Evolutionarily conserved transcriptional regulators control monoaminergic neuron development

Clifton Lewis[1,2,11], Matthew Goulty[1,3], Aniela Wroblewska[1], Nicola Croxall[4], David Onion[4], Sue Robinson[5], Robert P. Zinzen[6], Jordi Solana[7,8], Charlambos P. Kyriacou[1], Ezio Rosato[1], Roberto Feuda[1,9,10]

[1] Neurogenetics Group, Division of Genetics and Genome Biology, School of Biology and Biomedical Sciences, College of Life Sciences, University of Leicester, Leicester LE1 7RH, UK

[2] Centre for Chromosome Biology, Biomedical Sciences Building, University of Galway, Galway, H91 W2TY, Ireland

[3] The Francis Crick Institute, London, UK

[4] Faculty of Medicine & Health Sciences, Queen's Medical Centre, Nottingham NG7 2UH, UK

[5] Division of Molecular and Cell Biology, College of Life Sciences, University of Leicester, Leicester LE1 7RH, UK

[6] Systems Biology Imaging Technology Platform, Max Delbrück Centre for Molecular Medicine (MDC-BIMSB), Berlin, Germany

[7] Living Systems Institute, University of Exeter, Exeter, UK

[8] Department of Biosciences, University of Exeter, Exeter, UK

[9] Department of Biology, Geology and Environmental Science, University of Bologna, Bologna, Italy

[10] Correspondence: rf190@leicester.ac.uk / roberto.feuda@unibo.it (R.F.)

[11] Correspondence: clifton.lewis@leicester.ac.uk / clifton.lewis@universityofgalway.ie (C.L.)

Abstract

To what extent conserved developmental programs specify homologous cell types is a central question in biology. Here, we address this by focusing on reconstructing monoaminergic neuron development in *Drosophila melanogaster* embryo using time-resolved single-cell genomics, spatial transcript mapping with hybridisation chain reaction, and targeted metabolomics. We uncover a regulatory landscape in which specific transcription factors are activated before biosynthetic enzymes, establishing a prospective temporal architecture for monoaminergic fate specification. Comparative analyses of developmental single-cell atlases from zebrafish and sea urchin indicate that components of this machinery are conserved across ~550 million years of bilaterian evolution with orthologous transcription factors showing similar temporal dynamics. Together, these findings point to a putatively conserved regulatory core that interfaces with other context-dependent transcription factors; this interplay accommodates monoaminergic multifunction and subtype diversity across distinct neuroanatomies.

Introduction

Understanding whether conserved or divergent developmental programs underlie the specification of homologous cell types is a central goal in evolutionary and developmental biology. We address this question by focusing on monoaminergic neurons, a class of cells that produce neuromodulators such as dopamine and serotonin. Monoaminergic neurons provide a powerful system to study evolution and development for several reasons. They are molecularly defined by the expression of key biosynthetic enzymes (Libersat and Pflueger, 2004), they are widespread across bilaterians, their monoamine products can be precisely quantified, and the genes underlying their biosynthetic pathways have been shown to derive from a common evolutionary origin (Parent, 1984; Yamamoto and Vernier, 2011; Goulty *et al.*, 2023).

Although these neurons are extremely rare (<1% of neurons in fly and mammalian brains), they play fundamental roles in processes ranging from motor control and circadian rhythms to learning and mood (Libersat and Pflueger, 2004; Flames and Hobert, 2011; Monti, 2011; Waddell, 2013; Zheng *et al.*, 2025). Furthermore, their dysfunction contributes to major human neurological and psychiatric disorders, including Parkinson's disease, depression, and schizophrenia (Kurian *et al.*, 2011; Volkow *et al.*, 2011; Grace, 2016).

The developmental specification of monoaminergic neurons requires the coordinated activity of multiple transcription factors (TFs) and other regulatory genes organised into hierarchical gene regulatory networks (GRNs) (Hobert, 2008; Flames and Hobert, 2009, 2011). Studies across vertebrates and invertebrates have identified TFs that are critical for establishing monoaminergic fate (Flames and Hobert, 2011). In vertebrates, for example, LIM-homeodomain proteins Lmx1a/b, the ETS factor Fev/Pet-1, and Nkx2.2 are essential for serotonergic neuron specification (Cheng *et al.*, 2003; Ding *et al.*, 2003; Hendricks *et al.*, 2003). In *Drosophila*, the zinc-finger proteins Eagle (eg) and Huckebein (hkb) play similar roles in serotonergic differentiation (Dittrich *et al.*, 1997). In *C. elegans*, dopaminergic fate is controlled by terminal selector TFs such as the ETS transcription factor, AST-1 (Flames and Hobert, 2009). Previous work from our lab established that monoaminergic cell identity is defined by a conserved set of transcription factors shared across bilaterian species, suggesting that a common regulatory basis for these neurons originated early in animal evolution (Goulty *et al.*, 2025). Yet these candidate-based studies provide only

a partial view of the regulatory landscape underlying monoaminergic specification and their possible conservation. We therefore still lack a systems-level understanding of how monoaminergic neurons develop. Addressing these questions is essential for understanding whether homologous cell-types develop following a conserved transcriptional program.

Here to address this open question, we combine time-resolved single-cell transcriptomics, trajectory modelling, spatial transcript validation, and targeted metabolomics in embryonic neurogenesis of *Drosophila melanogaster*, and test for evolutionary conservation in zebrafish and sea urchin development. Our results reveal the developmental program of monoaminergic neurons in the *Drosophila* embryo, including key TFs—*fd59A*, *dmrt99B*, and *Lmx1a*—that pattern these lineages and differences in maturation rates between subtypes. We also identified several TF orthogroups conserved across *Drosophila*, sea urchin, and zebrafish. Comparison of monoaminergic developmental trajectory between *Drosophila* and zebrafish revealed a similar developmental cascade of TFs. Together, these findings support a putative conserved temporal logic for monoaminergic fate acquisition and begin to explain how developmental fate and function can align and/or diverge across species.

Results

A temporally resolved single-cell atlas captures neurogenesis in *Drosophila*. To understand the development of monoaminergic neurons, we performed single-cell RNA sequencing (scRNA-seq) on *Drosophila melanogaster* embryos collected at eight partially overlapping time windows spanning 0 to 22 hours after egg laying (AEL), with two biological replicates per window (**Figure 1A.**), covering the transition from neuroectoderm to fully differentiated embryonic neurons. To enrich for progenitor and early neural population, we used a SoxN-sfGFP transgenic line (VDRC #318062, Sarov *et al.*, 2016) to isolate proneural cells and neuroblasts, and anti-Elav immunostaining to enrich for differentiating neurons, followed by fluorescence-activated cell sorting (FACS) (**Figure S1.**). Single cell RNA-seq libraries were prepared on the 10X Genomics v3.1 platform and sequenced using Illumina NovaSeq X Plus.

After standard quality control, we retained 57,476 high-quality cells, resolving into transcriptionally distinct clusters. Cluster identities were assigned by mapping cluster-specific marker genes—selected as the top differentially expressed features

with adjusted p-value < 0.05—to *in situ* expression patterns from the Berkeley Drosophila Genome Project (BDGP) using an annotation pipeline adapted from Calderon *et al.* (2022). This analysis revealed a clear temporal transition: early clusters were enriched for maternal and early zygotic genes, whereas later stages featured organised, differentiated tissues; including the brain primordium, embryonic brain, and embryonic gut (**Figure 1B.** and **Table S1.**). Embedding all time points together revealed a continuous developmental trajectory (**Figure 1C.**), capturing the progressive emergence of neuronal identity.

Neurotransmitter subtype emergence reveals early co-expression and lineage complexity. Although our dataset is enriched for neural progenitors and differentiated neurons, cluster annotations revealed the presence of additional cell types (**Figure 1C.**). This likely reflects autofluorescence and/or putatively *SoxN*-driven GFP expression in the developing gut, leading to non-specific capture during FACS enrichment (**Figure S1.**).

To resolve distinct stages of neurogenesis, we used established markers - e.g. *dpn*, *wor*, *mir*, *elav*, *nSyb* - to identify neuroblasts, ganglion mother cells (GMCs), intermediate neural progenitors (INPs), and mature neurons in a population of ~40,000 cells (**Figure 2A.**, **Figure S3A.**) (Dillon *et al.*, 2022; Peng *et al.*, 2024). To capture the full neurogenic continuum, spanning from neuroblasts to mature neurons, we subsampled to ~10,000 high-quality cells, excluding disparate clusters (**Figure 2B.**, **Figure S3B.**). These cells were distributed across all the developmental time points, enabling the reconstruction of continuous differentiation trajectories (**Figure 2C.**). Among these terminal-state mature neurons, we identified seven neurotransmitter-defined populations: cholinergic (931 cells), GABAergic (816 cells), glutamatergic (173 cells), monoaminergic (77 cells), and three mixed populations co-expressing markers of GABA and glutamate (100 cells), acetylcholine and GABA (556 cells), or serotonin and GABA (388 cells), respectively (**Figure 2B.**).

