Diverse prey capture strategies in teleost larvae

Duncan S. Mearns[1,2,*], Sydney A. Hunt[1,3,*], Martin W. Schneider[1], Ash V. Parker[1], Manuel Stemmer[1], and Herwig Baier[1]

[1]Max Planck Institute for Biological Intelligence, Department Genes – Circuits – Behavior, Am Klopferspitz 18, 82152 Martinsried, Germany

[2]Present address: Princeton Neuroscience Institute, Princeton, NJ, USA

[3]Present address: Max Planck Institute of Animal Behavior, Am Obstberg 1, 78315 Radolfzell, Germany

*Equal contribution

Correspondence: herwig.baier@bi.mpg.de

Summary

Animal behavior is adapted to the sensory environment in which it evolved, while also being constrained by physical limits, evolutionary history, and developmental trajectories. The hunting behavior of larval zebrafish (*Danio rerio*), a cyprinid native to streams in Eastern India, has been well characterized. However, it is unknown if the complement and sequence of movements employed during prey capture by zebrafish is universal across freshwater teleosts. Here, we explore the syntax of prey capture behavior in larval fish belonging to the clade *Percomorpha*, whose last common ancestor with cyprinids lived ~240 million years ago. We compared the behavior of four cichlid species from Lake Tanganyika endemic to deep benthic parts of the lake (*Lepidiolamprologus attenuatus*, *Lamprologus ocellatus*, and *Neolamprologus multifasciatus*) or inhabiting rivers (*Astatotilapia burtoni*) with that of medaka (*Oryzias latipes*), a fish found in rice paddies in East Asia. Using high speed videography and neural networks, we tracked eye movements and extracted swim kinematics during hunting from larvae of these five species. Notably, we found that the repertoire of hunting movements of cichlids is broader than that of zebrafish, but shares basic features, such as eye convergence, positioning of prey centrally in the binocular visual field, and discrete prey capture bouts, including two kinds of capture strikes. In contrast, medaka swim continuously, track the prey monocularly without eye convergence, and position prey laterally before capturing them with a side swing. This configuration of kinematic motifs suggests that medaka may judge distance to prey predominantly by motion parallax, while cichlids and zebrafish may mainly use binocular visual cues. Together, our study documents the diversification of locomotor and oculomotor adaptations among hunting teleost larvae.

Introduction

In recent decades, biologists have increasingly relied on a handful of genetically tractable species to study questions related to behavioral mechanisms and their underlying neural circuitry (Devineni & Scaplen, 2022; Piggott et al., 2011; Zhu & Goodhill, 2023). However, it is often not clear to what extent these findings in model organisms can be generalized to other taxa or even closely related species. For example, recent comparative behavioral studies in drosophilids that diverged less than 40 mya found differences in both locomotion (York et al., 2022) and sequencing of spontaneous behaviors (Hernández et al., 2021). For vertebrates, the zebrafish larva has become a dominant model for understanding the neuronal circuits and pathways controlling innate behavior (Baier & Scott, 2009; Friedrich et al., 2010; Gahtan & Baier, 2004; Orger & de Polavieja, 2017; Portugues & Engert, 2011). While many aspects of the visuomotor transformations and underlying neural circuitry for prey capture have been revealed in this species (reviewed by Baier & Scott, 2024; Zhu & Goodhill, 2023), it is not known if this is the general solution for larval teleosts or a derived adaptation to the zebrafish's specific ecological niche.

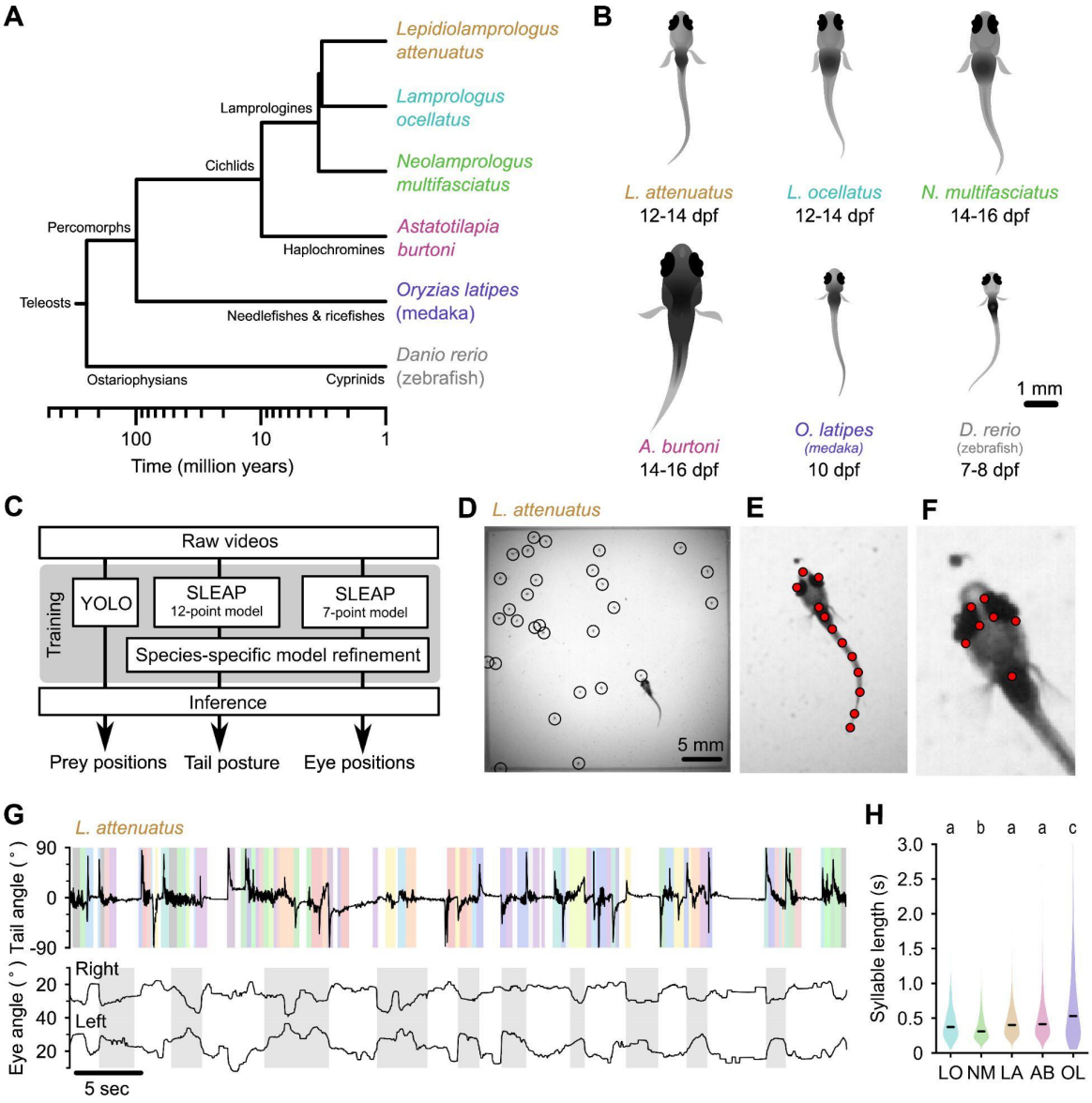


Figure 1. Tracking prey capture behavior in five species of fish larvae. (**A**) Phylogenetic relationship between percomorph species used in this study, and their relationship to zebrafish (a cyprinid). Based on Betancur-R et al. (2017). (**B**) Schematics of larvae of each species studied here (and zebrafish, *Danio rerio*, for reference). Scale bar: 1 mm. (**C**) Outline of tracking procedure. Raw video frames were analyzed with three different neural networks to extract tail pose, eye pose and prey position in each frame. Kinematic features were extracted from raw pose data and used for subsequent analyses. (**D-F**) Tracking output for each neural network for a single frame from a recording of an *L. attenuatus* larva, showing identified artemia (D), tail points (E) and eye points (F). (**G**) One minute of tail and eye tracking from *L. attenuatus* showing tail tip angles (top trace) and right and left eye angles (bottom traces) relative to midline. Bottom trace, gray shaded boxes: automatically identified prey capture periods, when the eyes are converged (see **Figure 3**). Top trace, shaded boxes: automatically identified and classified behavioral syllables. Color corresponds to cluster identity in **Figure 5**. (**H**) Distribution of syllable length across species for all fish combined. Black bars indicate median; letters denote significance groups (bootstrap test difference between medians with Bonferroni correction). See also **Supp. Fig. 1**.

Here, we applied advances in computational ethology, such as automated tracking with deep neural networks (Mathis et al., 2018; Pereira et al., 2019) and unsupervised analyses (Berman et al., 2014; Marques et al., 2018; Mearns et al., 2020; Wiltschko et al., 2015) to compare hunting behavior of zebrafish to the Japanese rice fish, medaka (*O. latipes*) (Mano & Tanaka, 2012) and four cichlids from Lake Tanganyika (El Taher et al., 2021; Higham et al., 2007). Medaka and cichlids belong to a diverse clade of teleosts known as percomorphs, whose last common ancestor with ostariophysi, such as zebrafish and Mexican cavefish, lived ~240 mya (Betancur-R et al., 2017) (**Fig. 1A**). While medaka and cichlids are only distantly related (~100 mya; Betancur-R et al., 2017), the haplochromine species and the three lamprologine species studied here diverged recently, within the past 3 million years (Ronco et al., 2021). We discovered that swim kinematics and the use of eye movements differ qualitatively and quantitatively between these species. This divergence may be driven, in part, by differences in the sensorineural mechanisms underlying prey detection. We believe that a detailed and quantitative description of behavior is an important first step toward uncovering underlying neural mechanisms and relating behavioral differences between species to their ecological niches.

Results

Artificial neural networks track swimming behavior and eye movements in percomorph larvae

At five to seven days post fertilization (dpf), zebrafish larvae robustly feed on small prey items such as paramecia. As juveniles, they feed on zooplankton such as brine shrimp (artemia), but these prey are too large larvae to consume. To stage-match larvae of different species to zebrafish, we first identified when they started feeding (**Fig. 1B**). Cichlid larvae started feeding later than zebrafish, from 12-14 dpf, and medaka (*O. latipes*) started feeding at approximately 10 dpf. We next adapted an experimental paradigm used to study prey capture in zebrafish for these other species (Mearns et al., 2020). Individual larvae were placed in chambers with prey items (either artemia or paramecia). We recorded each animal for 15 minutes using a high-speed camera (**Supp. Videos 1-6**, **Methods**). We then used neural networks to extract tail pose, eye movements, and prey locations from videos (**Fig. 1C**). We trained a 12-point SLEAP model (Pereira et al., 2022) to track the tail, and a 7-point model to track the eyes of larvae (**Fig. 1E,F; Supp. Vid. 1**). We tracked prey position using YOLO (Redmon et al., 2016; **Fig. 1D**). We subsequently used these estimated pose dynamics (**Fig. 2**) and prey location information to compare hunting and swimming across species.

Larvae of different species have diverse swim patterns

At the first-feeding stage, zebrafish larvae swim in discrete, discontinuous bouts (Budick & O'Malley, 2000), which merge into continuous swimming at the juvenile stage (Westphal & O'Malley, 2013). In contrast, *Danionella cerebrum*, a close relative of zebrafish, exhibits slow

continuous swimming as larvae (Rajan et al., 2022). It is not known how larval swimming patterns differ between more distantly related species.

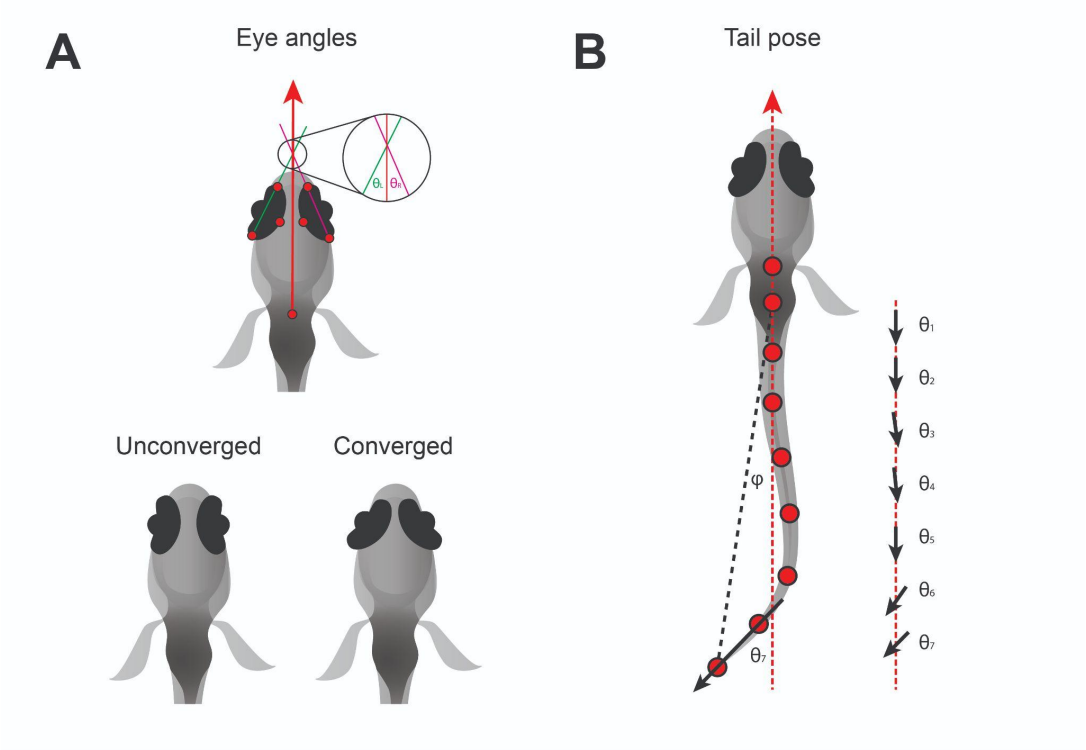


Figure 2: Pose information extracted from tracking data
(**A**) Definition of left (θ_L) and right (θ_R) eye vergence angles. The heading vector (red arrow) is the vector starting from the midpoint of the fins and passing through the midpoint of the eyes. The orientation of each eye is computed from the temporal and nasal points (long axis of the eye). The eye angle is the signed angle between the heading and long axis of the eye. Zero degrees indicates the long axis of the eye is parallel with the heading. Nasalward rotation of the eye increases the vergence angle. Eye converegnce angle is $\theta_L + \theta_R$. Bottom: schematic showing eyes in an unconverged and converged state. (**B**) Parameters used to estimate tail pose. Tail pose is measured starting from the middle of the swim bladder. We compute the vector between each consecutive pair of points along the tail. At each pair of points, we compute the signed angle between the corresponding vector and the midline of the fish (RHS: $\theta_{1\text{-}7}$, angles between black arrows and red dashed line), providing a 7-dimensional representation of the tail pose in each frame. For visualization purposes in **Fig. 1G**, **Supp. Fig. 1**, we plot the tail angle, which is the signed angle between the swim bladder and the final tail point relative to the heading (angle between dashed red and black lines, φ).

