv. https://doi.org/10.1101/2025.04.24.650462

Afting, C., Bhatti, N., Schlagheck, C., Salvador, E.S., Herrera-Astorga, L., Agarwal, R., **Suzuki, R.,** Hackert, N., Lorenz, H.-M., Zilova, L., Wittbrodt, J., Exner, T., 2025. Deep learning predicts tissue outcomes in retinal organoids. bioRxiv. https://doi.org/10.1101/2025.02.19.639061

Acknowledgment

First of all, I would like to express my deepest gratitude to my supervisor **Prof. Dr. Jochen Wittbrodt** for giving me the incredible opportunity to work in your lab. It all started with the email that I hesitated to send for whole week… I still vividly remember the day of my interview, presenting my master's thesis with shaky hands and a trembling pointer. That day, and the conversation we had at the table in your office, will always stay with me.
I'm so grateful to have had you as a supervisor. Whenever I felt like I'd reached a dead end, talking with you always turned things around. You have this remarkable ability to make things work out, and you somehow know exactly how to make that happen. You as a scientist is truly inspiring, always looking at the bigger picture, always finding the bright side of unexpected result. Not only about science but you also tought me what it means to be a thoughtful and generous leader. You always notice the small changes in people, make time for them, and check in to ensure everyone is doing okay. Most of all, you always gave me the push to move forward, even in uncertain times. Thank you, Jochen!

I would also like to extend my sincere thanks to **Prof. Dr. Ewan Birney** for the opportunity to work on this collaborative project. I also appreciate all the valuable input on statistics and analysis that you provided during TAC meetings and Indigene meetings.

I am very grateful to **Prof. Dr. Lazaro Centanin** and **Prof. Dr. Hilmar Bading** for joining on my Thesis Advisory Committee, for your thoughtful feedback, and for your continued support, including acting as my examiners.

I would also like to extend my thanks to **Prof. Dr. Kiyoshi Naruse** for sharing documents on old medaka papers, giving us the opportunity to learn about exciting research in Japan by organizing meetings, and even letting us join a wild medaka catch. It was fascinating to talk science with you.

I am grateful for **Dr. Felix Loosli** for providing embryos from KIT for the experiments.
I'm grateful for **Dr. Vladimir Benes** and **Mireia Osuna Lopez** for the great collaboration.
I would like to thank **Prof.Dr Stephan Neuhauss** for hosting me to perform OKR.
I am grateful for **Dr. Jingjing Zang** for teaching me OKR experiments and performing ERG.

I am grateful for the members of the Birney group: **Fanny, Tom, Ian, and Esther,** Thank you for your insightful input during the analysis. It greatly improved the quality of the results and deepened my understanding of the data. Especially **Esther-** Thank you for being such an easy person to communicate with, for your reliability, and for providing the data, I truly appreciated your support and dedication.

I thank **Frederike Seibold** for all of your support from the beginning even before I arrived in Germany, helping me navigate all the bureaucracy with your deep well of experience and

knowledge. I thank **Eva Hasel de Carvalho** thank you too—for supporting me with administrative matters and even stepping in to handle last-minute FedEx registration modification.

I greatly thank **Gero Hofmann** for building all the experimental setup of OMR and Light injury. You can build anything and it's just so impressive. I appreciated **Niklas Grammling** for building the LED light in OMR setup and reassembling of the monitor. Without you two's great skills this project would have been impossible.
I would like to thank **Damjan Kalšan** for developing the software and for being so easy to work with.

I am grateful for **Prof. Dr. Satoshi Ansai** for the valuable discussions on the OMR setup during challenging times, whose suggestions were instrumental in improving the setup. I'm also grateful for the expert input from **Prof. Dr. Hideaki Takeuchi**.
Many thanks to **Prof. Dr. Kiyoshi Naruse**, **Assis.Prof. Dr. Ai Shinomiy and fish care taking team** at NBRP for providing eggs, as well as for your warm hospitality and reliable, efficient work. Your support made our stay and experiments run incredibly smoothly.

I am thankful for the great work in the fish room of **Marzena Majewski, Rebecca Lipp, Erik Leist, Rachel Müller and Antonino Saraceno,** who have made an important contribution to the research group and the success of our projects.
I very much appreciate **Dr. Thomas Thumberger.** I had so much fun developing the assay with you. Thanks for your help with analysis and coding, all the crazy and brilliant idea and all the jokes made many fun memories in the lab.
I am thankful for **Dr. Tinatini Tavhelidse-Suck**, working on the manuscript with you was such a smooth process and I learned a lot from you!