Given that distinct TF combinations are known to contribute to neuronal subtype identity, we next sought to identify TFs associated with each neurotransmitter-defined population (**Figure 2E.**). Our differential TF expression analysis revealed that, with the exception of neuroblasts/GMCs, GABA/glutamatergic neurons, and monoaminergic neurons, most of these neuronal populations could not be unambiguously resolved based on TF expression alone. This outcome suggests that TFs in these other

141 lineages may lack strict subtype specificity, that neuronal identity may arise from more
142 complex regulatory interactions beyond TF expression, or that the dataset may not
143 have sufficient depth to fully resolve these distinctions. Monoaminergic neurons,
144 however, showed co-expression of multiple TFs, including *fd59A*, *dmrt99B*, *Lmx1a*,
145 and *Vsx1* and *Vsx2* (**Figure 2E.**). In contrast, the mixed serotonergic/GABAergic
146 population was marked by the co-expression of *CG8301* and *br* (**Figure 2E.**). These
147 findings define candidate regulatory genes underlying monoaminergic fate.
148
149 **RNA velocity reveals continuous differentiation trajectories.** To investigate the
150 transcriptional dynamics underlying monoaminergic neuron differentiation, we applied
151 RNA velocity analysis using scVelo (La Manno *et al.*, 2018; Bergen *et al.*, 2020) and
152 modelled lineage trajectories with CellRank (v2) (Lange *et al.*, 2022; Weiler *et al.*,
153 2024), which integrates splicing kinetics (Velocity Kernel), transcriptomic similarity
154 (Connectivity Kernel), and CytoTRACE dynamics to infer directed developmental
155 transitions (CytoTRACE Kernel). We first applied scVelo to the full dataset. Velocity
156 vector fields projected onto the UMAP embedding revealed a strong correlation
157 between transcriptional dynamics and the temporal structure of the dataset (**Figure**
158 **S4.**).

159 We then focused on the subclustered population spanning the full neurogenic
160 trajectory from neuroblasts to mature neurons (**Figure 2B–D.**). Application of scVelo
161 and CellRank to this dataset recovered clear trajectories for each subtype and a
162 smooth latent-time ordering (**Figure S5A.**). To interpret the underlying programs, we
163 identified velocity-inferred dynamical genes with the scVelo dynamical model and
164 ordered cells by latent time (**Figure S5B.**). We discretised latent time into 20 bins and,
165 for each bin, compiled the genes whose fitted expression/velocity profiles peaked
166 within that interval. Gene Ontology (GO) over-representation analyses (ORA) were
167 then performed using all expressed genes in the subset as background with
168 Benjamini–Hochberg correction. The enrichment landscape shows a coherent
169 progression: (i) early bins are dominated by cell-state remodelling and tissue-building
170 programs—RNA processing, cell cycle/proliferation, morphogenesis, and
171 organogenesis; (ii) intermediate bins transition toward neuronal differentiation, with
172 terms related to synapse formation and cell morphogenesis involved in neuron
173 differentiation; while (iii) late bins are enriched for synaptic and excitability modules,

including neuron recognition, axon guidance, and neuroactive ligand–receptor interaction, consistent with the acquisition of functional neuronal properties. Together, these GO trends align with the velocity-derived latent-time ordering and provide functional support for a continuous maturation from progenitor states to functionally specialised neurons (**Figure S5C.**).

To reconstruct the differentiation trajectory of monoaminergic neurons in greater detail, we applied CellRank to define initial and terminal states and to model transcriptional dynamics within this specific lineage (**Figure 3A.**). Genes encoding key enzymes required for monoaminergic neurotransmission—including *ple*, *Tdc1*, *Tdc2*, *Tbh*, and *Vmat*—were strongly upregulated toward the terminal state, as expected, marking the final stages of differentiation. With the exception of *Trhn* and *Ddc*, these enzymes exhibited high specificity for the monoaminergic lineage (**Figure S6.**), indicating subtype-restricted induction of the biosynthetic program. Similarly, top-ranking TFs associated with monoaminergic identity—including *fd59A*, *dmrt99B*, and *Lmx1a*—were also selectively expressed in the monoaminergic trajectory (**Figure 3B.**) and peaked at approximately the same latent time as the biosynthetic genes (**Figure 3A.**). This coordinated expression suggests a tightly coupled regulatory program in which TFs may directly activate terminal differentiation effectors. In contrast—and consistent with their role as responders to monoamine signals—genes encoding monoaminergic receptors showed broader expression across multiple neuronal lineages (**Figure S7.**). This pattern suggests that receptor deployment is decoupled from biosynthetic identity and instead reflects widespread postsynaptic responsiveness to monoamines.

Spatial validation supports coordinated transcriptional activation and resolves monoaminergic subtype identity. Although our scRNA-seq atlas resolves the major neuronal lineages and transmitter classes, monoaminergic neurons are a rare population in the *Drosophila* CNS (Busch *et al.*, 2009; Hartenstein *et al.*, 2017; Kasture *et al.*, 2018; Babski, Codianni and Bhandawat, 2024). As a result, it was difficult to clearly separate specific monoaminergic subtypes from transcriptomes alone. To both validate CellRank-inferred dynamics and add spatial specificity, we focused on two subtypes – serotonergic and dopaminergic - and performed hybridisation chain reaction (HCR) (Choi *et al.*, 2018)**.**

We assayed a panel that included most of the monoaminergic biosynthetic enzymes, their receptors, and six transcription factors (**Figure 4.**, **Figure S8-9.**). TFs were chosen based on differential expression within the monoaminergic lineage (**Figure 2C.**) and prior evidence for roles in monoaminergic fate. Spatial profiling showed that *fd59A*, *dmrt99B*, and *Lmx1a* are expressed in the CNS beginning at stages 10–12 (~8 h AEL), preceding the onset of biosynthetic gene expression. In contrast, monoaminergic enzymes were robustly expressed from stage 12 (~10 h AEL) onward, consistent with the timing of neuronal specification. Notably, this TF induction occurs slightly earlier than suggested by our latent-time analysis, which predicted a more coordinated co-activation with the enzymes. Receptors generally followed a similar onset, with most appearing from stage 12. A notable exception was *5-HT2A*, which was detected as early as stage 1 and in non-neural tissues, consistent with a recent study implicating it in early morphogenetic roles beyond neurotransmission (Karki *et al.*, 2023)**.**

Together, these data mostly corroborate the temporal ordering inferred by CellRank and, crucially, resolve subtype identity by localising expression to serotonergic or dopaminergic domains. They further suggest that these TFs, *fd59A*, *dmrt99B*, and *Lmx1a*, may act upstream to specifically initiate part of the monoaminergic program, with effector enzymes and receptors switching on subsequently.

Mass spectrometry reveals temporally distinct onset of monoamine production. Our transcriptional and spatial analyses suggests that by stage 12 (~9–10 h AEL), the *Drosophila* embryo expresses the biosynthetic enzymes and receptors required to produce and respond to monoamines. However, whether these neurons are functionally active at this stage remains unresolved. Previous studies have detected serotonin and dopamine via immunohistochemistry only from stage 17 (~15 h AEL) onwards (Lundell and Hirsh, 1994), raising questions about the relationship between transcriptional commitment and neurotransmitter availability. To resolve this dichotomy, we conducted targeted mass spectrometry across seven developmental windows spanning 0–22 h AEL. At each time point, we analysed ≥6 biological replicates (54 samples total) and quantified 19 analytes, including all canonical monoamines (**Table S2.**) GABA, glutamate, and acetylcholine **(Figure S11.).**