We found marked variation in the continuity of swimming in early percomorph larvae (**Fig. 1G; Supp. Fig. 1**). On one extreme, medaka (OL) and *A. burtoni* (AB) swam with a continuous "gliding" style characterized by sustained, uninterrupted tail undulations over many seconds with only short breaks between swimming episodes, similar to *Danionella*. On the other extreme, *L. ocellatus* (LO) and *N. multifasciatus* (NM) had an intermittent swimming style, often resting at the bottom of the chamber for minutes at a time with quiescent periods interrupted by short, rapid bursts of activity. *L. attenuatus* (LA) showed a behavior intermediate to these extremes, alternating

between rapid tail beating and quiescence, each lasting on the order of a few seconds, a discontinuous style similar to zebrafish.

Analyses across the animal kingdom suggest that behavior is organized into sub-second kinematic motifs ("syllables"), even when movement is continuous (Berman et al., 2014; Stephens et al., 2008; Tinbergen, 1951; Wiltschko et al., 2015). Inspecting periods of swimming across percomorph species revealed significant substructure within swimming episodes (**Fig. 1G; Supp. Fig. 1**), suggesting that the continuous swimming of some species may in fact consist of kinematically discrete syllables fused together. We segmented continuous tail traces into discrete syllables using a change detection algorithm adapted from Mearns et al., 2020 (see **Methods**), which revealed changes in swim kinematics occuring on the order of tens to hundreds of milliseconds (median seconds: 0.33, LO; 0.29, NM; 0.38, LA; 0.36, AB; 0.49, OL). Furthermore, medaka (OL) exhibited considerably longer syllables than cichlids ($p < 0.01$, bootstrap test difference between medians), while one cichlid species, NM, exhibited significantly shorter syllables than the others ($p < 0.05$) (**Fig. 1H**). These results highlight differences in how fish larvae pattern their swim episodes, which may relate to different sensorimotor strategies for prey detection and tracking.

<u>Larvae of different species use diverse strategies to hunt prey</u>

Zebrafish larvae converge their eyes (nasalward rotation, **Fig. 2A**) at the onset of hunting episodes, and keep their eyes converged over the course of prey capture (Bianco et al., 2011; Patterson et al., 2013). Eye convergence aids prey capture by bringing prey into a binocular zone in the central visual field (Gahtan et al., 2005; Gebhardt et al., 2019). In contrast, some fish species such as the blind cave fish, *Astyanax mexicanus*, capture prey that are positioned laterally to their mouth (Lloyd et al., 2018; Espinasa et al., 2023). This shift in capture strategy appears to have happened over a relatively short timescale when cavefish diverged from their surface ancestors, and consequently cavefish still converge their vestigial eyes even though they do not rely on vision for prey capture (Espinasa & Lewis, 2023).

We found that all four cichlid species examined here converged their eyes during prey capture, but medaka did not (**Fig. 3A, Supp. Videos 2-6**). Two peaks in the distribution of eye convergence angles were visible in LA (**Fig. 3A**, center), but other cichlids also spent significant time with their eyes converged more than 50 degrees. From the raw videos, we could see that all cichlid species did converge their eyes during prey capture (e.g., **Supp. Video 3** (LO), 00:06; **Supp. Video 4** (NM), 00:22; **Supp. Video 5** (AB), 00:40); however, this might not manifest as two peaks in the distribution of eye convergence angles if these species spent less time engaged in prey capture than LA and if variation in the exact placement of key points by SLEAP blurred the boundary between converged and unconverged states. To verify that two eye convergence states also existed in these species, but not in medaka (OL), we fitted Gaussian mixture models to the eye convergence data of each species with one or two components. For all species except OL, modeling the data using

two underlying distributions provided a better fit than a single distribution (**Fig. 3B**). A unimodal distribution of eye vergence angles in OL was also present when we plotted the joint distribution of left and right eye angles (**Supp. Fig. 2**), indicating that this species does not perform eye convergence, whereas for all cichlid species there are periods in the recordings when the eyes are converged.

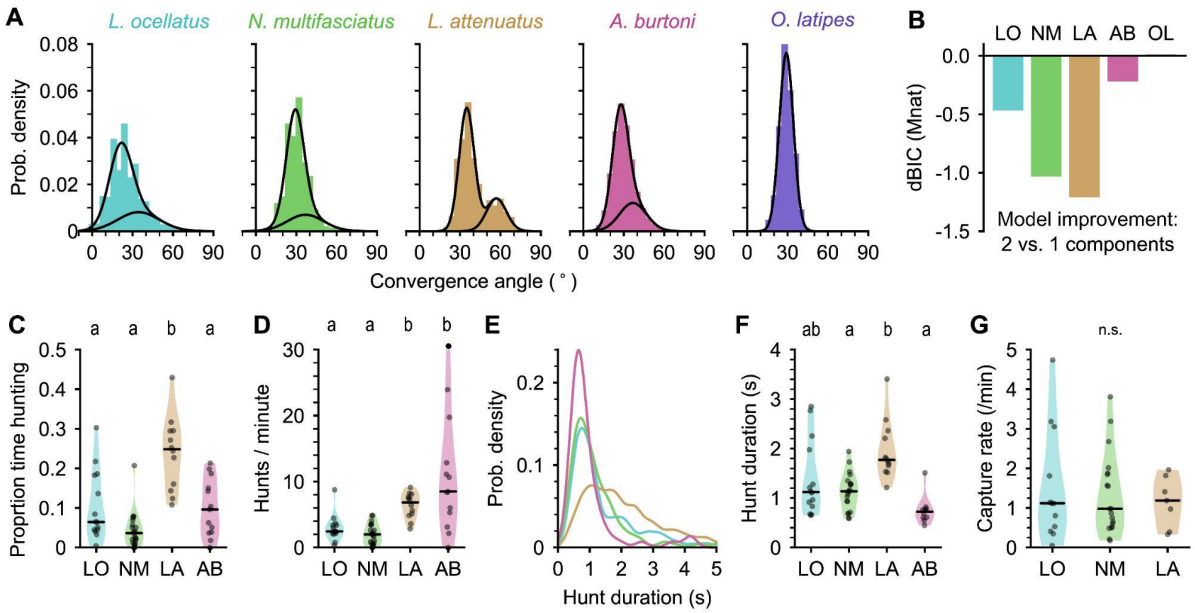


Figure 3. Eye movements during prey capture and statistics of hunting sequences. (**A**) Histograms of eye convergence angles for each species. The convergence angle is the angle between the long axes of the eyes. Shaded bars: normalized binned counts of convergence angles from all fish. Black lines: a best fit gaussian mixture model for each species (with one or two components). For all cichlid species, the data are better modeled as being drawn from two underlying distributions. For medaka (*O. latipes*), the data are better modeled with a single underlying distribution. (**B**) Improvement in fit for a two component over single component Gaussian mixture model, assessed using Bayesian inference criterion (BIC). The BIC is a measure of model fit, while punishing over-fitting. Lower values are better. Two mixtures provide a better fit over a single mixture for all species except medaka (OL). (**C**) Proportion of time spent by cichlids engaged in prey capture within the first five minutes of being introduced to the behavior arena. Points are single animals; black bar is the median; letters indicate significance (different letters indicate difference between groups, $\alpha = 0.05$). (**D**) Hunting rate, measured as the number of times eye convergence is initiated per minute, within the first five minutes. Points are single animals; black bar is the median; letters indicate significance. (**E**) Kernel density estimation of hunt durations for each species (all animals pooled). *A. burtoni* (pink) hunts skew shorter, *L. attenuatus* (brown) hunts tend to be longer. (**F**) Median hunt duration for each animal compared across species. Black bar is the median across animals; letters indicate significance (**G**) Capture rate (number of artemia consumed) per minute over the first five minutes for three cichlid species. Black bar is the median across animals. N.s. indicate no statistically significant difference between groups.

Using eye convergence to identify hunting periods, we found that different cichlid species spent different amounts of time engaged in prey capture (**Fig. 3C**). LO, NM, and AB spent less than 10% of the time hunting (median proportions: 0.064, LO; 0.036, NM; 0.096, AB). In contrast, LA, a piscivorous species as adults, spent about a quarter of the time actively engaged in prey capture, with all individuals tested spending at least 10% of their time and one individual spending over 40% of the time hunting (median: 0.25, p-values < 0.05 comparing LA to all other species, Mann-Whitney U test with Bonferroni correction).

Surprisingly, although LO, NM and AB all spent comparable total time engaged in prey capture, AB initiated hunts at a much higher rate, similar to LA (**Fig. 3D**) (median hunts/min: 2.45, LO; 1.98, NM; 6.85, LA; 8.49, AB; p-values < 0.05 comparing either LO or NM to LA or AB, Mann-Whitney U test with Bonferroni correction). AB hunts were also shorter than other species, while LA hunts tended to be longer (**Fig. 3E-F**). LO and NM individual hunt durations were similar to each other and intermediate to the other species (medians across fish (seconds): 1.12, LO; 1.14, NM; 1.77, LA; 0.73, AB). Despite these differences in time engaged in prey capture, species consumed prey at similar rates of approximately one artemia per minute (**Fig. 3G**).

Taken together, our results highlight differences among closely related cichlids in hunting behavior. LA are the most persistent hunters, initiating prey capture often and spending a longer time engaged in the behavior once initiated (**Supp. Vid. 2**). In contrast, AB prey capture dynamics are characterized by a high rate of short-duration hunting episodes (**Supp. Vid. 5**). Both LA and AB prey capture behaviors are examples of an active hunting strategy. On the other hand, LO and NM initiate prey capture rarely, spending much time resting at the bottom of the chamber, and moving only occasionally when they dart towards prey (**Supp. Vid. 3&4**), characteristic of a sit-and-wait predation strategy.

<u>Cichlids center prey within a strike zone</u>

Zebrafish larvae strike at prey once it is localized in the center of their visual field and ~0.5 mm away (the "strike zone") (Mearns et al., 2020; Patterson et al., 2013). This centering behavior is impaired when animals are blinded in one eye (Mearns et al., 2020). To test whether cichlids similarly center prey within a strike zone, we identified the most likely targeted artemia during each hunting episode and studied how the prey moved through the visual field as hunting sequences progressed (**Fig. 4A**, see **Methods**) (n events: 176, LO; 361, NM; 756, LA). This revealed that, within each prey capture sequence, prey became increasingly localized to the near central visual field over time. Computing kernel density estimates of prey distributions at the beginning, middle, and end of each hunting episode revealed that cichlids initiate prey capture at a wide range of distances, and that they center prey in the visual field in the early stages of a hunting sequence (**Fig. 4B**). For instance, the azimuthal angle of prey decreased over the start of the hunting episode in LO and NM ($p < 0.01$, bootstrap test difference between medians), (**Fig. 4D**), but not over the second half ($p > 0.05$) . Curiously, LA often initiated prey capture when

artemia were already in the center of the visual field (**Fig. 4A,B&D**), which may be due to them re-starting hunts on recently expelled prey (see description of spitting behavior below and **Supp. Vid. 6**). On average, the cichlids' eyes converged when prey were ~5 mm away (median distance, mm: 5.43, LO; 3.98, NM; 4.17, LA). However, all species could detect prey up to 15 mm away (**Fig. 4C**), and the distance to the prey decreases gradually over the course of the hunting episode, in contrast to centering which tends to occur at the beginning.

In all cichlid species analyzed, prey were highly localized to a central strike zone ~1-2 mm away at the end of prey capture (**Fig. 4B-C**). The tails of the distributions in **Fig. 4C** are likely due to aborted hunting events. Together, these results demonstrate that cichlids are able to detect prey at a large distance and possess the fine sensorimotor control required to localize prey to a strike zone.

<u>Fish larvae of different species share a common pose space</u>

Zebrafish larvae preferentially perform different kinds of swims during prey capture and exploration (Borla et al., 2002; McElligott & O'Malley, 2005). During prey capture, different swim types mediate distinct transformations of the visual scene: J-turns center prey in the visual field, approach swims bring prey to the strike zone, and strikes are deployed to capture prey. The identification and classification of these distinct swims has been aided by finding low-dimensional representations of behavior (Johnson et al., 2020; Marques et al., 2018; Mearns et al., 2020).

Due to neural and mechanical constraints, it is often possible to capture variation in tail shape over time with a greatly reduced number of dimensions, which serves to eliminate tracking noise and aid analysis (Berman et al., 2014; Mearns et al., 2020; Stephens et al., 2008; York et al., 2022). Performing principal components analysis (PCA) on all species combined revealed that tail pose dynamics are similarly low-dimensional (**Fig. 5A**, black line; 3 PCs explain > 90% variance). These results also hold when each species is analyzed individually (**Fig. 5A**, colored lines). As with zebrafish larvae (Girdhar et al., 2015; Mearns et al., 2020), the principal components correspond to a harmonic series (**Fig. 5B**), with the first PC capturing "turniness", and PCs 2 and 3 capturing tail oscillations during locomotion. The first three PCs were similar across species (**Fig. 5C**), allowing swim kinematics to be compared across species.