I would like to thank **Tanja Kellner** for always knowing every protocol in well detail with key tips and with precision. You're not only impressive in crafting skill but also impressively efficient in everything you do. I'm also really grateful for your help with egg collection and shipping preparation.
I am also grateful for **Beate Wittbrodt** for not only for your help with experiments, but also for your support in preparing media, collecting eggs, and even taking on kitchen duties. Your help allowed me to focus on the project, and I'm truly grateful. It was always so reassuring to have someone with such cumulative experience in Bay 1, your presence was a real source of strength.

All the 5th floor + members (now and then)
- Rashi, Laura, Bettina, Philip, Cassian, Christina, Jana, Encarni, Kaisa, Javi, Jasmin, Anthi, Kristaps, Fariha, Ayse, Natalia, Sankeert, Alex, Omar, Clara, Andread, and Verena - I started writing message for every one of you but I soon realized that time and space would never be enough to write the so many things that I'm grateful for about you all! In short, thank you for being there—not just as colleagues, but as kind, supportive, and caring people who made the stress and challenges feel lighter. You brought humor and

encouragement into this journey, and I will always remember that. - **Rashi, Cassian, Encarni, Kaisa, Bettina, Laura, and Jana** - Thank you for being my late-night or weekend lab companions and for supporting me through difficult times. I truly don't know how I would have managed without your encouragement and friendship. I'm so grateful for your help, even when you were busy yourselves. - **Bettina, Jana, Philip, & Thomas** - Thank you for all the fun time in Japan and in London! All the mid night presentation preparation to late night trouble shooting in the lab, and huge bottle of whiskey will not be forgotten. - **Lucie & Venera** - Thank you for the inspiring pep talks and for showing me how one can be both a dedicated scientist and a caring parent. - **Christina, Kaisa** - you two always inspired me to do sports which was helpful for keep going under the stress time. - **Joergen** - I'm so grateful for all the discussion on the analysis. - **Anthi** - your passion to science has always been so motivating - **Kristaps** - let's keep bouldering! - **Encarni, Rashi, Cassian** - You guys always kept me sane during tough time, through all the emotional support and deep talks, thank you!

- Encarni, Bettina, Cassian, Rashi, Thomas, Eva, Esther, Jingjing, and Tetsuya - Thank you so much for proofreading my thesis. Even under the time constraint. Your insightful comments and feedback improved it so much.

I am thankful to these people outside of the work who have helped shape the person I am today and encouraged me along the path that led me here.
- **Prof. Dr. Verena Schröder -** I'm so grateful for having worked with you for the internship. You inspired me in so many aspects and am so grateful for encouraging me to take a step forward for myself. - **Valerie** - you were the first person to put the idea of me living in Germany when I was still 17 years old, and having known that you were in the same country made me feel like I'm not alone. - **Greetje** - thank you for always being full on kind heart and giving me slices of oranges when I needed the most (when I was sick). Your emotional intelligence, insights in life, and kind heart have supported me through tough times.
I am grateful for - **Keith** - my Canadian dad - for showing it's never too late to start learning something new (started learning python at the age of 82.) - **Tetsuya** - your passion and enthusiasm to science and your determined work ethics always inspire me. I'm so lucky and grateful to have you in my life, who's reliable, caring, knowing always what's right and always being there for me. Thank you for always making me feel truly at home, no matter where I am.

I'd like to thank to long lasting friendship back home in Japan (**Chi-chan, Izumi, Mai, Miki, Mamiko, Saorin, Nami, Koharu Sachee, Yusuke and all my friends in Japan**), I am deeply grateful to each and every one of you. Thank you for always longing for my return, encouraging me, and supporting me with all your warm heart. Knowing that I have a place to return to has given me the strength to keep going even far away from home across the ocean.

Finally, I would like to express my heartfelt gratitude to my parents and family for always supporting me, raising me with love, pushing me to take on new challenges, and always standing by my side. 心から感謝しています。いつもありがとう。

Declaration

Herewith I declare that I have written the PhD thesis "Genetic Variation Shapes Visual-Motor Behavior: Insights from High-Throughput Optomotor Response Screening in Medaka Inbred Population" on my own and with no other sources and aids than quoted.

Heidelberg, 2025

List of Figures

List of Tables