240 Our results revealed distinct temporal profiles across monoaminergic subtypes
241 (**Figure 5.; Figure S10.**). Histamine and tyramine were constitutively present
242 throughout development, suggesting their production may occur independently of
243 canonical neurogenesis. In contrast, both dopamine and serotonin exhibited
244 temporally restricted accumulation, with sharp increases after stage 12—closely
245 aligning with the transcriptional induction of their biosynthetic enzymes. Although
246 dopamine and serotonin followed similar temporal trajectories, their relative levels
247 differed, likely reflecting variations in biosynthetic efficiency or substrate availability. It
248 is important to note that this effect may also reflect differences in the detection
249 capabilities for these molecules.
250 Together, these findings indicate that transcriptional programs broadly predict
251 the functional potential of monoaminergic neurons, but the timing and extent of
252 neurotransmitter production are modulated by subtype-specific biochemical properties
253 and/or regulatory constraints.
254
255 **Conserved transcriptional dynamics underpin monoaminergic neuron**
256 **development in *Drosophila* and zebrafish.** Monoaminergic neurons are widely
257 distributed across bilaterian animals, yet whether they arise from a conserved
258 developmental program remains unclear. Our findings in *Drosophila* suggest that a
259 specific set of TFs are activated during monoaminergic neuron differentiation,
260 providing an opportunity to test the evolutionary conservation of this regulatory
261 program. To address this question, we reanalysed published developmental single-
262 cell RNA-seq data from *Danio rerio* and sea urchin, *Strongylocentrotus purpuratus*.
263 For *Danio rerio*, we reprocessed a developmental time series spanning embryonic to
264 larval stages (10 hours post-fertilisation [hpf] 232 to 10 days post-fertilisation [dpf];
265 Lange *et al.*, 2024). For *S. purpuratus*, we merged two complementary datasets,
266 covering the 8-cell embryo through the 72 hpf larval stages (Foster, Oulhen and
267 Wessel, 2020; Perillo *et al.*, 2020). Central nervous system cell types were annotated
268 using orthologous marker genes applied previously in *Drosophila* (**Figure 2D.**) and
269 identified monoaminergic neuron populations based on expression of canonical
270 biosynthetic genes (**Figure S12A-B.**).
271 Applying the same differential-expression (DE) criteria used in fly (adjusted $p <$
272 0.05; $|\log_2 FC| \geq 0.25$), we identified transcription factors (TFs) enriched in zebrafish
273 and sea urchin monoaminergic clusters analysed independently (**Figure 6A.**; **Table**

274 **S3.**). We analysed clusters separately, rather than collapsing them into a single
275 monoaminergic group, to gather an inclusive list of TFs that define monoaminergic
276 identity across subtypes. To enable cross-species comparisons despite lineage-
277 specific duplications, genes were mapped to bilaterian-level eggNOG orthogroups.
278 Orthologs of key fly regulators were recovered in vertebrates, including *lmx1a* paralogs
279 (*lmx1a*, *lmx1al*, *lmx1ba*). In total, 15 TF orthogroups were differentially expressed in
280 both fly and zebrafish monoaminergic neurons, and 16 TF orthogroups were shared
281 between fly and sea urchin (**Figure 6C.**; **Table S4.**), defining a set of putative
282 conserved regulators. Notably, six TF orthogroups—*CG8301*, *Lmx1a*/*CG4328*, *Hr3*,
283 *Eip78C*, *net*, and *Pdp1*—were shared across all three species comparisons,
284 underscoring a potentially conserved core program for monoaminergic identity. The
285 partial overlap of the remaining TFs suggests that, while a conserved regulatory core
286 might exist, additional factors have diverged or been co-opted in lineage-specific
287 contexts

288 To test whether this conservation was unique to monoaminergic neurons, we
289 repeated the orthogroup-level analysis between *Drosophila* and zebrafish for other
290 neurotransmitter classes (Acetylcholine, GABA, Glutamate; **Figure S13.**). The number
291 of shared DE TF orthogroups was of similar magnitude across classes, despite vastly
292 different neuronal population sizes, supporting a model in which each transmitter
293 lineage may share a small, conserved set of regulators that is then diversified by
294 lineage- and species-specific additions.

295 We next asked whether not only TF identity but also their temporal activation is
296 conserved. Using the comparable time-resolved zebrafish dataset, we applied RNA
297 velocity and CellRank to reconstruct differentiation dynamics. Several orthologous
298 TFs, including *lmx1* paralogs, showed late upregulation along the zebrafish
299 monoaminergic trajectory, closely mirroring the dynamics observed in *Drosophila*
300 (**Figure 6B.**). Thus, some of the regulatory content and timing of activation are aligned
301 between insects and vertebrates, at least in these representative species.

302 Together, these findings support a model in which monoaminergic neurons
303 emerge from a potential conserved transcriptional program, with shared regulators,
304 such as *Lmx1a*, but critically has undergone significant species-specific elaborations.

305 ## Discussion

306 The extent to which homologous cell types are specified by conserved developmental
307 and transcriptional programs remains a central question in evolutionary developmental
308 biology. Monoaminergic neurons, which play fundamental roles in neurotransmission
309 across bilaterians, offer a powerful model for addressing this issue. Previous
310 comparative transcriptomic analyses have suggested that monoaminergic cell identity
311 is defined by a shared set of transcription factors, indicating an origin in the bilaterian
312 stem lineage (Goulty *et al.*, 2025). Yet, whether the developmental processes that give
313 rise to these neurons are likewise conserved—or instead represent cases of
314 convergent evolution—has remained unresolved. Additionally, by determining whether
315 the developmental program underlying these neurons is shared across species, we
316 further clarify the hypothesis of a conserved evolutionary origin. In this study, we
317 characterise the transcriptional program governing monoaminergic neuron
318 development in *Drosophila* embryos and reveal that key regulatory components are
319 shared with vertebrate and deuterostome systems. Our findings suggest that
320 monoaminergic neurons are specified by an evolutionarily conserved developmental
321 framework that emerged early in bilaterian evolution.

322 Within *Drosophila*, we identify a regulatory program involving *fd59A*, *dmrt99B*,
323 *Vsx1*, *Vsx2*, and *Lmx1a*. These TFs are selectively expressed in the monoaminergic
324 lineage and become active around stage 12, coincident with the onset of biosynthetic
325 enzyme expression (*ple*, *Tdc2*, *Tbh*, *Vmat*). Spatial validation by HCR confirmed the
326 sequential activation of these genes and showed that TF induction precedes
327 biosynthetic gene expression, supporting their role as primary regulators of
328 monoaminergic fate. This organisation mirrors vertebrate serotonergic development,
329 where Lmx1a/b and Pet-1/Fev initiate comparable programs (Ding *et al.*, 2003;
330 Hendricks *et al.*, 2003; Flames and Hobert, 2011), pointing to deep evolutionary
331 parallels.

332 Although biosynthetic gene expression initiates around stage 12, metabolomic
333 profiling reveals pronounced differences in monoamine accumulation. Dopamine and
334 serotonin both increase after stage 12, yet dopamine accumulates more rapidly—likely
335 reflecting the higher catalytic efficiency of Dopa decarboxylase (Ddc) for L-3,4-
336 dihydroxyphenylalanine (L-DOPA) than for 5-hydroxytryptophan (5-HTP), the
337 respective intermediates in the dopamine and serotonin biosynthetic pathways

338 (Montioli and Borri Voltattorni, 2021). In contrast, octopamine remains undetectable
339 despite enzyme expression, suggesting either technical limitations or additional
340 biological constraints such as substrate availability, enzymatic kinetics, or post-
341 transcriptional control. These findings indicate that transcriptional activation alone is
342 insufficient for transmitter synthesis, pointing to additional regulatory layers that
343 modulate monoamine output. However, these interpretations remain constrained by
344 the detection sensitivity and chemical properties of the specific metabolites analysed.

345 Our comparative analysis of monoaminergic neuron development in
346 *Drosophila*, zebrafish, and sea urchins reveals a deeply conserved transcriptional
347 program, underscoring a shared regulatory logic that likely emerged early in bilaterian
348 evolution. The identification of six conserved TF orthogroups—*CG8301*,
349 *Lmx1a*/*CG4328*, *Hr3*, *Eip78C*, *net*, and *Pdp1*—establishes a core genetic program for
350 specifying monoaminergic identity. This finding supports the hypothesis that these
351 neurons represent potentially homologous cell types that originated early in bilaterian
352 evolution. Core regulatory modules specifying neuronal identity appear to have been
353 established in the bilaterian ancestor and maintained across deep evolutionary time
354 (Arendt, 2005; Arendt *et al.*, 2019; Rapti, 2022; Sachkova, Modepalli and Kittelmann,
355 2025). For instance, Lmx1a has a well-documented and critical role in the
356 development of dopaminergic neurons in vertebrates, where it regulates key effector
357 genes such as tyrosine hydroxylase (*Th*) and the dopamine transporter (*DAT*) to
358 ensure proper differentiation and maturation (Cai *et al.*, 2009; Doucet-Beaupré, Ang
359 and Lévesque, 2015). This role has not previously been demonstrated in non-
360 vertebrates, suggesting that *Lmx1a* may represent an ancient regulator of
361 monoaminergic fate specification—although convergent evolution of similar
362 transcriptional strategies cannot be excluded.