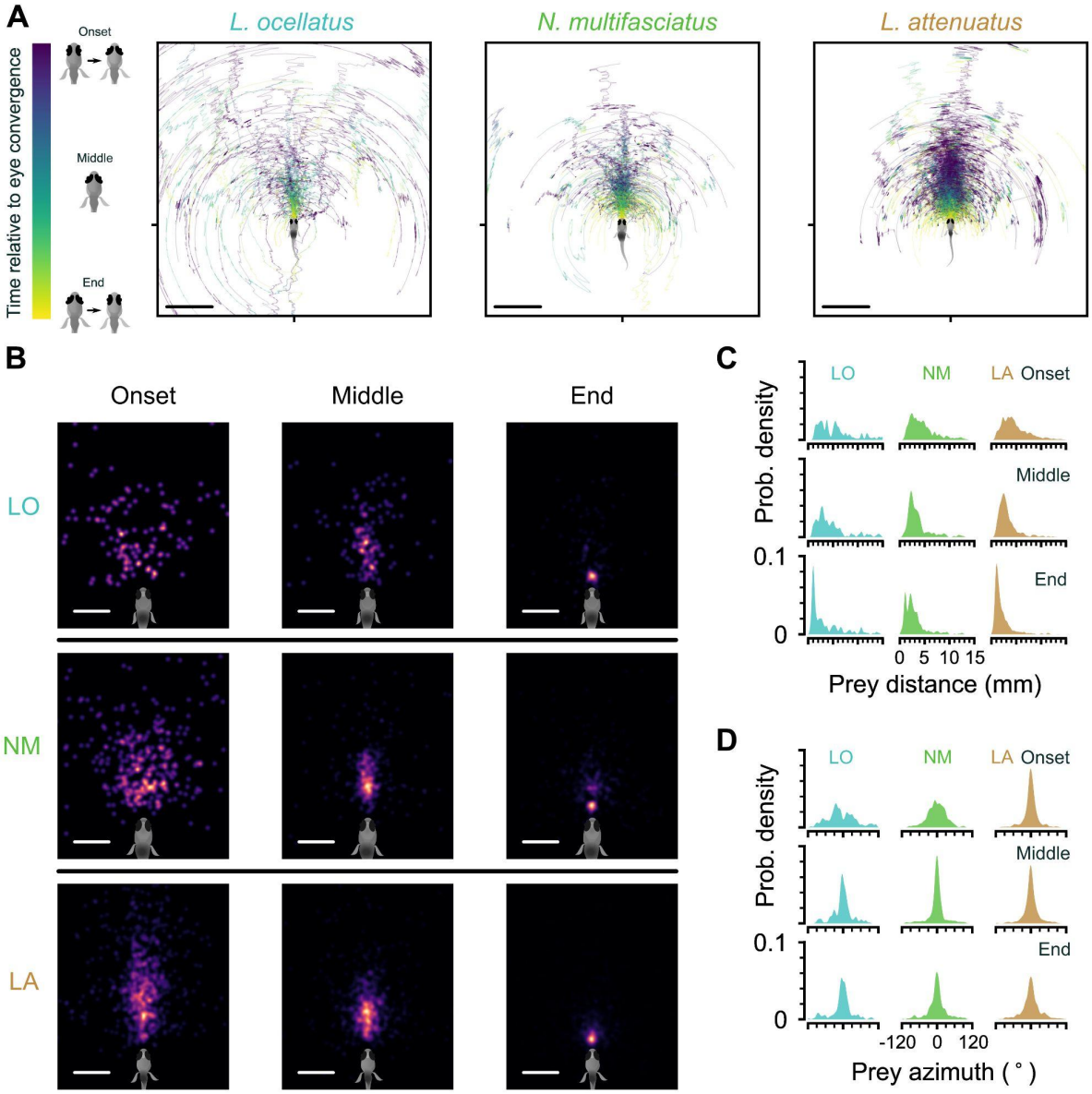


Figure 4. Location of prey in the visual field during prey capture in cichlids. (**A**) Trajectories of prey in the visual field for all automatically identified and tracked hunting events. Each prey is represented by a single line that changes in color from blue to yellow from the onset of eye convergence to when the eyes de-converge (end). Middle indicates the midpoint between eye convergence onset and end. Scale bar: 5 mm. (**B**) Kernel density estimation of the distribution of prey in the visual field across all hunting events. Rows: individual species. Columns: snapshots showing the distribution of hunted prey items at the beginning, middle and end of hunting sequences. Scale bar: 2 mm. (**C-D**) Kernel density estimation of prey distance (C) and azimuthal angle from the midline (D) at the onset (top), in the middle (center) and at the end (bottom) of hunting episodes. Each column shows the distribution from all events for a single species.

206 Hunting sequences in cichlids share kinematic motifs with zebrafish but are more complex

207 A common low-dimensional pose space in drosophilids has aided with the identification of
208 evolutionary trajectories in behavior (York et al., 2022). We therefore wanted to investigate the
209 trajectories through pose space of cichlids and medaka (OL) to understand if they share common
210 behaviors during hunting , or if their behaviors have diverged over evolution.

211 For each species, we projected the time series of rostrocaudal tail angles for each distinct bout onto
212 the first three principal components (Mearns et al., 2020). We then performed affinity propagation
213 to identify clusters corresponding to different swim types and performed t-SNE to visualize these
214 clusters in a low-dimensional embedding. LA spent the most time engaged in hunting in our
215 experiments, and so we chose to focus on this species in particular when discussing the cichlid
216 behavioral repertoire (**Fig. 5D-G)**; however, we performed similar analyses with all species **(Supp.**
217 **Fig. 3**), which revealed behavioral syllables shared across species (discussed below). We also
218 found notable diversity in the number of syllables across species, ranging from 11 (LO) to32 (AB),
219 which is on par with or greater than the bout types that have been found in zebrafish larvae
220 (Marques et al., 2018; Mearns et al., 2020). Multiple factors may contribute to this threefold
221 difference in the number of behavioral syllables we identify between cichlid species. First, we note
222 that species with the most clusters identified (29 and 32 in LA and AB, respectively) performed
223 more hunting sequences than LO or NM (**Fig. 3D**). If the highest variation in swim kinematics
224 occurs during prey capture, we could be undersampling the behavioral repertoire of species such
225 as LO. Similarly, we could be undersampling the diversity of spontaneous swims in species that
226 spend a lot of time sitting on the bottom of the behavioral chamber. Alternatively, different cluster
227 numbers could represent real variation in the behavioral repertoire of species, reflecting different
228 hunting strategies. As noted above, LO and NM tend to hunt with a sit-and-wait strategy, and
229 therefore may perform fewer kinds of behavior syllables when darting toward prey. In contrast,
230 AB and LA are much more active hunters, and therefore may draw on a richer behavioral repertoire
231 to orient towards, pursue and switch between targets.

232 Similar to zebrafish, we found that cichlids have behavioral syllables that they preferentially
233 perform during prey capture (**Fig. 5F**). Some of these swims share characteristics with hunting
234 bouts previously described in zebrafish, such as J-turns and approach swims, but we also identify
235 new types that do not have clear analogs in zebrafish (**Fig. 5G**). For example, cichlid larvae often
236 "hover" in place, oscillating their tail without any forward movement (**Fig. 5G**). Such behaviors
237 tend to be longer lasting than approach swims. We speculate that hovers may aid to keep animals
238 in place during hunting by preventing them from being moved by convection currents, or by
239 breaking their forward momentum following more rapid forward swims. Alternatively they may
240 allow animals to pause between sensorimotor decisions during hunting sequences while keeping
241 the tail motile, reducing the latency of the motor system to respond to descending motor
242 commands. Interestingly, hovering is also present in some adult cichlids that use their lateral line

243 to sense prey in the subsurface (Schwalbe et al., 2016). During hunting, cichlids will alternate
244 between approaches and these hover swims.

245 Another feature of cichlid hunting behavior not present in zebrafish is tail coiling. Here, the tail
246 coils into an S-shape over many hundreds of milliseconds leading up to a strike. The coil is released
247 rapidly during the strike, allowing fish to spring toward the prey (**Fig. 5G,** bottom center **& 4I;**
248 **Supp. Vid. 2,** 00:12; **Supp. Vid. 4,** 00:31). In addition to such capture springs, we have also
249 observed suction captures (**Supp. Vid. 3**, 00:33) and ram-like attack swims (**Supp. Vid. 2**, 00:45)
250 in cichlids.

251 Finally, we found that cichlids exhibit a range of post-strike maneuvers not present in zebrafish.
252 These include swimming backwards to re-center prey in the visual field in cases where capture
253 strikes miss (**Supp. Vid. 8**), and expelling prey ("spitting") that has already been captured (**Supp.**
254 **Vid. 7**). Spitting behavior often triggered another attempt to hunt and capture the same prey, and
255 was particularly common in LA. This behavior likely explains the high incidence of hunt initiations
256 when prey are already in the central visual field in this species (**Fig. 4A,B&D**). The function of
257 spitting is not clear, and it might indicate that the prey are slightly too large for LA to swallow.
258 We also find that cichlid captures often occur in two phases, with the prey first being caught by
259 the pharyngeal jaws (a second, internal set of jaws present in cichlids; Liem 1973) before being
260 ingested. Together, these results reveal a richer behavioral repertoire for pursuing, capturing, and
261 reorienting towards missed prey in cichlids than in zebrafish.

262 <u>Cichlids share two common capture strikes that are distinct from medaka side-swing strikes</u>

263 Studies in invertebrates have demonstrated that behavior can evolve through changing the
264 kinematics of a behavior performed in a given sensory context (Ding et al., 2019), changing which
265 sensory cues drive behavior (Seeholzer et al., 2018), or changing the transition frequencies
266 between a common set of kinematic motifs (Hernández et al., 2021). It is not known to what extent
267 these factors play a role in the evolution of vertebrate behaviors, nor the evolutionary timescales
268 over which behaviors diverge to the point of becoming distinct. To address this question, we
269 focused on the capture strike kinematics of cichlids and medaka (OL). The capture strike of
270 zebrafish larvae has been extensively characterized, consisting of two distinct types: the attack
271 swim and S-strike. These strikes are deployed variably, dependent on experience and different
272 prey distances (Lagogiannis et al., 2020; McClenahan et al., 2012; Mearns et al., 2020; Patterson
273 et al., 2013). At the week-old stage, attack swims are the most common capture maneuver, and
274 consist of a symmetric oscillation of the tail that resembles spontaneous forward swims. During
275 an S-strike, the tail does not oscillate, but rather bends into an S-shape and then straightens rapidly,
276 causing larvae to lunge toward prey.

277 Medaka (OL) swimming is nearly continuous, and their eyes do not converge during prey capture.
278 Nevertheless, they may still perform different behaviors during spontaneous swimming and prey
279 capture. Investigating clusters of behavior in OL revealed types of swim that are similar to hunting

bouts of cichlids and zebrafish. These included J-turns, approach swims, and slow swims (**Fig. 5H**). The presence of these swims suggests that OL, like other species studied, may also have a suite of behaviors they preferentially perform during prey capture. Remarkably, we found that OL did not strike at prey in the central visual but rather captured prey in the lateral visual field with a side-swing behavior (**Fig. 5H-J**). Although the eyes do not converge, there is sometimes a slight nasalward rotation of the ipsilateral eye leading up to and during the side swing (**Fig. 5I**).

To test whether the side-swing behavior was unique to medaka (OL), we clustered capture strikes of cichlids and OL after projecting their tail kinematics into the common pose space, setting the number of clusters to the number of species (**Fig. 5A**). We reasoned that if bouts were similar within each species but different between species, cluster identity would correlate with species identity. In contrast, if species shared similar kinematic motifs, clusters would contain bouts from all species. We found that strikes from all cichlid species were similar to each other, while the side-swing strike was highly enriched in OL (**Fig. 5K**).

This analysis uncovered at least two kinematically distinct capture strikes in cichlids. The slower of these behaviors is similar to the zebrafish attack swim, characterized by a symmetric undulation of the tail about the midline which propels the fish towards the prey, while the faster corresponds to the capture spring. The S-shape of the tail leading up to the capture spring shares structural similarities with the S-strike of zebrafish larvae (Mearns et al., 2020), differing primarily in the time course over which they occur. The S-strike of zebrafish may represent an accelerated form of the capture spring, forming and releasing over tens of milliseconds rather than hundreds of milliseconds seen in cichlids or, vice versa, the capture spring may represent an augmentation of the S-strike.

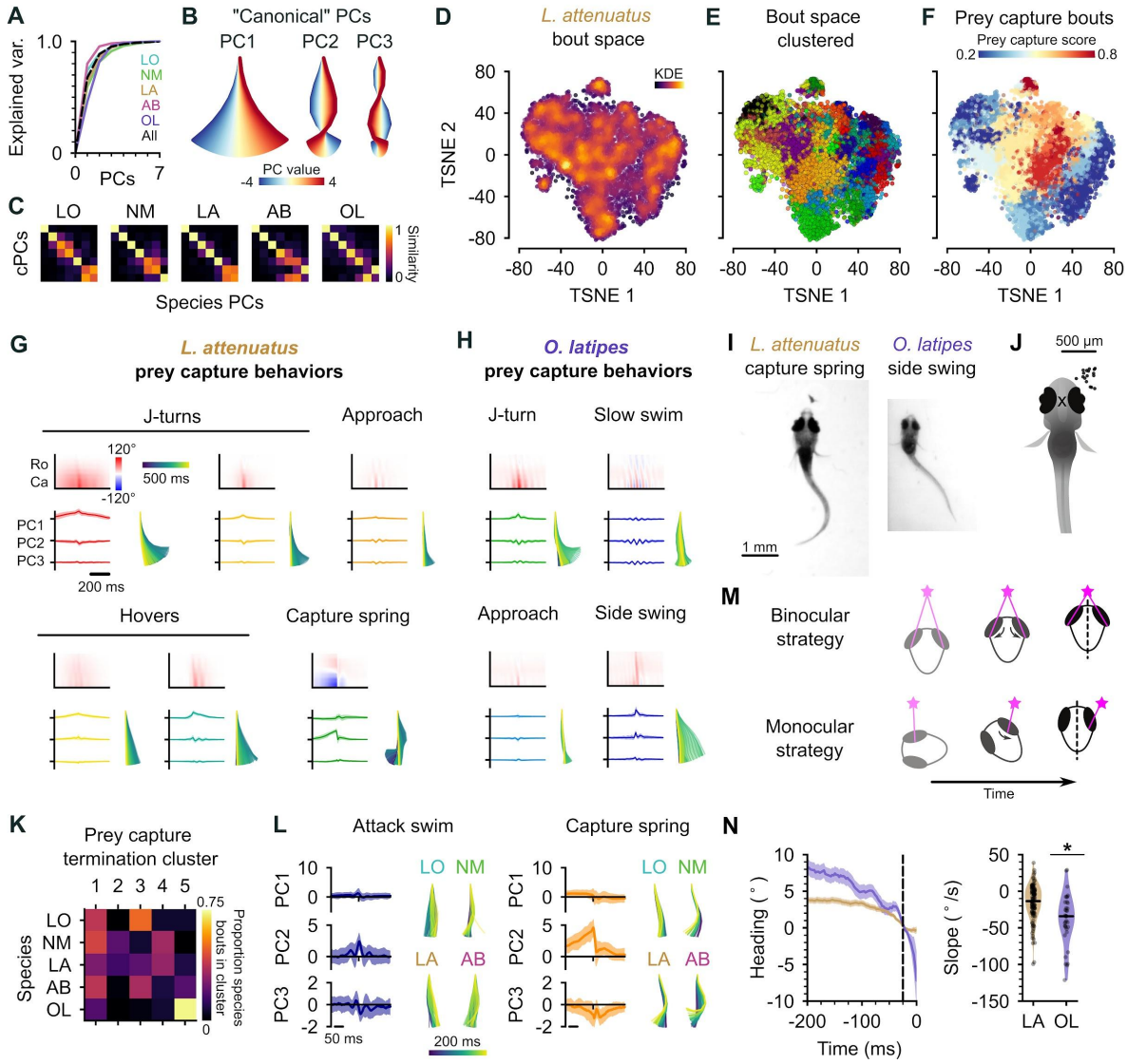


Figure 5. Interspecies comparison of prey capture strategies.