363 Importantly, this conservation is not absolute. We also observed substantial
364 regulatory divergence, with several TFs associated with monoaminergic differentiation
365 being unique to specific lineages. This pattern may reflect a dual-layered evolutionary
366 process—an ancient, conserved core program governing fate commitment,
367 complemented by lineage-specific regulators that refine neuronal subtype identity
368 (Hobert, 2008; Arendt *et al.*, 2019). Alternatively, the conserved TF set itself may
369 specify distinct monoaminergic identities across species, while divergent TFs act as
370 complementary components that integrate these core programs into broader gene-
371 regulatory networks, coordinating cross-talk and aligning specification with species-

specific neural architectures (Flames and Hobert, 2011; Rapti, 2022). While some apparent differences may arise from technical factors (e.g., sequencing depth, detection sensitivity, or annotation), others likely represent genuine regulatory divergence. Given the substantial variation in neuroanatomical organisation among insects, vertebrates, and echinoderms, such divergence may reflect adaptive remodelling of the regulatory landscape to distinct neural architectures.

The temporal dynamics of conserved TFs suggest that these genes govern monoaminergic fate commitment, whereas lineage-specific TFs modulate neuronal subtype identity in a species-dependent manner. The hierarchical relationships between conserved and lineage-specific TFs remain to be elucidated. Determining whether these factors operate through feedforward or feedback regulatory loops—a recurrent feature of GRN architecture that stabilises neuronal identity—will be crucial (Flames and Hobert, 2011; Peter and Davidson, 2011). Understanding how conserved GRNs are modified to generate lineage-specific phenotypes will be essential for uncovering the general principles that govern cell-type diversification across evolution.

Our mass-spectrometry data indicate a partial decoupling between monoamine abundance and biosynthetic enzyme expression. While enzyme kinetics may account for some of the observed differences in transmitter accumulation, it remains unclear whether such decoupling is a general feature of monoaminergic neurons across taxa. Moreover, the functions of early monoamine signalling—potentially enabled by precocious receptor expression (e.g., 5-HT2A)—during embryogenesis remain unresolved.

Integrating transcriptomic, metabolomic, and spatial data across additional bilaterian taxa will be important for testing the universality of these patterns. Comparative perturbation experiments—targeting conserved TFs such as Lmx1a or Hr3 in non-model systems—could determine whether shared transcriptional logic indeed translates into conserved functional outcomes.

In conclusion, our study uncovers a conserved temporal program for monoaminergic neuron specification, in which early transcription factor activation orchestrates biosynthetic gene expression to align fate determination with functional maturation. By linking transcriptional programs, transmitter synthesis, and receptor deployment across species, our work provides a framework for understanding how conserved and divergent GRNs generate neuromodulatory diversity and how neuronal fate and function are developmentally coupled.

406 ## Methods

407 ### Fly husbandry

408 All *Drosophila melanogaster* stocks were maintained on standard cornmeal–yeast–
409 agar medium at 18°C or 25°C, with relative humidity held at approximately 40%. We
410 employed a transgenic SoxN–sfGFP line (VDRC #318062), which was reared
411 continuously at 25°C (Sarov *et al.*, 2016).

412

413 ### Embryo collection

414 Adult flies were transferred into population cages fitted with agar plates prepared from
415 mango juices, each supplemented with fresh yeast paste to boost fecundity. To
416 eliminate circadian-driven fluctuations in gene expression, cages were kept in constant
417 darkness (DD) beginning immediately upon transfer (Wijnen *et al.*, 2006). Following a
418 48-hour acclimation under DD, embryo collections was commenced, ensuring that all
419 developmental stages progressed without light-induced transcriptional variability.

420 Embryos were collected over eight successive, overlapping time windows
421 covering the full span of embryonic development; between 0-22 hours after egg laying
422 (AEL). Collections were performed in biological duplicate to increase data reliability
423 and resolution of transitional stages. Agar plates were replaced regularly to maintain
424 optimal oviposition conditions, and collected embryos remained in the same DD
425 incubator for precise aging.

426 Given the necessity of preserving RNA integrity for 10X Genomics single-cell
427 RNA-sequencing, we avoided formaldehyde-based fixatives. Instead, we applied
428 dithio-bis(succinimidyl propionate) (DSP; Sigma-Aldrich), a reversible cross-linker
429 compatible with downstream library preparation. A 50 mg/mL DSP stock was made in
430 DMSO, then diluted by adding 5 mL Dulbecco's PBS (DPBS; Gibco) to the 100 µL
431 stock, followed by sterilisation through a 0.2 µm syringe filter. Collected embryos were
432 fixed in this working solution, preserving stage-specific transcriptomes and enabling
433 reliable reversal of cross-links prior to sequencing.

434

435 ### Embryo Harvesting, Dechorionation, and Fixation

436 Collected embryos were gently dislodged from agar plates by rinsing with distilled
437 water and transferred into a 30 µm nylon mesh sieve using a fine paintbrush. Any
438 residual yeast paste was washed away under a gentle stream of distilled water to

439 minimise contamination. Dechorionation was performed by submerging the embryos
440 in 50% bleach (Arco Essentials) for 3 minutes, then immediately rinsing them
441 thoroughly with distilled water until all bleach was removed.
442 For fixation, equal volumes of the DSP working solution and heptane (Fisher
443 Scientific) were combined in a glass scintillation vial. Dechorionated embryos were
444 added to this biphasic mixture and incubated on a gyroscopic rocker at room
445 temperature for 1 hour, ensuring gentle but continuous agitation. After fixation, the
446 lower aqueous phase was discarded, and cold methanol (VWR) was added to the
447 heptane layer to initiate devitellinisation. Embryos were vortexed for 2 minutes to
448 facilitate complete removal of the vitelline membrane, then allowed to settle to the
449 bottom of the vial. Finally, embryos were transferred to 1.5 mL Eppendorf tubes and
450 washed three times in fresh methanol. Fixed embryos were stored at –20°C until
451 further processing.
452
453 **Immunofluorescence**
454 All immunostaining was performed at 4°C under stringent RNase-free conditions:
455 Protector RNase Inhibitor (Roche, RNAINH-RO) was added to all buffers at 5 U/mL
456 during brief washes and 20 U/mL for all longer incubations, and both bovine serum
457 albumin (BSA; Gibco) and normal goat serum (NGS; Vector Laboratories) were heat-
458 inactivated at 60°C for 30 minutes to eliminate residual RNases. Prior to antibody
459 staining, DSP–methanol–fixed embryos were gradually rehydrated through an ethanol
460 series into PBT (PBS + 0.1% Triton X-100; Sigma-Aldrich T8787) and then washed
461 three times for 10 minutes each in PBT containing RNase inhibitor, with gentle rocking
462 to ensure complete removal of the methanol fixative. To block non-specific binding,
463 embryos were incubated overnight at 4°C in PBT supplemented with 1% heat-
464 inactivated BSA and 2% heat-inactivated NGS. The embryos were then transferred
465 into fresh blocking buffer (PBT + inhibitor, 1% BSA, 2% NGS) containing primary
466 antibodies: rabbit α-GFP (1:100; Invitrogen) to label SoxN–sfGFP^{+} neurogenic cells,
467 and rat α-Elav (1:100; DSHB) to detect post-mitotic neurons. After overnight incubation
468 at 4°C with gentle agitation, excess primary antibody was removed by three 10-minute
469 washes in PBT. Secondary detection was performed by incubating embryos overnight
470 in the same blocking buffer supplemented with goat α-rabbit Alexa Fluor 488 (1:500;
471 Invitrogen), mouse α-rat PE (1:500; Invitrogen), and the nuclear counterstain DRAQ5

472 (4 µL/mL; Thermo Scientific). Finally, stained embryos were washed three times in
473 PBT (10 minutes each) and stored at 4 °C in PBT until mounting and imaging.
474
475 **Microscopy validation and processing**
476 Images were acquired using a Zeiss LSM 980 Airyscan 2 microscope in Airyscan
477 super-resolution mode, with the default settings applied for Airyscan image
478 processing. A Plan-Apochromat 10x/0.45 or Plan-Apochromat 20x/0.8 objective was
479 employed for image acquisition (University of Leicester Advanced Imaging Facility).
480 Post-processing was performed using a custom FIJI (Schindelin *et al.*, 2012) macro
481 for orienting the embryos along anatomical axes (Embryo Processing IF Script).
482
483 **Single-cell dissociation**
484 Embryos were transferred into an ice-cold 15 mL Dounce homogeniser in PBS
485 containing Protector RNase Inhibitor, and all subsequent steps were performed at 4°C
486 to preserve RNA integrity. After three gentle washes with cold Dulbecco's PBS
487 (DPBS), the supernatant was removed and replaced with PBS supplemented with
488 0.04% BSA (PBS/BSA). Embryos were then mechanically disrupted by 10–15 strokes
489 with the tight-fitting pestle, and the homogenate was collected into a fresh tube. To
490 maximise cell recovery, the homogeniser was rinsed with additional PBS/BSA, and
491 the combined suspension was centrifuged at 3,000 × g for 3 minutes. The supernatant
492 was discarded, and the pellet was resuspended in PBS/BSA.
493 To remove large debris, the resuspended cells were centrifuged at 40 × g for 3
494 minutes, after which the supernatant—containing single cells—was transferred to a
495 new tube and spun at 1,000 × g for 3 minutes. The resulting pellet was gently triturated
496 in PBS/BSA by passing the suspension through a 22-gauge needle with 20 strokes to
497 dissociate any remaining aggregates, then filtered through a 100 µm nylon mesh. A
498 final spin at 1,000 × g for 3 minutes pelleted the cleaned single-cell suspension, which
499 was resuspended in PBS/BSA and kept on ice prior to fluorescence-activated cell
500 sorting (FACS). Parallel control samples—unstained; stained with only primary or only
501 secondary antibodies; and combinations thereof with DRAQ5 nuclear stain—were
502 prepared and processed identically for gating and compensation during FACS
503 analysis.
504

505 **FACS**

506 Cells were sorted on a Beckman Coulter MoFlo Astrios cell sorter (Nottingham)
507 equipped with a 100 µm nozzle, operating in “Purify” precision mode with a 1–2 drop
508 envelope to maximise purity of collected fractions. Single-cell suspensions were
509 maintained at 4°C throughout sorting by housing the entire flow chamber in a water-
510 cooled environment, thereby minimising heat-induced stress and preserving cell
511 integrity. Samples were deposited into a two-way tube holder pre-chilled on ice.

512 Gating strategies were defined using unstained and single-stained controls to
513 eliminate autofluorescent events. We applied broad gates to capture both strongly and
514 weakly fluorescent populations, acknowledging the continuum of marker expression.
515 $SoxN^+$ and $Elav^+$ cells—representing early neurogenic and post-mitotic neuronal
516 cohorts, respectively—were collected together. This inclusive approach ensured
517 recovery of the full spectrum of developing neural cells.

518

519 **Reverse cross-linking**

520 After FACS, the cells were subjected to a reverse cross-linking step to remove the
521 DSP fixative and prepare them for subsequent library preparation. The sorted cells
522 were first centrifuged at 4°C at 1,000 x g for 3 minutes, and the supernatant was
523 discarded. The cell pellet was then resuspended in PBS/BSA solution and incubated
524 with 50 mM dithiothreitol (DTT, Sigma-Aldrich) at 37°C for 30 minutes to break the
525 cross-links formed by DSP. Following incubation, the cells were centrifuged twice at
526 4°C at 1,000 x g for 3 minutes to thoroughly remove any residual DTT.

527

528 **10X Genomics Library Preparation**

529 The library preparation was conducted using the 10X Genomics Chromium Next GEM
530 Single Cell 3' Reagent Kit v3.1 (10X Genomics). Prior to the preparation, cell
531 concentration was assessed with a Luna FX7 Cell Counter, and cells were adjusted
532 to a concentration of approximately 500-1500 cells/µl. Approximately 8000 cells per
533 sample were used to generate single-cell gel beads in emulsion (GEMs), targeting
534 approximately 5000 single-cell GEMs per sample. GEM emulsions were created with
535 the 10X Chromium Controller, followed by complementary DNA (cDNA) synthesis and
536 library preparation according to the manufacturer's protocol. cDNA amplification was
537 performed for 12 cycles, and library amplification during the PCR indexing stage was
538 similarly carried out for 12 cycles using the Dual Index Kit TT Set A (10X Genomics).

539 Quantification of cDNA was performed using Qubit dsDNA HS assays (Invitrogen) and
540 D5000 HS ScreenTape (Agilent) on an Agilent Technologies TapeStation 4200, while
541 library quantification was conducted with Qubit dsDNA HS assays and D1000
542 ScreenTape on the same platform. Libraries were prepared following the
543 manufacturer's guidelines and were subsequently sequenced externally by Novogene
544 on the Illumina NovaSeq X Plus platform, utilising 150 bp paired-end reads.
545
546 **HCR Validation**
547 Custom probe sets targeting mRNAs of key monoaminergic pathway enzymes,
548 receptors, and TFs were designed from the longest isoform FlyBase annotations and
549 synthesised by Molecular Instruments. Briefly, embryos were first equilibrated in
550 hybridisation buffer (v3.1, Molecular Instruments) at 37°C for 30 min to block non-
551 specific sites, then incubated overnight (12–16 h) at 37°C with 4 pmol of each probe
552 set in fresh hybridisation buffer, following the manufacturer guidelines. Following this,
553 unbound probes washed, and fluorophore-conjugated HCR hairpins (Molecular
554 Instruments) were snap-cooled and applied to embryos in amplification buffer, then
555 incubated for 12–16 h at room temperature in the dark. Post-amplification, embryos
556 were washed to remove excess hairpins.
557 Finally, embryos were counterstained with DAPI (1:1000, Sigma-Aldrich) during the
558 last wash, mounted in antifade medium (Vectashield; Vector Laboratories), and
559 imaged by confocal microscopy and processed using a custom macro script in FIJI
560 (Embryo Processing HCR Script).
561
562 **Targeted Metabolomics**
563 Embryos from seven consecutive developmental windows were rinsed thoroughly in
564 distilled water to remove yeast, excess liquid discarded, weighed, and flash-frozen in
565 liquid nitrogen. Samples were shipped on dry ice to the EMBL Metabolomics Core
566 Facility (Heidelberg, Germany), where Bernhard Drotleff performed all downstream
567 processing.
568 Each frozen sample (75 mg/mL) was extracted on dry ice by adding 80%
569 acetonitrile (with internal standards and 1% formic acid), then homogenised using 1
570 mm zirconia beads in a bead-beater (FastPrep-24) at 6.0 m/s (5 × 30 s, 5 min rest).
571 After incubation at –20 °C for 20 min, lysates were vortexed and centrifuged (15 000

572 × g, 10 min, 4 °C). Supernatants were transferred to glass vials and analysed within 2
573 h.
574 LC–MS/MS was carried out on an Agilent Infinity II BioLC coupled to a Sciex
575 QTRAP 6500+ (positive ESI). Chromatography used an Atlantis Premier BEH Z-HILIC
576 column (2.1 × 100 mm, 1.7 µm) at 40 °C, with mobile phases A (90:10 H_2O:ACN) and
577 B (98:2 ACN:H_2O), both buffered with 25 mM ammonium formate (pH 3.0). The 13
578 min gradient (0–1 min 100% B; 1–8 min to 60% B; 8–10 min 60% B; 10–10.5 min to
579 100% B; 10.5–13 min 100% B) ran at 0.3 mL/min, with 5 µL injections at 4 °C. Ion
580 source parameters were: curtain gas 40 psi, nebulizer gas 40 psi, heater gas 40 psi,
581 ion spray voltage 5 kV, source temperature 400 °C, medium collision gas. Data were
582 acquired in scheduled MRM mode (90 s windows, 600 ms cycle). Samples were
583 analysed in randomised order, with pooled QCs for system equilibration and every fifth
584 sample to monitor performance; a processed blank defined background. Raw peak
585 areas were extracted and processed in SciexOS. Data analysis and visualisation was
586 done in R using custom scripts (Metabolomics R Script).
587
588 **scRNA-seq Analysis: Data alignment**
589 Raw sequencing data were processed using Cell Ranger count (v9.0.1, 10X
590 Genomics). Reads were aligned to a custom *Drosophila melanogaster* reference
591 (Ensembl release BDGP6.46.110) into which the GFP transgene sequence had been
592 incorporated. Cell Ranger generated UMI-collapsed gene–cell count matrices for each
593 library.
594
595 **scRNA-seq Analysis: Preprocessing and Quality Control**
596 Downstream quality control and preprocessing were carried out in Seurat (v5.2.0; full
597 pipeline available at GitHub page). Briefly, cells with fewer than 800 or more than
598 100,000 UMIs, fewer than 200 or more than 8,000 detected genes, or mitochondrial
599 transcript content exceeding 8–10% were excluded. After log-normalisation and
600 identification of highly variable features, JackStraw analysis was used to select
601 significant principal components for dimensionality reduction. Uniform Manifold
602 Approximation and Projection (UMAP) was then performed on these PCs. To remove
603 artifactual multiplets, DoubletFinder (v2.0.6, McGinnis, Murrow and Gartner, 2019)
604 was applied with an expected doublet rate of 3.9%, consistent with 10X Genomics
605 loading recommendations.