(**A**) Cumulative explained variance for the "canonical" principal components (PCs) obtained from all species (black dotted line), and PCs for each species individually (colored lines). In all cases, three PCs explain >90% of the variance in tail shape. (**B**) "Eigenfish" of the first three canonical PCs. Each principal component represents a vector of angles from the base to the tip of the tail (oriented with the fish facing up). At a given moment, the shape of the tail can be described as a linear combination of these vectors. Colors correspond to the tail shape obtained by scaling each PC from -4 to 4 standard deviations from the mean. (**C**) Each species' eigendecomposition compared against the canonical PCs computed for all species together. Color intensity represents the cosine similarity between pairs of vectors. The strong diagonal structure (particularly in the first three PCs) shows that similar PCs are obtained by analyzing species separately or together. (**D**) Behavioral space of *L. attenuatus*. Each point represents a single bout. Bouts are projected onto the first three PCs, aligned to the peak distance from the origin in PC space and then projected into a two-dimensional space using TSNE. Color intensity represents density of surrounding points in the embedding. (**E**) Clustered behavioral space of *L. attenuatus*. Clusters (colors) are computed via affinity propagation independently of the embedding. (**F**) Prey capture and spontaneous bouts in *L. attenuatus*. Prey capture score is the

probability that the eyes are converged at the peak of each bout. Bouts are colored according to the mean prey capture score for their cluster. Blue: clusters of bouts that only occur during spontaneous swimming; red: clusters of bouts that only occur during prey capture. (**G-H**) Example prey capture clusters from *L. attenuatus* (G) and medaka, *O. latipes* (H). For each cluster, top left: mean rostrocaudal bending of the tail over time; bottom left: time series of tail pose projected onto first three PCs (mean +/- standard deviation); bottom right: reconstructed tail shape over time for mean bout. (**I**) Representative frames of an *L. attenuatus* capture spring (left) and *O. latipes* (medaka) side swing (right), highlighting differences in tail curvature between these behaviors. (**J**) Location of prey in the visual field (black dots) immediately prior to the onset of a side swing. All events mirrored to be on the right. X marks the midpoint of the eyes, aligned across trials. (**K**) Comparison of clustered hunt termination bouts from all species. Termination bouts from all species were sorted into five clusters based on their similarity. Rows show the proportion of bouts from each species that were assigned to each cluster (columns). Cichlid bouts are mixed among multiple clusters, while medaka bouts (OL) mostly sorted into a single cluster. (**L**) Capture strikes in cichlids. Representative examples of tail kinematics during attack swims (left) and capture springs (right) from each species, including time series of tail pose projected onto the first three PCs (mean +/- standard deviation, all species combined) shown for each type of strike. (**M**) Two hypotheses for distance estimation make different predictions of how heading (black dotted line) changes over time as fish approach prey (pink star). Top: fish maintain prey in the central visual field and use binocular cues to judge distance. Bottom: fish "spiral" in towards prey, using motion parallax to determine distance. Black arrows indicate motion of prey stimulus across the retina. (**N**) Change in heading over time leading up to a capture strike for *L. attenuatus* (LA, brown, n=113) and medaka (OL, purple, n=36). Left: time series of heading. Zero degrees represents the heading 25 ms prior to the peak of the strike (black dotted line). Mean ± s.e.m. across hunting events. Right: comparison of the rate of heading change, computed as the slope of a line fit to each hunting event 200-25 ms prior to the peak of the strike. The heading decreases more rapidly leading up to a strike in medaka than in *L. attenuatus*.

302 <u>Different prey capture strategies may reflect different sensorineural solutions to prey detection</u>

303 Centralized prey and converged eyes in cichlids suggest that the mechanism they use to determine
304 the distance to prey could be similar to that of zebrafish. One possibility is that these species use
305 binocular disparity to judge depth (Qian, 1997). In contrast, we hypothesized that medaka (OL)
306 might use a different, monocular strategy to judge the distance to prey, such as motion parallax
307 (Yoonessi & Baker, 2011). These two mechanisms of depth perception make different predictions
308 about how larvae should approach prey to make best use of visual cues (**Fig. 5M**). To investigate
309 these possibilities, we computed the change in heading leading up to a strike as a proxy for the
310 change in visual angle of the prey for cichlids and medaka (**Fig. 5N**). We found that the rate of
311 change in heading was greater in medaka than in cichlids (median degrees/second: -13.4, LA; -
312 34.1, OL; $p < 0.001$, Mann-Whitney U-test). These results are consistent with cichlids using
313 binocular cues, such as binocular disparity, to determine distance to prey, while medaka may
314 employ a monocular strategy.

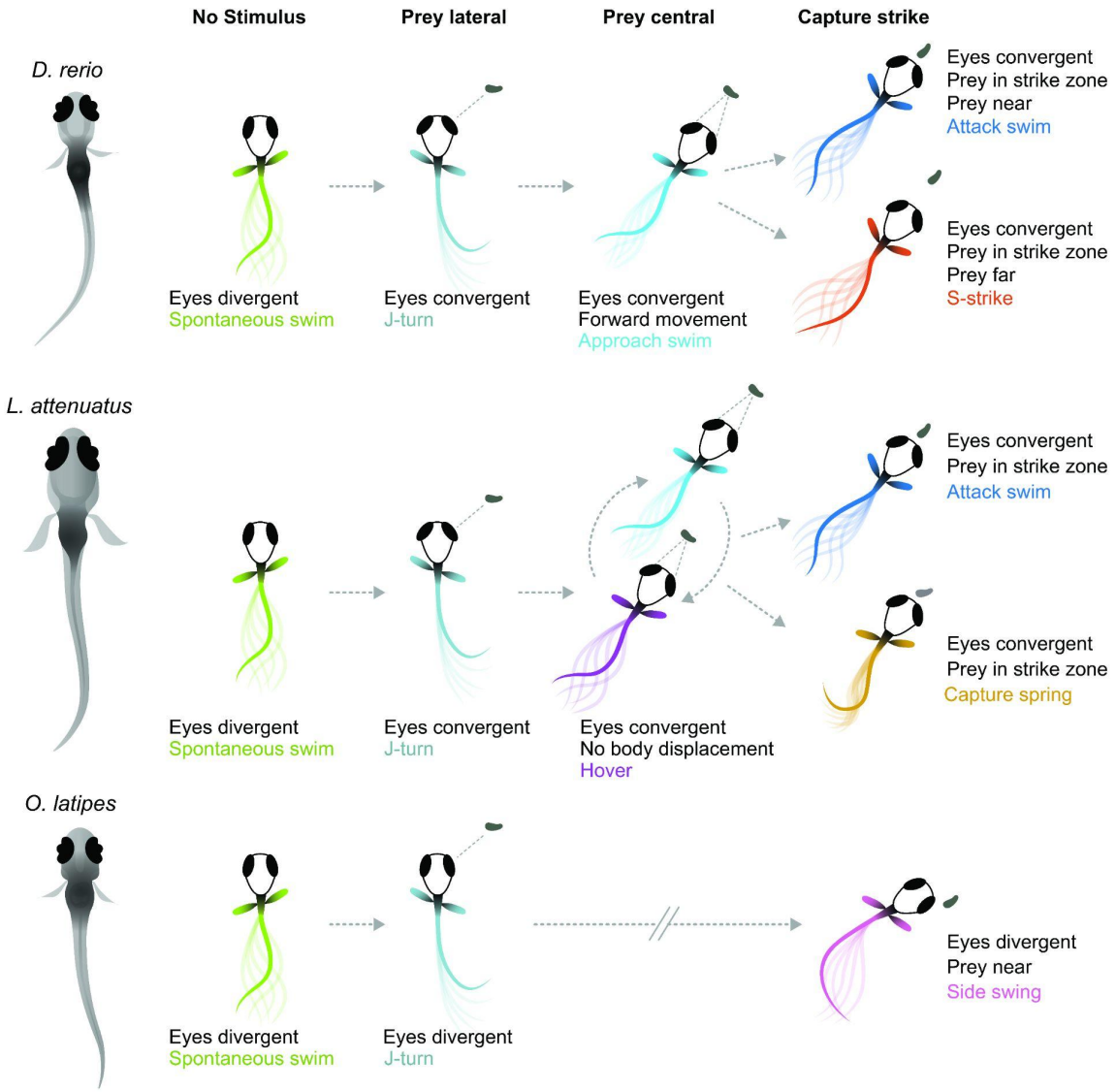


Figure 6. Schematic of hunting strategies in teleost larvae.
Top: prey capture in zebrafish larvae (*Danio rerio*) begins with eye convergence and a J-turn to orient towards the prey. The prey is then approached with a series of low amplitude approach swims. Once in range in the central visual field, the prey is captured with an attack swim or an S-strike. Middle: prey capture in cichlids (represented by *L. attenuatus*) also begins with eye convergence and a J-turn. Prey is approached with a wide variety of tail movements. The prey is captured when it is in the central visual field with either an attack swim, or a capture spring, during which the tail coils over several hundreds of milliseconds. Bottom: prey capture in medaka (*O. latipes*) begins with reorienting J-like turns, but these do not centralize prey in the visual field. Instead, the prey is kept lateral in the visual field and is captured with a side swing.

Discussion

Here we have provided the first description of the hunting behavior of larval cichlids and medaka, popular model systems for understanding the neural and genetic basis of behavior (Shima & Mitani, 2004; Jordan et al., 2021; Johnson et al., 2023). We have found that cichlids share the same general hunting strategy to zebrafish: striking at prey localized to a binocular strike zone while the eyes are converged, with two distinct capture bout types whose kinematics are remarkably similar given these species diverged 240 million years ago (see **Fig. 1A**). In contrast, medaka, which are phylogenetically intermediate to zebrafish and cichlids, deploy a different hunting strategy: striking at prey laterally in the monocular visual field with a side-swing behavior. **Fig. 6** provides a schematic summary of similarities and differences among the teleost larvae investigated. Characterization of prey capture behavior in a wider range of species could reveal whether the similarities between zebrafish and cichlids represent conservation of an ancestral hunting strategy or convergent evolution.

What sensorineural mechanisms might underlie these two different hunting strategies? The side swing of medaka is reminiscent of hunting in (adult) blind cavefish (Lloyd et al., 2018), which position prey laterally prior to strikes using their mechanosensory lateral line. Adult cichlids are known to use their lateral line for prey capture (Schwalbe et al., 2012), and some percomorph species exhibit both an S-strike-like and a sideways capture strategy as adults (New, 2002) Therefore, it is conceivable that medaka larvae do not rely on vision to the same extent for hunting, but use their lateral line instead. However, lateral-line use is not uncoupled from eye convergence across the teleost clade. Zebrafish larvae can still hunt in the dark, as do blind zebrafish *lakritz* mutants, albeit with greatly reduced efficacy, and with their eyes converged (Gahtan et al., 2005; Mearns et al., 2020; Patterson et al., 2013). Early-stage cave fish, which evolved from sighted surface ancestors and are completely blind, exhibit vestigial eye convergence movements during prey capture before their eyes fully degenerate (Espinasa et al., 2023; Espinasa & Lewis, 2023). This comparative evidence suggests that the teleost brain employs cues from both sensory modalities, if available, to locate prey, but that eye movements are a poor predictor of the dominance of vision.

On the other hand, species that attack centrally located prey could be using different visual cues to judge prey distance than those that use a side swing to ingest food: binocular disparity (Qian, 1997) vs. motion parallax (Yoonessi & Baker, 2011). Zebrafish use binocular information (Gahtan et al., 2005; Gebhardt et al., 2019; Henriques et al., 2019) and stationary differences such as brightness and contrast (Khan et al., 2023) to estimate depth and distance. Here we have shown that cichlid larvae also likely use binocular cues, while medaka's approach to prey conforms to a monocular strategy (see **Fig. 5M,N**). Strikingly, cichlid and zebrafish larvae swim in bouts with intermittent pauses, while medaka swim continuously with longer bouts ("gliding"). We speculate that the pauses may serve to sample the distance to the prey by comparing its position across the two eyes. In contrast, to sample motion cues, it is advantageous to perform continuous glides and swim

lateral to the prey up until it is close enough for a sideways strike. Thus, natural selection may have favored specific locomotor (bouts vs. glides) and oculomotor (convergent vs. divergent) adaptations of behavioral control depending on the dominant distance measurement mechanism (binocular vs. monocular) (**Fig. 6**). While the behavioral evidence for this conjunction is still circumstantial and calls for further comparative work, this scenario makes testable predictions about its neurobiological and genetic implementation. Cichlids and medaka may also use size, brightness and contrast cues to estimate prey distance in addition to the binocular disparity and monocular parallax mechanisms we have postulated here.

We have demonstrated that eye convergence is not always a hallmark of prey capture in fish larvae, and that multiple, kinematically distinct capture swims exist at these early stages in different species. We speculate that ocular vergence angles and swim kinematics for prey capture might co-evolve, with S-strike-like captures and attack swims being associated with convergent eyes and side-swing captures occuring in the absence of eye convergence. These two strategies may have been present in a common ancestor of teleosts, with one or the other becoming dominant in certain lineages. Comparing neural circuitry across these (now) experimentally tractable animals could help pinpoint evolvable circuit nodes ("hotspots" of evolutionary change; Roberts et al., 2022; Seeholzer et al., 2018) and genetic loci that underlie prey localization strategies and swim kinematics in fishes.