606 **scRNA-seq Analysis: Integration and Cluster Annotation**

607 Biological replicates from each developmental time point were integrated using
608 Harmony (v1.2.3, Korsunsky *et al.*, 2019) to correct batch effects while preserving
609 biological variation. Clusters were initially defined in the integrated UMAP space and
610 then annotated following a modified version of the Calderon *et al.* (2022) pipeline:
611 marker genes for each cluster, selected as the top differentially expressed features
612 with adjusted $P < 0.05$, were compared against curated spatial expression terms from
613 the Berkeley Drosophila Genome Project (BDGP) using Fisher's exact test to identify
614 over-represented *in situ* annotations (Tomancak *et al.*, 2002, 2007; Hammonds *et al.*,
615 2013; Calderon *et al.*, 2022). After Bonferroni correction, significant enrichment terms
616 were manually reviewed and consolidated to assign each cluster to a defined
617 neurogenic or neuronal identity. This combined data-driven and expert-curated
618 approach ensured robust annotation of cell states across *Drosophila* embryonic
619 neurogenesis.

620 To preserve the continuum of developmental variation, cells from all time points
621 were merged without additional batch-correction; standard integration algorithms
622 tended to over-correct and attenuate genuine developmental signals. The merged
623 dataset was then reclustered in Seurat, and each resulting cluster was assigned an
624 identity based on the modal annotation of its constituent cells from the individual time-
625 point analyses. This majority-vote approach ensured that clusters reflected the
626 predominant cell state across development.

627

628 **scRNA-seq Analysis: Neuronal-Specific Annotation**

629 We next focused on delineating specific neuronal lineages. Clusters corresponding to
630 the progression from neuroblasts to mature neurons were reannotated using
631 expression of subtype markers defined in Dillon *et al.* (2022) Peng *et al.* (2024). This
632 approach yielded fine-grained, cell-type annotations that aligned with, and extended,
633 the broader developmental ontology derived from BDGP enrichment annotation
634 pipeline.

635

636 **scRNA-seq Analysis: Trajectory Inference via RNA Velocity**

637 To infer cellular dynamics and developmental trajectories, Cell Ranger BAM files were
638 reprocessed with Velocyto (v0.17.17, La Manno *et al.*, 2018) to quantify spliced and
639 unspliced transcripts, producing Loom matrices. Next, the Seurat-annotated object

640 was converted to an AnnData structure using sccustomize (v3.0.1, Samuel Marsh *et*
641 *al.*, 2024), and merged with the velocity matrices preserving metadata, embeddings,
642 and clustering labels. Downstream modeling was carried out in scVelo (v0.3.3, La
643 Manno *et al.*, 2018; Bergen *et al.*, 2020) using the dynamical model: data were first
644 filtered and normalised, moments were computed on the neighbourhood graph, and
645 velocity vectors were estimated for each cell. Finally, velocity-stream embeddings
646 were projected onto the UMAP embedding to visualise inferred developmental
647 progressions.
648
649 **scRNA-seq Analysis: Subclustering and Lineage Dynamic Analysis**
650 To resolve finer neuronal subtypes, we isolated cells from the maturing neuronal
651 lineage and performed a second round of clustering in Seurat. Clusters were
652 reannotated based on the expression of established subtype marker genes, and dot
653 plots were generated to verify marker specificity. We then recalculated RNA-velocity
654 on this subset in scVelo (v0.3.3, La Manno *et al.*, 2018; Bergen *et al.*, 2020), re-
655 deriving the first- and second-order moments, velocity vectors, velocity-stream
656 embedding, and latent-time projections. These velocity maps were overlaid on the
657 UMAP to visualise intra-lineage transitions. Finally, CellRank was applied to the
658 subclustered dataset—integrating connectivity, velocity, and CytoTRACE kernels—to
659 infer both initial (progenitor) and terminal (mature neuron) fates and compute fate
660 probabilities within the neuronal lineage (Reuter *et al.*, 2018; Lange *et al.*, 2022;
661 Reuter, Klein and Lange, 2022; Weiler *et al.*, 2024).
662
663 **scRNA-seq Analysis: Regulatory dynamics**
664 To uncover TFs that drive subtype specification, we performed differential expression
665 analysis restricted to TFs annotated in the AnimalTFDB for *Drosophila melanogaster*
666 (Shen *et al.*, 2023). Candidate TFs with significant up-regulation in each neuronal
667 subtype were identified and ranked. Furthermore, leveraging the latent-time ordering
668 from scVelo, we modelled the dynamic expression of all genes along developmental
669 trajectories. We then refined this analysis using CellRank's multi-kernel framework—
670 combining connectivity, velocity, and CytoTRACE information—to pinpoint TFs whose
671 expression kinetics most strongly correlate with the latent time. This layered approach
672 highlighted key regulators of neuronal differentiation for subsequent validation using
673 HCR.

674

675 **Comparative scRNAseq analysis: Dataset Acquisition and Preprocessing**

676 To explore the evolutionary conservation of these regulators, we reanalysed published
677 single-cell RNA-seq datasets from *Danio rerio* and *Strongylocentrotus purpuratus*. The
678 zebrafish atlas spans 10 developmental stages from 10 hours post-fertilisation (hpf) to
679 10 days post-fertilisation (dpf) (Lange *et al.*, 2024). The sea urchin dataset was
680 generated from combining two datasets, comprising 10 time points from the 8-cell
681 stage to the 72 hpf larval stage (Foster, Oulhen and Wessel, 2020; Perillo *et al.*, 2020).
682 For zebrafish, raw 10x Genomics sequencing data were reprocessed with Cell Ranger
683 (v9.0.1; 10x Genomics) against the Ensembl *D. rerio* reference
684 (Danio_rerio.GRCz11.114), retaining both spliced and unspliced reads. Resulting
685 BAM files were converted to spliced/unspliced count matrices using Velocyto
686 (v0.17.17, (La Manno *et al.*, 2018), yielding loom files for downstream analysis. For
687 sea urchin, we used the author-provided count matrices and accompanying metadata
688 without remapping.

689

690 **Comparative scRNAseq analysis: Data Integration and Annotation**

691 Rather than reclustering the full datasets, we focused on neural lineages. For
692 zebrafish, central nervous system (CNS) populations were subset using the original
693 annotations provided in the source atlas (https://zebrahub.sf.czbiohub.org/data). For
694 sea urchin, we subset to neurogenic clusters based on the authors' labels.

695 Cross-species re-annotation was performed with orthologous marker genes
696 derived using eggNOG-based orthology mapping (described below), with the
697 *Drosophila melanogaster* subclustering marker set as the anchor. For sea urchin,
698 orthologues were called from the *S. purpuratus* proteome corresponding to the
699 reference genome used for mapping (GCF_000002235.4; Spur_4.2). Zebrafish
700 markers were referenced to Ensembl *D. rerio* GRCz11 (release 114). Monoaminergic
701 neuron populations were identified by expression of the orthologous canonical
702 biosynthetic enzymes and transporters.

703

704 **Comparative scRNAseq analysis: Trajectory and Dynamic Expression**
705 **Analyses**

706 RNA velocity was estimated using scVelo (Bergen *et al.*, 2020) and lineage inference
707 was refined with CellRank (v2) (Weiler *et al.*, 2024) by integrating connectivity and

708 velocity information. Latent-time ordering was used to model dynamic gene expression
709 trajectories and identify TFs with temporal activation patterns correlating with neuronal
710 differentiation.
711
712 **Comparative scRNAseq analysis: Differential Expression and Orthology**
713 **Mapping**
714 Candidate TFs enriched in each neuronal subtype were identified using the same
715 thresholds applied to the *Drosophila* dataset (adjusted $p < 0.05$; $|\log2FC| \geq 0.25$). For
716 cross-species comparison, we re-annotated the proteomes of zebrafish and
717 *Drosophila* with eggNOG (v5.0, Huerta-Cepas *et al.*, 2019), assigning genes to its
718 hierarchical orthogroup framework. We then conducted all comparative analyses at
719 the Bilateria-level orthogroups, which consolidate one-to-many and many-to-one
720 relationships, enabling robust mapping of putative orthologs and fair TF comparisons
721 across species.
722
723 **Comparative scRNAseq analysis: Comparative and Class-specific Analyses**
724 Shared and species-specific regulators were identified by comparing differentially
725 expressed TF orthogroups across *Drosophila* and zebrafish monoaminergic clusters.
726 To contextualise these findings, the orthogroup-level analysis was extended to other
727 neurotransmitter lineages (acetylcholine, GABA, glutamate).
728