Key Resources Table

Reagent type (species) or resource	Designation	Source or reference	Additional information
Lamprologus ocellatus	LO	Other	From Alex Jordan, Max Planck Institute of Animal Behavior, Konstanz
Neolamprologus multifasciatus	NM	Bose & Windorfer et al. (2020)	From Alex Jordan, Max Planck Institute of Animal Behavior, Konstanz
Lepidiolamprologus attenuatus	LA	Other	From Alex Jordan, Max Planck Institute of Animal Behavior, Konstanz
Astatotilapia burtoni	AB	Other	From Alex Jordan, Max Planck Institute of Animal Behavior, Konstanz
Oryzias latipes (wild-type)	Medaka; OL	Koester, et al. (1997)	From Joachim Wittbrodt, University of Heidelberg
Social LEAP Estimates Animal Pose	SLEAP	Pereira et al. (2019); Pereira et al. (2022)	
You Only Look Once	YOLO	Redmon et al. (2016)	
Python code		Mearns et al. (2020)	
Custom analysis code (Python 3)		This paper	
Original data		Mearns, Duncan; Hunt, Sydney; Baier, Herwig (2024), "Diverse prey capture strategies in teleost larvae", Mendeley Data, V1, doi: 10.17632/m6zs9bm7w3.1	

Methods

Experimental system

The following species were used in this study: *Lamprologus ocellatus* (n = 14), *Lepidiolamprologus attenuatus* (n = 12), *Neolamprologus multifasciatus* (n = 22) and *Astatotilapia burtoni* (n = 14); medaka, *Oryzias latipes* (n = 3). Adult fish were raised at the Max Planck Institute for Biological Intelligence. The cichlid colonies were originally provided by the Max Planck Institute for Animal Behavior, Konstanz, Germany. *L. ocellatus* were housed in 50 liter tanks with one male and two females and a shell for each individual for egg laying. *L. attenuatus* were housed in groups (10-15 mix-sexed individuals) within a 50 l tank with a half flower pot for egg laying. *N. multifasciatus* were housed in large groups (20-30 mix-sexed individuals) within a 160 l tank with shells for egg laying. *A. burtoni* were housed in 160 l tanks in groups of 15 (1 male and 14 females) with plants to encourage courtship arenas. *O. latipes* were raised like zebrafish. All behavioral observations were non-invasive and were performed in accordance with applicable laws and governmental regulations.

Adult cichlids were fed live artemia, kept under a 13 hour:11 hour day-night cycle in 27°C water (pH 8.2, conductivity ~550 µS) on sand substrate. Egg-laying behavior was monitored, and clutches were collected at 7-8 days post-fertilization (dpf) from the shell (*L. ocellatus*, *N. multifasciatus*), the flower pot half (*L. attenuatus*), or gently released from the mouth of the mother (*A. burtoni*). *O. latipes* embryos were harvested from natural crosses. Clutches were incubated at 28°C in sterile petri dishes with filtered fish facility water. Larvae were assayed as soon as they displayed robust feeding behaviors: *L. ocellatus* and *L. attenuatus* at 12 dpf, *N. multifasciatus* at 14 dpf, *A. burtoni* at 14-16 dpf, and *O. latipes* at 10 dpf.

Video acquisition

High-speed (300-500 frames per second) videography data of single larvae of each species were collected as described previously (Mearns et al., 2020). Behavior chambers were molded out of 2% agarose and measured 30 mm x 30 mm for cichlids or 20 mm x 20 mm for medaka. Single larvae were introduced into chambers with a drop of food culture (*Artemia salina* for cichlids, ~30 each; *Paramecium multimicronucleatum* (Carolina Biological Supply Company, Burlington, NC) for medaka). Each animal was recorded for 10-15 minutes using a high-speed camera (EoSens CL MC1362, Mikrotron; objective: Sigma 50 mm f/2.8 ex DG Macro, Japan), backlit with a custom-built infrared LED array.

Tracking

To track the tail and eyes we trained neural networks using SLEAP (Social Leap Estimates Animal Poses) (Pereira et al., 2022). We designed a 12-point skeleton to track the tail (left eye center, right eye center, rostral tip, swim bladder and eight points along the tail), and a 7-point skeleton to track the eyes (swim bladder, and a nasal, central and temporal point on the perimeter of each eye).

Tail tracking: For tail tracking, a single instance model was trained for *L. attenuatus* on a total of 701 labeled frames from 8 videos across 8 fish. 290 frames were manually labeled and 411 were corrected across two rounds of human-in-the-loop training. This model was used to track the tail in *L. ocellatus* and *N. multifasciatus* in addition to *L. attenuatus*. To track *A. burtoni* and medaka,

the *L. attenuatus* base model was trained with an additional 345 and 190 frames, respectively, from each of these species.

Eye tracking: For the eyes, we first tracked 4,900 frames of *L. attenuatus* behavior using an unsupervised tracking algorithm. These points were used to train a base model. This base model was further refined for each species using human-in-the-loop on an additional 402 frames, *A. burtoni*; 331 frames, *O. latipes*; and 314 frames, *L. ocellatus* (this model was also used to track the eyes in *N. multifasciatus*).

Prey tracking: Artemia were tracked using a single neural network developed for object detection, YOLO (You Only Look Once) (Redmon et al., 2016). We trained the model on 128 frames. Due to the lower prey contrast in the *A. burtoni* and medaka videos, we could not track prey location in these videos. Prey locations prior to the medaka swing were manually annotated in 17 frames where the targeted paramecium was clearly visible.

To obtain continuous tracks of the most likely targeted prey, we tracked prey location across all frames of episodes of eye convergence. We tracked identities to single artemia across frames using the Hungarian algorithm. We assumed that the artemia nearest to the larva in the anterior visual field in the 100 ms leading up to eye deconvergence was the targeted prey.

Statistical analysis

All analysis was performed using custom-written Python code (libraries: numpy, pandas, scikit-learn, scipy, opencv).

We used $\alpha = 0.05$ with a Bonferroni correction when assessing significance. In figures 2 and 4 we used the Mann-Whitney U-test from the scipy.stats module. For bootstrap tests in figures 1 and 3, we compared the difference in median between groups and computed the probability of observing this value or greater under the null hypothesis. To build distributions under the null hypothesis, we recomputed the test statistic from median-adjusted data (sampling 10% of the data with replacement) 100,000 times.

Eye convergence analysis

Computing eye convergence angles: The angle of each eye in each frame was calculated as the angle between heading vector of the fish (vector between the midpoint of the finsand midpoint of the two eyes) and the long axis of the eye (vector between the temporal and nasal points). Zero signifies the long axis of the eye is parallel with the heading vector of the fish; positive angles signify nasalward rotation. The convergence angle was computed as the sum of these two angles (positive convergence indicates both eyes are rotated inwards towards the midline). Eye angles of $< -10°$ (diverged) or $> 60°$ converged indicated tracking errors and were excluded from further analysis.

Gaussian mixture modeling: To determine when the eyes were converged, we fitted a Gaussian mixture model to the distribution of convergence angles for each species. We tested models with either one or two mixtures to determine whether any discernible eye convergence was present. We used a Bayesian Inference Criterion to determine whether one or two mixtures resulted in a better

fit. For all cichlid species, a two-mixture model outperformed a single Gaussian distribution. Convergence phases coincided with hunting events. For medaka (*O. latipes*), a two-mixture model performed marginally worse than the single distribution model.

Identification of hunting episodes: To identify periods of hunting in cichlids, we computed the probability of eye convergence in each frame as the probability it belonged to the Gaussian mixture with the higher mean. This yielded a time series of eye convergence probability, which we smoothed with a 200 ms rolling mean. We defined a hunting episode as any continuous period where the probability of eye convergence was above 50%, with the additional condition that that convergence probability was above 80% for at least half this period.

Tail kinematic analysis

We tracked eight equally spaced points from the tail base to the tail tip in each frame using the whole body SLEAP model (see above). In each frame, we represented the shape of a tail as a seven-dimensional vector, computed by calculating the angle between a line through each consecutive pair of tail points and the heading vector (from the midpoint of the fins to the midpoint between the eyes).

Principal component analysis (PCA) and data curation: To reduce the dimensionality of pose data, we performed PCA as described previously (Mearns et al., 2020). We concatenated the seven-dimensional tail vectors in time (all animals pooled, grouped by species, or for all species combined). We z-scored the data matrix prior to PCA by subtracting the mean angle at each point and dividing by the standard deviation. This ensures equal weighting of all points of the tail during PCA, since otherwise loadings in the principal components would be biased towards points closer to the tail tip, potentially missing subtle but important curvature in the rostral trunk.

PCA provides a set of principal components (PCs; orthogonal vectors, representing some loading across the input dimensions), each associated with an eigenvalue (variance in the data captured by that component). Each principal component can be represented as an "eigenfish" by mapping the raw loadings multiplied by some scalar back into the original space (Mearns et al., 2020). In all cases, we found that three components were sufficient to capture over 90% of the variance in the data, and so for subsequent analyses we represented tail kinematics by a time varying three-dimensional vector representing the projection of the tail shape onto the first three PCs in each frame.

Since we were unable to track every tail point in every frame of our data, we performed some interpolation to fill in missing frames in our tracking. For any period where there was a gap of four or fewer frames (≤ 10 ms) of missing tracking data, we inferred the tail shape using a cubic spline interpolation in each PC dimension. Gaps greater than four frames were excluded from our analyses. We found that tracking errors were easiest to detect by inspecting the weight of the third PC in each frame. Therefore, we excluded frames where the absolute value of this PC exceeded 10 standard deviations (or 5, in the case of *A. burtoni*).

We linearly interpolated tail kinematics so that all trials had a common sample rate in cases where the frame rates of videos did not match.

Segmentation: We adapted our previously published bout detection algorithm for zebrafish larvae (Mearns et al., 2020) to segment the more continuous swimming behavior of the larval species used in this study. This algorithm involves five steps:

1. The tail is represented as a single-dimensional time series by computing the L2 norm in PC space for each frame.

2. The absolute value of the derivative of this trace was convolved with a Gaussian kernel (St. Dev: 40 ms). This produces a smooth trace with peaks whenever the shape of the tail is changing rapidly in a non-cyclic manner (i.e., at transitions between behaviors).

3. A threshold was automatically computed by performing a kernel density estimation across all values of this trace (bandwidths: 0.005, NM, AB & OL; 0.01, LA; 0.02, LO). If this generated a bimodal distribution, we set the threshold to the antimode (local minimum) of this distribution; otherwise, we set the threshold to the first inflexion point after the mode.

4. Peaks in the traces (step 2) that were above the threshold (step 3) were detected (scipy.signal.find_peaks function). Behavioral syllables ("bouts") were segmented at local minima between these peaks.

5. We performed a round of cleaning to exclude noisy bouts. For each bout, we computed the standard deviation in each PC dimension (low standard deviations in every dimension suggest no tail movement but rather noise in tracking). We then fit a Gaussian mixture model to the three-dimensional space of these standard deviations, and excluded bouts predicted to belong to the mixture with the smallest standard deviation across PC dimensions. For each species we manually set the number of mixture components to exclude most noisy bouts while keeping real movement in the dataset.

This pipeline identified the following numbers of bouts in each species: 2,067 (LO); 11,237 (NM); 10,445 (LA); 9,796 (AB); 3,437 (OL).

Alignment: Having segmented behavior into individual bouts, we sought to align these bouts for clustering and embedding. Our segmentation algorithm produced bouts of varying duration; however, we reasoned that the most important features of a bout that would distinguish it from others would occur when the tail is in the "most extreme" or characteristic position, so we aligned all bouts to their peak distance in PC space (see segmentation step 1). Then, we truncated or zero-padded the aligned bouts such that we fully encompassed 80% of bout starts (before the peak) and 80% of bout ends (after the peak).

To aid the identification of behavioral clusters we chose to ignore the left/right orientation of bouts. Therefore, we flipped bouts such that at their peaks the tail was always oriented in the same direction. We performed an initial clustering step using ward hierarchical clustering with 30 clusters. For this clustering step, we included each bout and its mirror image to produce symmetric clusters. To determine the directionality of a bout, we computed the mean value of the first PC at the peak of each cluster and used either the original or mirrored version of the bout accordingly.

We found flipping bouts based on cluster was more robust than flipping each bout individually, since in many instances the value of the first PC is close to zero at the bout peak, and we wanted to ensure all similar bouts were oriented in the same direction to aid embedding and subsequent clustering.

Clustering: To identify behavioral clusters, we performed affinity propagation on the aligned, flipped kinematic motifs. We treated each motif as a single vector, concatenating the time series for each PC. Affinity propagation computes a set of clusters (represented by "exemplar" data points) for a given dataset at a given preference value. Lower preference values produce fewer clusters, so we used the minimum similarity between data points, which is a standard metric for computing a preference value (Frey & Dueck, 2007). This approach yielded between 11 and 32 clusters per species (11, LO; 23, NM; 29, LA; 32, AB; 17, OL), which is a reasonable number to expect in these data, being the same order as the number of behaviors that have been identified in zebrafish larvae (Marques et al., 2018).

Embedding: To visualize behavior maps, we embedded motifs in a two-dimensional space using t-distributed stochastic neighbor embedding (t-SNE) (Maaten & Hinton, 2008), concatenating the time series for each PC such that each motif was represented by a single vector.

Prey capture bouts: To assign bouts a prey capture score, we computed the probability of eye convergence (see *eye convergence analysis*, above) at the peak of each bout. We then averaged these probabilities over all bouts belonging to a given cluster to compute the prey capture score for that cluster. Thus, a prey capture score of zero signifies zero probability of eye convergence for all bouts within a cluster, and a prey capture score of one indicates a 100% probability of eye convergence for all bouts within a cluster.

Capture strike analysis

Identifying putative strikes: To identify putative capture strikes in cichlids, we found all bouts whose peak coincided with the end of a hunting episode (see *eye convergence analysis*, above), allowing the peak to occur up to 100 ms before and up to 200 ms after the detected end of eye convergence. Since we could not use eye convergence to detect prey capture in medaka, we manually inspected clusters obtained from analyzing tail movements, and found that one of these clusters corresponded to a strike behavior (side swing). For cross-species comparisons, all data were interpolated to a common frame rate of 400 fps. We performed an initial round of clustering on these putative strike bouts (Ward hierarchical clustering with n clusters = 8) and removed small clusters containing fewer than 20 bouts. After this filtering, we had 341 putative strike bouts (16, LO; 31, NM; 191, LA; 59, AB; 44, OL).

Clustering strikes: To only consider tail kinematics around the moment of the strike, we truncated bouts to a 150 frame window (375 ms) around the peak. We then clustered these bouts with Ward hierarchical clustering, with the number of clusters set to five (same as the number of species). To determine whether each species had its own unique strike behavior, or whether the same behaviors were shared across species, we computed the confusion matrix using the cluster labels and species identity labels. Two clusters corresponded to abort behaviors in cichlids, two clusters corresponded to cichlid capture strikes (attack swim and capture spring), and one cluster corresponded to medaka side swings.

Change in heading during prey capture: To compute the change in heading leading up to a strike, we considered a 200 ms window before the peak of each strike bout. So as to not bias our analysis to the rapid changes in heading at the moment of the medaka side swing, we only analyzed heading changes up to 25 ms before the peak of each bout. We virtually rotated and mirrored all capture sequences such that the heading 25 ms before the bout peak was zero, and the mean heading prior to this moment was positive. For each event, we then computed the slope of a regression line (least squares fit).

Author contributions

D.S.M. and H.B. conceived the project. D.S.M. and M.W.S. collected the data. S.A.H. trained models and performed tracking. A.V.P. and M.S. maintained cichlid colonies and provided larvae for experiments. D.S.M. and S.A.H. analyzed the data. D.S.M. and H.B. wrote the manuscript with input from all authors.

Competing interests

No competing interests declared.

Funding

Funding was provided by the Max Planck Society. M.W.S. was supported by a Boehringer Ingelheim Fonds Fellowship.

Acknowledgements

We thank Alex Jordan for sharing cichlid stocks and Joachim Wittbrodt for sharing medaka stocks; Jessica Zung, Qing Wang and Greg Marquart for critical feedback on the manuscript; Abdelrahmen Adel for technical assistance training neural networks; Krasimir Slanchev and the animal caretakers at the Max Planck Institute for Biological Intelligence for assistance raising and breeding fish; and Julia Kuhl for illustrations.

References

Baier, H., & Scott, E. K. (2009). Genetic and optical targeting of neural circuits and behavior—Zebrafish in the spotlight. *Current Opinion in Neurobiology*, *19*(5), 553–560. https://doi.org/10.1016/j.conb.2009.08.001

Baier, H., & Scott, E. K. (2024). The Visual Systems of Zebrafish. *Annual Review of Neuroscience*, *47*, 255–276. https://doi.org/10.1146/annurev-neuro-111020-104854

Berman, G. J., Choi, D. M., Bialek, W., & Shaevitz, J. W. (2014). Mapping the stereotyped behaviour of freely moving fruit flies. *Journal of The Royal Society Interface*, *11*(99), 20140672. https://doi.org/10.1098/rsif.2014.0672

Betancur-R, R., Wiley, E. O., Arratia, G., Acero, A., Bailly, N., Miya, M., Lecointre, G., & Ortí, G. (2017). Phylogenetic classification of bony fishes. *BMC Evolutionary Biology*, *17*(1), 162. https://doi.org/10.1186/s12862-017-0958-3

Bianco, I. H., Kampff, A. R., & Engert, F. (2011). Prey Capture Behavior Evoked by Simple Visual Stimuli in Larval Zebrafish. *Frontiers in Systems Neuroscience*, *5*, 101. https://doi.org/10.3389/fnsys.2011.00101

Borla, M. A., Palecek, B., Budick, S., & O'Malley, D. M. (2002). Prey Capture by Larval Zebrafish: Evidence for Fine Axial Motor Control. *Brain Behavior and Evolution*, *60*(4), 207–229. https://doi.org/10.1159/000066699

Budick, S. A., & O'Malley, D. M. (2000). Locomotor repertoire of the larval zebrafish: Swimming, turning and prey capture. *The Journal of Experimental Biology*, *203*(Pt 17), 2565–2579. https://doi.org/10.1242/jeb.203.17.2565

Devineni, A. V., & Scaplen, K. M. (2022). Neural Circuits Underlying Behavioral Flexibility: Insights From Drosophila. *Frontiers in Behavioral Neuroscience*, *15*. https://www.frontiersin.org/articles/10.3389/fnbeh.2021.821680

Ding, Y., Lillvis, J.L., Cande, J., Berman, G.J., Arthur, B.J., Long, X., Xu, M., Dickson, B.J., Stern, D.L., 2019. Neural Evolution of Context-Dependent Fly Song. Current Biology 29, 1089-1099.e7. https://doi.org/10.1016/j.cub.2019.02.019

El Taher, A., Böhne, A., Boileau, N., Ronco, F., Indermaur, A., Widmer, L., & Salzburger, W. (2021). Gene expression dynamics during rapid organismal diversification in African cichlid fishes. *Nature Ecology & Evolution*, *5*(2), 243–250. https://doi.org/10.1038/s41559-020-01354-3

Espinasa, L., Diamant, R., Vinepinsky, E., & Espinasa, M. (2023). Evolutionary modifications of Astyanax larval prey capture (LPC) in a dark environment. *Zoological Research*, *44*(4), 750–760. https://doi.org/10.24272/j.issn.2095-8137.2022.466

Espinasa, L., & Lewis, K.-A. (2023). Eye convergence is evoked during larval prey capture (LPC) without visual stimulus and in blind cavefish. *Subterranean Biology*, *46*, 47–60. https://doi.org/10.3897/subtbiol.46.105707

Frey, B. J., & Dueck, D. (2007). Clustering by Passing Messages Between Data Points. *Science*, *315*(5814), 972–976. https://doi.org/10.1126/science.1136800

Friedrich, R. W., Jacobson, G. A., & Zhu, P. (2010). Circuit Neuroscience in Zebrafish. *Current Biology*, *20*(8), R371–R381. https://doi.org/10.1016/j.cub.2010.02.039

Gahtan, E., & Baier, H. (2004). Of lasers, mutants, and see-through brains: Functional neuroanatomy in zebrafish. *Journal of Neurobiology*, *59*(1), 147–161. https://doi.org/10.1002/neu.20000

Gahtan, E., Tanger, P., & Baier, H. (2005). Visual prey capture in larval zebrafish is controlled by identified reticulospinal neurons downstream of the tectum. *The Journal of*

Neuroscience: The Official Journal of the Society for Neuroscience, *25*(40), 9294–9303. https://doi.org/10.1523/JNEUROSCI.2678-05.2005

Gebhardt, C., Auer, T. O., Henriques, P. M., Rajan, G., Duroure, K., Bianco, I. H., & Del Bene, F. (2019). An interhemispheric neural circuit allowing binocular integration in the optic tectum. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-13484-9

Girdhar, K., Gruebele, M., & Chemla, Y. R. (2015). The Behavioral Space of Zebrafish Locomotion and Its Neural Network Analog. *PLOS ONE*, *10*(7), e0128668. https://doi.org/10.1371/journal.pone.0128668

Henriques, P. M., et al. (2019). Nucleus isthmi is required to sustain target pursuit during visually guided prey-catching. *Current Biology 29*(11), 1771-1786.

Hernández, D. G., Rivera, C., Cande, J., Zhou, B., Stern, D. L., & Berman, G. J. (2021). A framework for studying behavioral evolution by reconstructing ancestral repertoires. *eLife*, *10*, e61806. https://doi.org/10.7554/eLife.61806

Higham, T. E., Hulsey, C. D., Říčan, O., & Carroll, A. M. (2007). Feeding with speed: Prey capture evolution in cichilds. *Journal of Evolutionary Biology*, *20*(1), 70–78. https://doi.org/10.1111/j.1420-9101.2006.01227.x

Johnson, R. E., Linderman, S., Panier, T., Wee, C. L., Song, E., Herrera, K. J., Miller, A., & Engert, F. (2020). Probabilistic Models of Larval Zebrafish Behavior Reveal Structure on Many Scales. *Current Biology*, *30*(1), 70-82.e4. https://doi.org/10.1016/j.cub.2019.11.026

Khan, B., et al. (2023). Zebrafish larvae use stimulus intensity and contrast to estimate distance to prey. *Current Biology 33* (15), 3179-3191.

Liem, K. F. (1973). Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws. *Systematic Zoology 22*(4), 425-441.

Lloyd, E., Olive, C., Stahl, B. A., Jaggard, J. B., Amaral, P., Duboué, E. R., & Keene, A. C. (2018). Evolutionary shift towards lateral line dependent prey capture behavior in the blind Mexican cavefish. *Developmental Biology*, *441*(2), 328–337. https://doi.org/10.1016/j.ydbio.2018.04.027

Maaten, L. van der, & Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research*, *9*(86), 2579–2605.

Mano, H., & Tanaka, Y. (2012). Size specificity of predation by Japanese medaka Oryzias latipes on Daphnia pulex. *Journal of Freshwater Ecology*, *27*(2), 309–313. https://doi.org/10.1080/02705060.2012.659222

Marques, J. C., Lackner, S., Félix, R., & Orger, M. B. (2018). Structure of the Zebrafish Locomotor Repertoire Revealed with Unsupervised Behavioral Clustering. *Current Biology: CB*, *28*(2), 181-195.e5. https://doi.org/10.1016/j.cub.2017.12.002

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), Article 9. https://doi.org/10.1038/s41593-018-0209-y

McElligott, M. B., & O'Malley, D. M. (2005). Prey Tracking by Larval Zebrafish: Axial Kinematics and Visual Control. *Brain Behavior and Evolution*, *66*(3), 177–196. https://doi.org/10.1159/000087158

Mearns, D. S., Donovan, J. C., Fernandes, A. M., Semmelhack, J. L., & Baier, H. (2020). Deconstructing Hunting Behavior Reveals a Tightly Coupled Stimulus-Response Loop.

Current Biology, *30*(1), 54-69.e9. https://doi.org/10.1016/j.cub.2019.11.022

New, J. G. (2002). Multimodal Integration in the Feeding Behaviors of Predatory Teleost Fishes. *Brain Behavior and Evolution*, *59*(4), 177–189. https://doi.org/10.1159/000064905

Orger, M. B., & de Polavieja, G. G. (2017). Zebrafish Behavior: Opportunities and Challenges. *Annual Review of Neuroscience*, *40*(1), 125–147. https://doi.org/10.1146/annurev-neuro-071714-033857

Patterson, B. W., Abraham, A. O., MacIver, M. A., & McLean, D. L. (2013). Visually guided gradation of prey capture movements in larval zebrafish. *The Journal of Experimental Biology*, *216*(Pt 16), 3071–3083. https://doi.org/10.1242/jeb.087742

Pereira, T. D., Aldarondo, D. E., Willmore, L., Kislin, M., Wang, S. S.-H., Murthy, M., & Shaevitz, J. W. (2019). Fast animal pose estimation using deep neural networks. *Nature Methods*, *16*(1), Article 1. https://doi.org/10.1038/s41592-018-0234-5

Pereira, T. D., Tabris, N., Matsliah, A., Turner, D. M., Li, J., Ravindranath, S., Papadoyannis, E. S., Normand, E., Deutsch, D. S., Wang, Z. Y., McKenzie-Smith, G. C., Mitelut, C. C., Castro, M. D., D'Uva, J., Kislin, M., Sanes, D. H., Kocher, S. D., Wang, S. S.-H., Falkner, A. L., … Murthy, M. (2022). SLEAP: A deep learning system for multi-animal pose tracking. *Nature Methods*, *19*(4), Article 4. https://doi.org/10.1038/s41592-022-01426-1

Piggott, B. J., Liu, J., Feng, Z., Wescott, S. A., & Xu, X. Z. S. (2011). The Neural Circuits and Synaptic Mechanisms Underlying Motor Initiation in C. elegans. *Cell*, *147*(4), 922–933. https://doi.org/10.1016/j.cell.2011.08.053

Portugues, R., & Engert, F. (2011). Adaptive Locomotor Behavior in Larval Zebrafish. *Frontiers in Systems Neuroscience*, *5*. https://www.frontiersin.org/articles/10.3389/fnsys.2011.00072

Qian, N. (1997). Binocular disparity and the perception of depth. *Neuron*, *18*(3), 359–368. https://doi.org/10.1016/s0896-6273(00)81238-6

Rajan, G., Lafaye, J., Faini, G., Carbo-Tano, M., Duroure, K., Tanese, D., Panier, T., Candelier, R., Henninger, J., Britz, R., Judkewitz, B., Gebhardt, C., Emiliani, V., Debregeas, G., Wyart, C., & Del Bene, F. (2022). Evolutionary divergence of locomotion in two related vertebrate species. *Cell Reports*, *38*(13), 110585. https://doi.org/10.1016/j.celrep.2022.110585

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). *You Only Look Once: Unified, Real-Time Object Detection* (arXiv:1506.02640). arXiv. https://doi.org/10.48550/arXiv.1506.02640

Roberts, R. J. V., Pop, S., & Prieto-Godino, L. L. (2022). Evolution of central neural circuits: State of the art and perspectives. *Nature Reviews. Neuroscience*, *23*(12), 725–743. https://doi.org/10.1038/s41583-022-00644-y

Ronco, F., Matschiner, M., Böhne, A., Boila, A., Büscher, H. H., El Taher, A., Indermaur, A., Malinsky, M., Ricci, V., Kahmen, A., Jentoft, S., & Salzburger, W. (2021). Drivers and dynamics of a massive adaptive radiation in cichlid fishes. *Nature*, *589*(7840), Article 7840. https://doi.org/10.1038/s41586-020-2930-4

Schwalbe, M. A., Bassett, D. K., Webb, J. F. (2012). Feeding in the dark: lateral-line-mediated prey detection in the peacock cichlid *Aulonocara stuartgranti*. *Journal of Experimental Biology 215*, 2060-2071. https://doi.org/10.1242/jeb.065920

Schwalbe, M. A. B., Sevey, B. J., & Webb, J. F. (2016). Detection of artificial water flows by the lateral line system of a benthic feeding cichlid fish. *Journal of Experimental Biology*,

219(7), 1050–1059. https://doi.org/10.1242/jeb.136150

Seeholzer, L. F., Seppo, M., Stern, D. L., & Ruta, V. (2018). Evolution of a central neural circuit underlies Drosophila mate preferences. *Nature*, *559*(7715), 564–569. https://doi.org/10.1038/s41586-018-0322-9

Stephens, G. J., Johnson-Kerner, B., Bialek, W., & Ryu, W. S. (2008). Dimensionality and Dynamics in the Behavior of C. elegans. *PLOS Computational Biology*, *4*(4), e1000028. https://doi.org/10.1371/journal.pcbi.1000028

Tinbergen, N. (1951). *The study of instinct* (pp. xii, 237). Clarendon Press/Oxford University Press.