729 ## Data availability

730 All data supporting this study are available on Zenodo. Processed FACS data are
731 deposited at https://doi.org/10.5281/zenodo.17449471, targeted metabolomics data at
732 https://doi.org/10.5281/zenodo.17449991, and processed single-cell RNA-seq data at
733 https://doi.org/10.5281/zenodo.17441737. All datasets are released under the MIT
734 License. Analysis pipelines and scripts are available at
735 https://github.com/cliftonlewis/2025_Drosophila_scRNAseq_EmbryoNeurogenesis_
736 Monoamine.
737

738 ## Author Contributions

739 R.F. conceived the project. C.L. performed the majority of sample preparation for
740 scRNA-seq, HCR, and targeted metabolomics, and carried out the bioinformatics

741 analyses. M.G. provided guidance on the bioinformatics analyses. A.W. assisted with
742 optimization of HCR protocols. N.C. and D.O. contributed to FACS sorting. S.R.
743 supported the initial project setup. R.P.Z. provided the initial protocols for targeted
744 scRNA-seq in *Drosophila* embryos. J.S. contributed to project development. E.R. and
745 C.P.K. supervised and supported the project. R.F. was the primary supervisor. C.L.
746 and R.F. wrote and edited the manuscript.
747

748 **Acknowledgements**

749 We gratefully acknowledge the support of the Royal Society through a University
750 Research Fellowship (UF160226 and URF/R/221011) and a Research Grant
751 (RGF\R1\181012) awarded to R.F. C.L. is supported by a BBSRC-MIBTP PhD
752 Studentship (Grant number MIBTP2020: BB/T00746X/1). The authors wish to thank
753 the technical and scientific staff of several core facilities for their invaluable support.
754 We are especially grateful for the assistance received from the University of Leicester
755 Advanced Imaging Facility (RRID: SCR_020967; BBSRC grant: BB/S019510/1) and
756 specifically thank Dr. Kees Straatman for his expert microscopy advice. This research
757 used the ALICE High Performance Computing Facility at the University of Leicester.
758 We thank the Flow Cytometry Facility, particularly Dr. David Onion and Nicola Croxall,
759 for their support with the Beckman Coulter MoFlo Astrios Flow Cytometer. We also
760 thank the DEEP-SEQ Genomics Facility, specifically Dr. Christopher Moore and Dr.
761 Nadine Holmes, for their expert assistance with single-cell genomics, including 10X
762 Genomics support. Additionally, we are deeply appreciative of Dr. Bernhard Drotleff
763 from the EMBL Metabolomics Facility for running the targeted metabolomics
764 experiments. We would also like to thank Dr. Diego Calderon (University of California
765 San Francisco) for help with adapting the cell type annotation pipeline. Finally, we
766 would like to thank the Feuda Lab for their thoughtful proofreading and constructive
767 feedback on the manuscript.

References

Arendt, D. (2005) “Genes and homology in nervous system evolution: Comparing gene functions, expression patterns, and cell type molecular fingerprints,” *Theory in Biosciences*, 124(2), pp. 185–197. Available at: https://doi.org/10.1016/j.thbio.2005.08.002.

Arendt, D. *et al.* (2019) “Evolution of neuronal types and families,” *Current Opinion in Neurobiology*, 56, pp. 144–152. Available at: https://doi.org/10.1016/j.conb.2019.01.022.

Babski, H., Codianni, M. and Bhandawat, V. (2024) “Octopaminergic descending neurons in *Drosophila*: Connectivity, tonic activity and relation to locomotion,” *Heliyon*, 10(9), p. e29952. Available at: https://doi.org/10.1016/j.heliyon.2024.e29952.

Bergen, V. *et al.* (2020) “Generalizing RNA velocity to transient cell states through dynamical modeling,” *Nature Biotechnology*, 38(12), pp. 1408–1414. Available at: https://doi.org/10.1038/s41587-020-0591-3.

Busch, S. *et al.* (2009) “A map of octopaminergic neurons in the Drosophila brain,” *The Journal of Comparative Neurology*, 513(6), pp. 643–667. Available at: https://doi.org/10.1002/cne.21966.

Cai, J. *et al.* (2009) “The role of Lmx1a in the differentiation of human embryonic stem cells into midbrain dopamine neurons in culture and after transplantation into a Parkinson’s disease model,” *Stem Cells (Dayton, Ohio)*, 27(1), pp. 220–229. Available at: https://doi.org/10.1634/stemcells.2008-0734.

Calderon, D. *et al.* (2022) “The continuum of *Drosophila* embryonic development at single-cell resolution,” *Science*, 377(6606), p. eabn5800. Available at: https://doi.org/10.1126/science.abn5800.

Campos-Ortega, J.A. and Hartenstein, V. (1985) “Stages of Drosophila Embryogenesis,” in Campos-Ortega, J. A. and Hartenstein, V., *The Embryonic Development of Drosophila melanogaster*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 9–84. Available at: https://doi.org/10.1007/978-3-662-02454-6_3.

Cheng, L. *et al.* (2003) “Lmx1b, Pet-1, and Nkx2.2 coordinately specify serotonergic neurotransmitter phenotype,” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 23(31), pp. 9961–9967. Available at: https://doi.org/10.1523/JNEUROSCI.23-31-09961.2003.

Choi, H.M.T. *et al.* (2018) “Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust,” *Development*, 145(12), p. dev165753. Available at: https://doi.org/10.1242/dev.165753.

Dillon, N. *et al.* (2022) “Single cell RNA-seq analysis reveals temporally-regulated and quiescence-regulated gene expression in Drosophila larval neuroblasts,” *Neural Development*, 17(1), p. 7. Available at: https://doi.org/10.1186/s13064-022-00163-7.

807 Ding, Y.-Q. *et al.* (2003) “Lmx1b is essential for the development of serotonergic
808 neurons,” *Nature Neuroscience*, 6(9), pp. 933–938. Available at:
809 https://doi.org/10.1038/nn1104.

810 Dittrich, R. *et al.* (1997) “The differentiation of the serotonergic neurons in the
811 Drosophila ventral nerve cord depends on the combined function of the zinc finger
812 proteins Eagle and Huckebein,” *Development*, 124(13), pp. 2515–2525. Available at:
813 https://doi.org/10.1242/dev.124.13.2515.

814 Doucet-Beaupré, H., Ang, S.-L. and Lévesque, M. (2015) “Cell fate determination,
815 neuronal maintenance and disease state: The emerging role of transcription factors
816 Lmx1a and Lmx1b,” *FEBS Letters*, 589(24PartA), pp. 3727–3738. Available at:
817 https://doi.org/10.1016/j.febslet.2015.10.020.

818 Flames, N. and Hobert, O. (2009) “Gene regulatory logic of dopamine neuron
819 differentiation,” *Nature*, 458(7240), pp. 885–889. Available at:
820 https://doi.org/10.1038/nature07929.

821 Flames, N. and Hobert, O. (2011) “Transcriptional Control of the Terminal Fate of
822 Monoaminergic Neurons,” *Annual Review of Neuroscience*, 34(Volume 34, 2011),
823 pp. 153–184. Available at: https://doi.org/10.1146/annurev-neuro-061010-113824.

824 Foster, S., Oulhen, N. and Wessel, G. (2020) “A single cell RNA sequencing
825 resource for early sea urchin development,” *Development*, 147(17), p. dev191528.
826 Available at: https://doi.org/10.1242/dev.191528.

827 Goulty, M. *et al.* (2023) “The monoaminergic system is a bilaterian innovation,”
828 *Nature Communications*, 14(1), p. 3284. Available at: https://doi.org/10.1038/s41467-
829 023-39030-2.

830 Goulty, M. *et al.* (2025) “Monoaminergic neurons share transcriptional identity across
831 Bilaterian animals.” bioRxiv, p. 2025.10.10.679534. Available at:
832 https://doi.org/10.1101/2025.10.10.679534.

833 Grace, A.A. (2016) “Dysregulation of the dopamine system in the pathophysiology of
834 schizophrenia and depression,” *Nature Reviews. Neuroscience*, 17(8), pp. 524–532.
835 Available at: https://doi.org/10.1038/nrn.2016.57.

836 Hammonds, A.S. *et al.* (2013) “Spatial expression of transcription factors in
837 Drosophilaembryonic organ development,” *Genome Biology*, 14(12), p. R140.
838 Available at: https://doi.org/10.1186/gb-2013-14-12-r140.

839 Hartenstein, V. *et al.* (2017) “Developmental analysis of the dopamine-containing
840 neurons of the Drosophila brain,” *The Journal of comparative neurology*, 525(2), pp.
841 363–379. Available at: https://doi.org/10.1002/cne.24069.

842 Hendricks, T.J. *et al.* (2003) “Pet-1 ETS gene plays a critical role in 5-HT neuron
843 development and is required for normal anxiety-like and aggressive behavior,”
844 *Neuron*, 37(2), pp. 233–247. Available at: https://doi.org/10.1016/s0896-
845 6273(02)01167-4.