Westphal, R. E., & O'Malley, D. M. (2013). Fusion of locomotor maneuvers, and improving sensory capabilities, give rise to the flexible homing strikes of juvenile zebrafish. *Frontiers in Neural Circuits*, *7*, 108. https://doi.org/10.3389/fncir.2013.00108

Wiltschko, A. B., Johnson, M. J., Iurilli, G., Peterson, R. E., Katon, J. M., Pashkovski, S. L., Abraira, V. E., Adams, R. P., & Datta, S. R. (2015). Mapping Sub-Second Structure in Mouse Behavior. *Neuron*, *88*(6), 1121–1135. https://doi.org/10.1016/j.neuron.2015.11.031

Yoonessi, A., & Baker, C. L. (2011). Contribution of motion parallax to segmentation and depth perception. *Journal of Vision*, *11*(9), 13. https://doi.org/10.1167/11.9.13

York, R. A., Brezovec, L. E., Coughlan, J., Herbst, S., Krieger, A., Lee, S.-Y., Pratt, B., Smart, A. D., Song, E., Suvorov, A., Matute, D. R., Tuthill, J. C., & Clandinin, T. R. (2022). The evolutionary trajectory of drosophilid walking. *Current Biology*. https://doi.org/10.1016/j.cub.2022.05.039

Zhu, S. I., & Goodhill, G. J. (2023). From perception to behavior: The neural circuits underlying prey hunting in larval zebrafish. *Frontiers in Neural Circuits*, *17*, 1087993. https://doi.org/10.3389/fncir.2023.1087993

Supplementary Materials

Supplementary Figure 1: Diversity of swimming behavior in teleost larvae. Left: representative five minutes of swimming from each species. Black bar indicates a ten second window expanded on the right. Colors indicated automatically segmented and classified bouts within each swimming episode.

Supplementary Figure 2: 2D histograms of eye angles. Each point represents the angle of the left and right eyes in a given frame. Color corresponds to the probability a point belongs to a convergent state (red: converged; blue: not converged; yellow: equal likelihood).

Supplementary Figure 3: Behavioral spaces for each species. Left: t-SNE embedding and cluster identity of automatically segmented bouts. Right: example clusters from each species, showing average time series in first three PCs (top), rostrocaudal tail bending (middle), and tail kinematics (bottom) for each cluster. For cichlids, behaviors on the right are more likely to occur during prey capture and behaviors on the left are more likely to occur during spontaneous swimming.

Supplementary Video 1: Pose tracking in freely swimming *L. attenuatus*, slowed 10x. Tail points are shown in red and eye points are shown in cyan.

Supplementary Video 2: Example of *L. attenuatus* freely swimming and hunting artemia, played at half speed.

Supplementary Video 3: Example of *L. ocellatus* freely swimming and hunting artemia, played at half speed.

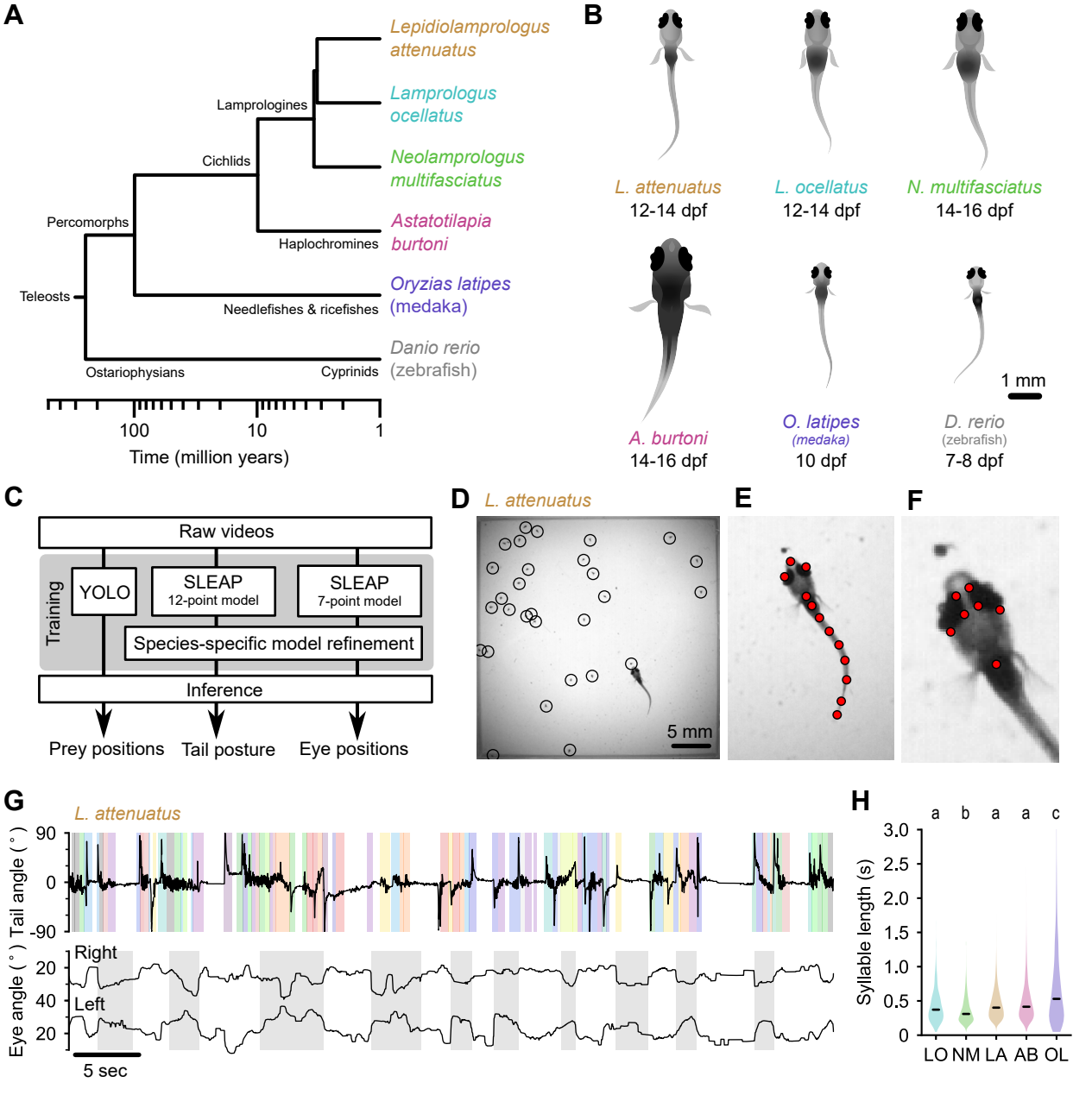
Supplementary Video 4: Example of *N. multifasciatus* freely swimming and hunting artemia, played at half speed.

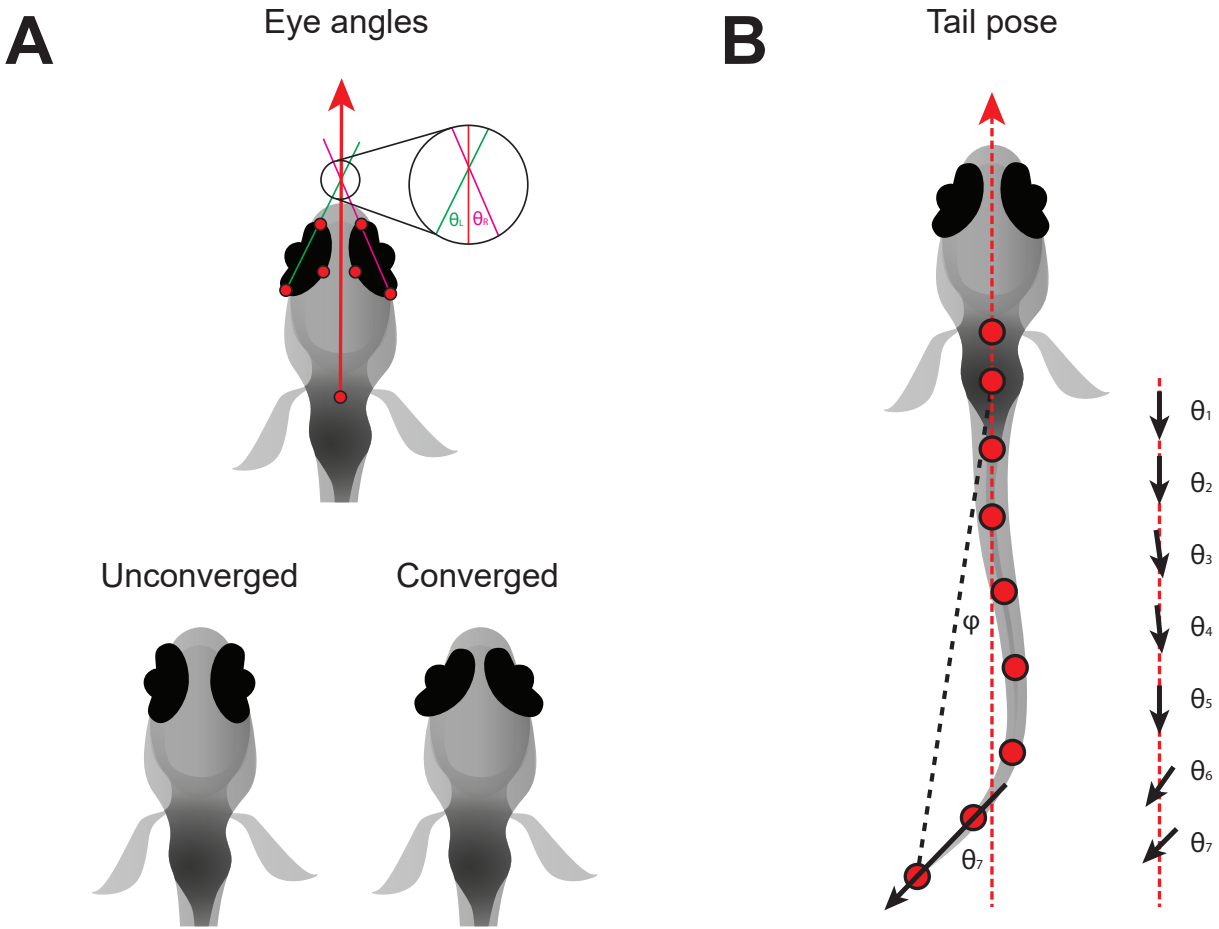
Supplementary Video 5: Example of *A. burtoni* freely swimming and hunting artemia, played at half speed.

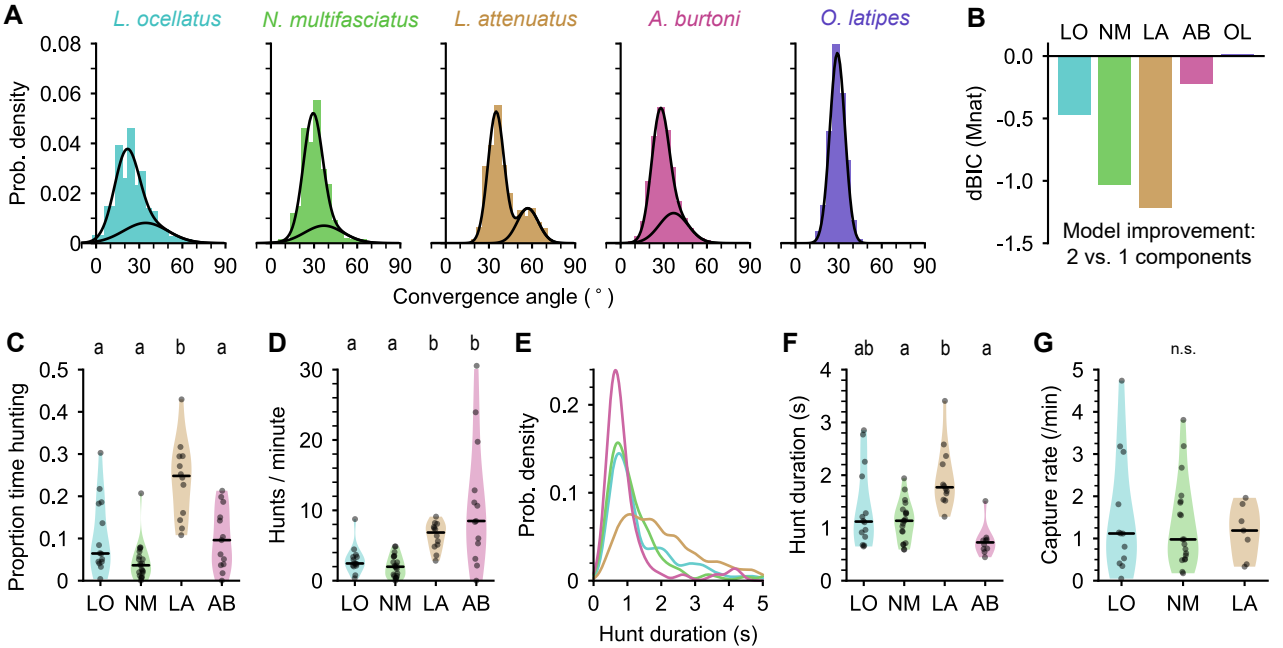
Supplementary Video 6: Example of medaka (*O. latipes*) freely swimming and hunting paramecia, played at half speed. Notice the side-swings of the head during prey capture and the lack of eye convergence.

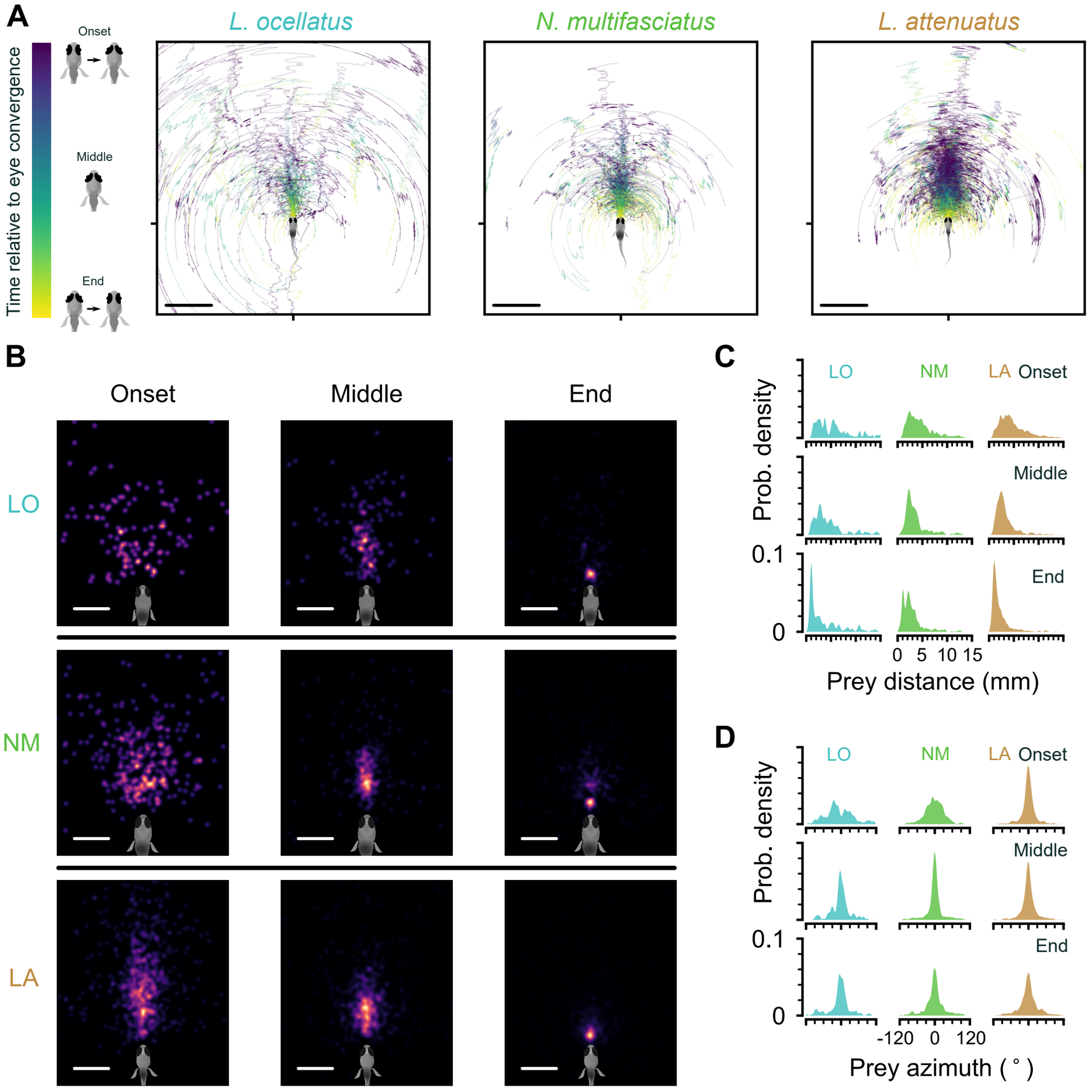
Supplementary Video 7: *L. attenuatus* spitting behavior, played at half speed.

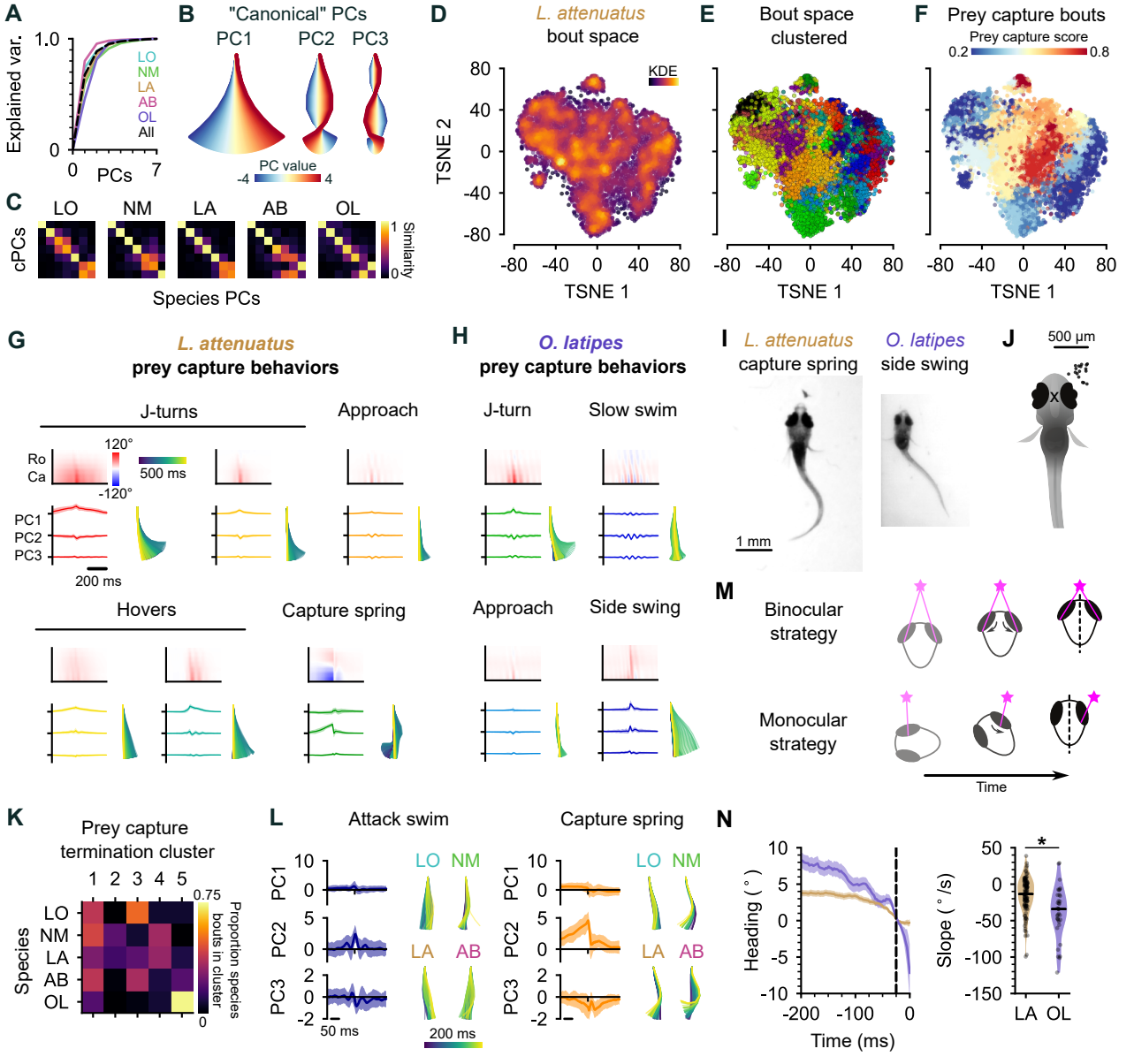
Supplementary Video 8: *A. burtoni* swimming backwards during prey capture, played at half speed.

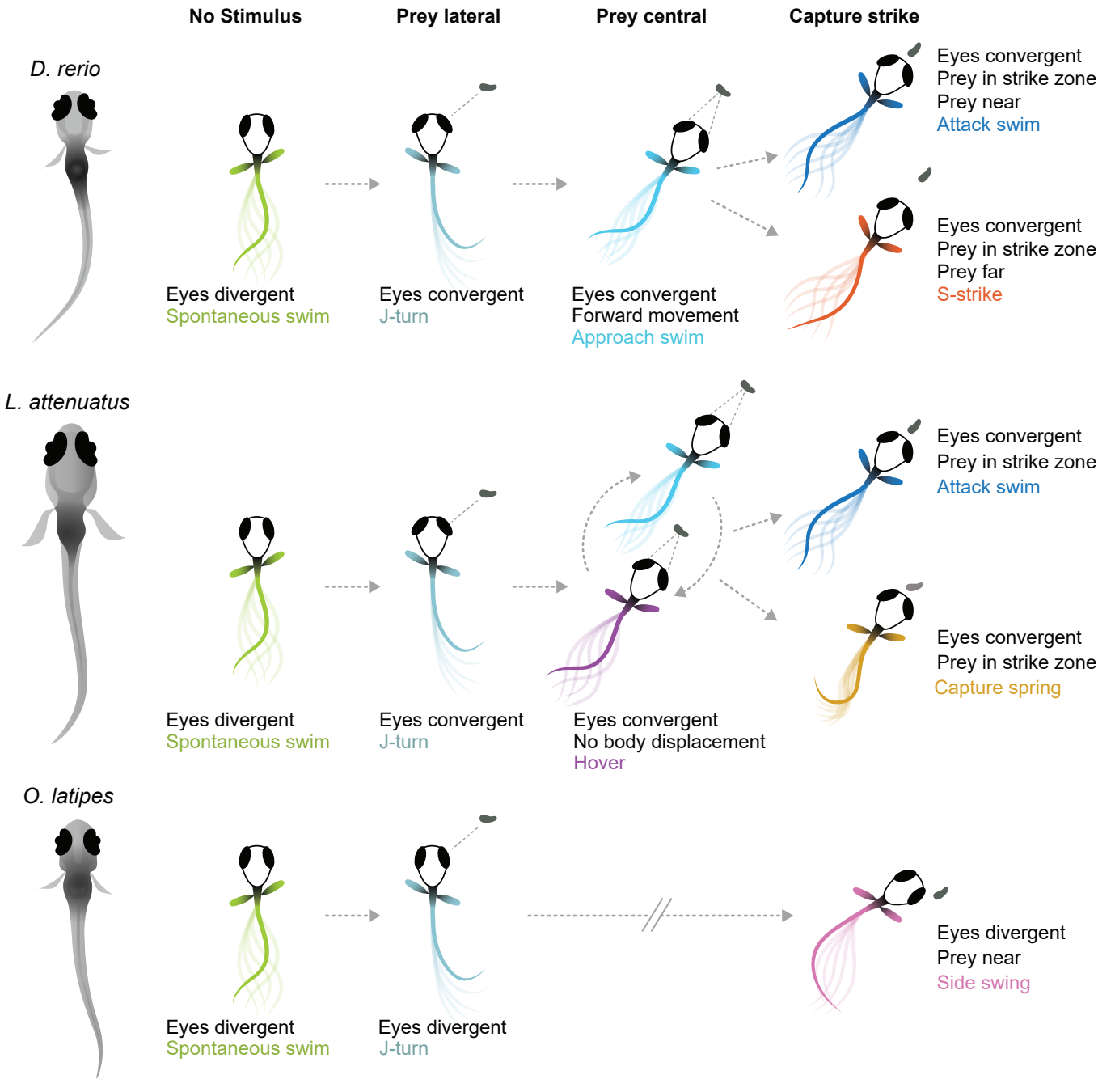

A
Lepidiolamprologus attenuatus
Lamprologus ocellatus
Neolamprologus multifasciatus
Astatotilapia burtoni
Oryzias latipes (medaka)
Danio rerio (zebrafish)
Lamprologines
Cichlids
Percomorphs
Haplochromines
Teleosts
Needlefishes & ricefishes
Ostariophysians
Cyprinids
100
10
1
Time (million years)
B
L. attenuatus
12-14 dpf
L. ocellatus
12-14 dpf
N. multifasciatus
14-16 dpf
A. burtoni
14-16 dpf
O. latipes
(medaka)
10 dpf
D. rerio
(zebrafish)
7-8 dpf
1 mm
C
Raw videos
Training
YOLO
SLEAP
12-point model
SLEAP
7-point model
Species-specific model refinement
Inference
Prey positions
Tail posture
Eye positions
D
L. attenuatus
5 mm
E
F
G
L. attenuatus
Tail angle (°)
90
0
-90
Eye angle (°)
Right
Left
20
40
20
5 sec
H
a
b
a
a
c
Syllable length (s)
3.0
2.5
2.0
1.5
1.0
0.5
0
LO
NM
LA
AB
OL


A
Eye angles
θL
θR
Unconverged
Converged
B
Tail pose
φ
θ7
θ1
θ2
θ3
θ4
θ5
θ6
θ7


A
L. ocellatus
N. multifasciatus
L. attenuatus
A. burtoni
O. latipes
Prob. density
0.08
0.06
0.04
0.02
0
0 30 60 90
Convergence angle (°)
B
LO NM LA AB OL
dBIC (Mnat)
0.0
-0.5
-1.0
-1.5
Model improvement:
2 vs. 1 components
C
a a b a
Proprtion time hunting
0.5
0.4
0.3
0.2
0.1
0
LO NM LA AB
D
a a b b
Hunts / minute
30
20
10
0
LO NM LA AB
E
Prob. density
0.2
0.1
0
0 1 2 3 4 5
Hunt duration (s)
F
ab a b a
Hunt duration (s)
4
3
2
1
0
LO NM LA AB
G
n.s.
Capture rate (/min)
5
4
3
2
1
0
LO NM LA


A
Onset
Middle
End
Time relative to eye convergence
L. ocellatus
N. multifasciatus
L. attenuatus
B
Onset
Middle
End
LO
NM
LA
C
LO
NM
LA
Onset
Middle
End
Prob. density
0.1
0
0 5 10 15
Prey distance (mm)
D
LO
NM
LA
Onset
Middle
End
Prob. density
0.1
0
-120 0 120
Prey azimuth (°)


A
Explained var.
1.0
0
0
PCs
7
LO
NM
LA
AB
OL
All
B
"Canonical" PCs
PC1
PC2
PC3
PC value
-4
4
C
LO
NM
LA
AB
OL
cPCs
Species PCs
Similarity
1
0
D
L. attenuatus
bout space
KDE
TSNE 2
TSNE 1
E
Bout space
clustered
TSNE 1
F
Prey capture bouts
Prey capture score
0.2
0.8
TSNE 1
G
L. attenuatus
prey capture behaviors
J-turns
Approach
Ro
Ca
120°
-120°
500 ms
PC1
PC2
PC3
200 ms
Hovers
Capture spring
H
O. latipes
prey capture behaviors
J-turn
Slow swim
Approach
Side swing
I
L. attenuatus
capture spring
O. latipes
side swing
1 mm
J
500 µm
M
Binocular
strategy
Monocular
strategy
Time
K
Prey capture
termination cluster
1 2 3 4 5
LO
NM
LA
AB
OL
Species
0.75
0
Proportion species
bouts in cluster
L
Attack swim
Capture spring
PC1
PC2
PC3
LO
NM
LA
AB
50 ms
200 ms
N
Heading (°)
Time (ms)
-200
-100
0
Slope (°/s)
*
LA
OL


No Stimulus
Prey lateral
Prey central
Capture strike
D. rerio
Eyes divergent
Spontaneous swim
Eyes convergent
J-turn
Eyes convergent
Forward movement
Approach swim
Eyes convergent
Prey in strike zone
Prey near
Attack swim
Eyes convergent
Prey in strike zone
Prey far
S-strike
L. attenuatus
Eyes divergent
Spontaneous swim
Eyes convergent
J-turn
Eyes convergent
No body displacement
Hover
Eyes convergent
Prey in strike zone
Attack swim
Eyes convergent
Prey in strike zone
Capture spring
O. latipes
Eyes divergent
Spontaneous swim
Eyes divergent
J-turn
Eyes divergent
Prey near
Side swing