Hobert, O. (2008) “Regulatory logic of neuronal diversity: terminal selector genes and selector motifs,” *Proceedings of the National Academy of Sciences of the United States of America*, 105(51), pp. 20067–20071. Available at: https://doi.org/10.1073/pnas.0806070105.

Huerta-Cepas, J. *et al.* (2019) “eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses,” *Nucleic Acids Research*, 47(D1), pp. D309–D314. Available at: https://doi.org/10.1093/nar/gky1085.

Karki, S. *et al.* (2023) “Serotonin signaling regulates actomyosin contractility during morphogenesis in evolutionarily divergent lineages,” *Nature Communications*, 14(1), p. 5547. Available at: https://doi.org/10.1038/s41467-023-41178-w.

Kasture, A.S. *et al.* (2018) “Big Lessons from Tiny Flies: Drosophila melanogaster as a Model to Explore Dysfunction of Dopaminergic and Serotonergic Neurotransmitter Systems,” *International Journal of Molecular Sciences*, 19(6), p. 1788. Available at: https://doi.org/10.3390/ijms19061788.

Korsunsky, I. *et al.* (2019) “Fast, sensitive and accurate integration of single-cell data with Harmony,” *Nature Methods*, 16(12), pp. 1289–1296. Available at: https://doi.org/10.1038/s41592-019-0619-0.

Kurian, M.A. *et al.* (2011) “The monoamine neurotransmitter disorders: an expanding range of neurological syndromes,” *The Lancet. Neurology*, 10(8), pp. 721–733. Available at: https://doi.org/10.1016/S1474-4422(11)70141-7.

La Manno, G. *et al.* (2018) “RNA velocity of single cells,” *Nature*, 560(7719), pp. 494–498. Available at: https://doi.org/10.1038/s41586-018-0414-6.

Lange, M. *et al.* (2022) “CellRank for directed single-cell fate mapping,” *Nature Methods*, 19(2), pp. 159–170. Available at: https://doi.org/10.1038/s41592-021-01346-6.

Lange, M. *et al.* (2024) “A multimodal zebrafish developmental atlas reveals the state-transition dynamics of late-vertebrate pluripotent axial progenitors,” *Cell*, 187(23), pp. 6742-6759.e17. Available at: https://doi.org/10.1016/j.cell.2024.09.047.

Libersat, F. and Pflueger, H.-J. (2004) “Monoamines and the Orchestration of Behavior,” *BioScience*, 54(1), pp. 17–25. Available at: https://doi.org/10.1641/0006-3568(2004)054%255B0017:MATOOB%255D2.0.CO;2.

Lundell, M.J. and Hirsh, J. (1994) “Temporal and spatial development of serotonin and dopamine neurons in the Drosophila CNS,” *Developmental Biology*, 165(2), pp. 385–396. Available at: https://doi.org/10.1006/dbio.1994.1261.

McGinnis, C.S., Murrow, L.M. and Gartner, Z.J. (2019) “DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors,” *Cell Systems*, 8(4), pp. 329-337.e4. Available at: https://doi.org/10.1016/j.cels.2019.03.003.

Monti, J.M. (2011) “Serotonin control of sleep-wake behavior,” *Sleep Medicine Reviews*, 15(4), pp. 269–281. Available at: https://doi.org/10.1016/j.smrv.2010.11.003.

Montioli, R. and Borri Voltattorni, C. (2021) “Aromatic Amino Acid Decarboxylase Deficiency: The Added Value of Biochemistry,” *International Journal of Molecular Sciences*, 22(6), p. 3146. Available at: https://doi.org/10.3390/ijms22063146.

Parent, A. (1984) “Functional Anatomy and Evolution of Monoaminergic Systems,” *American Zoologist*, 24(3), pp. 783–790. Available at: https://doi.org/10.1093/icb/24.3.783.

Peng, D. *et al.* (2024) “Organogenetic transcriptomes of the Drosophila embryo at single cell resolution,” *Development (Cambridge, England)*, 151(2), p. dev202097. Available at: https://doi.org/10.1242/dev.202097.

Perillo, M. *et al.* (2020) “Regulation of dynamic pigment cell states at single-cell resolution,” *eLife*, 9, p. e60388. Available at: https://doi.org/10.7554/eLife.60388.

Peter, I.S. and Davidson, E.H. (2011) “Evolution of gene regulatory networks controlling body plan development,” *Cell*, 144(6), pp. 970–985. Available at: https://doi.org/10.1016/j.cell.2011.02.017.

Rapti, G. (2022) “Open Frontiers in Neural Cell Type Investigations; Lessons From Caenorhabditis elegans and Beyond, Toward a Multimodal Integration,” *Frontiers in Neuroscience*, 15. Available at: https://doi.org/10.3389/fnins.2021.787753.

Reuter, B. *et al.* (2018) “Generalized Markov State Modeling Method for Nonequilibrium Biomolecular Dynamics: Exemplified on Amyloid β Conformational Dynamics Driven by an Oscillating Electric Field,” *Journal of Chemical Theory and Computation*, 14(7), pp. 3579–3594. Available at: https://doi.org/10.1021/acs.jctc.8b00079.

Reuter, B., Klein, M. and Lange, M. (2022) “pyGPCCA - python GPCCA: Generalized Perron Cluster Cluster Analysis package to coarse-grain reversible and non-reversible Markov state models.” Zenodo. Available at: https://doi.org/10.5281/zenodo.6914001.

Sachkova, M., Modepalli, V. and Kittelmann, M. (2025) “The Deep Evolutionary Roots of the Nervous System,” *Annual Review of Neuroscience*, 48(Volume 48, 2025), pp. 311–329. Available at: https://doi.org/10.1146/annurev-neuro-112723-040945.

Samuel Marsh *et al.* (2024) “samuel-marsh/scCustomize: Version 3.0.1.” Zenodo. Available at: https://doi.org/10.5281/ZENODO.5706430.

Sarov, M. *et al.* (2016) “A genome-wide resource for the analysis of protein localisation in Drosophila,” *eLife*. Edited by H.J. Bellen, 5, p. e12068. Available at: https://doi.org/10.7554/eLife.12068.

Schindelin, J. *et al.* (2012) “Fiji: an open-source platform for biological-image analysis,” *Nature Methods*, 9(7), pp. 676–682. Available at: https://doi.org/10.1038/nmeth.2019.

Shen, W.-K. *et al.* (2023) “AnimalTFDB 4.0: a comprehensive animal transcription factor database updated with variation and expression annotations,” *Nucleic Acids Research*, 51(D1), pp. D39–D45. Available at: https://doi.org/10.1093/nar/gkac907.

Tomancak, P. *et al.* (2002) “Systematic determination of patterns of gene expression during Drosophila embryogenesis,” *Genome Biology*, 3(12), p. research0088.1. Available at: https://doi.org/10.1186/gb-2002-3-12-research0088.

Tomancak, P. *et al.* (2007) “Global analysis of patterns of gene expression during Drosophilaembryogenesis,” *Genome Biology*, 8(7), p. R145. Available at: https://doi.org/10.1186/gb-2007-8-7-r145.

Volkow, N.D. *et al.* (2011) “Motivation deficit in ADHD is associated with dysfunction of the dopamine reward pathway,” *Molecular Psychiatry*, 16(11), pp. 1147–1154. Available at: https://doi.org/10.1038/mp.2010.97.

Waddell, S. (2013) “Reinforcement signalling in Drosophila; dopamine does it all after all,” *Current Opinion in Neurobiology*, 23(3), pp. 324–329. Available at: https://doi.org/10.1016/j.conb.2013.01.005.

Weiler, P. *et al.* (2024) “CellRank 2: unified fate mapping in multiview single-cell data,” *Nature Methods*, 21(7), pp. 1196–1205. Available at: https://doi.org/10.1038/s41592-024-02303-9.

Wijnen, H. *et al.* (2006) “Control of Daily Transcript Oscillations in Drosophila by Light and the Circadian Clock,” *PLOS Genetics*, 2(3), p. e39. Available at: https://doi.org/10.1371/journal.pgen.0020039.

Yamamoto, K. and Vernier, P. (2011) “The Evolution of Dopamine Systems in Chordates,” *Frontiers in Neuroanatomy*, 5. Available at: https://doi.org/10.3389/fnana.2011.00021.

Zheng, Q. *et al.* (2025) “Bidirectional histone monoaminylation dynamics regulate neural rhythmicity,” *Nature*, 637(8047), pp. 974–982. Available at: https://doi.org/10.1038/s41586-024-08371-3.

Figures

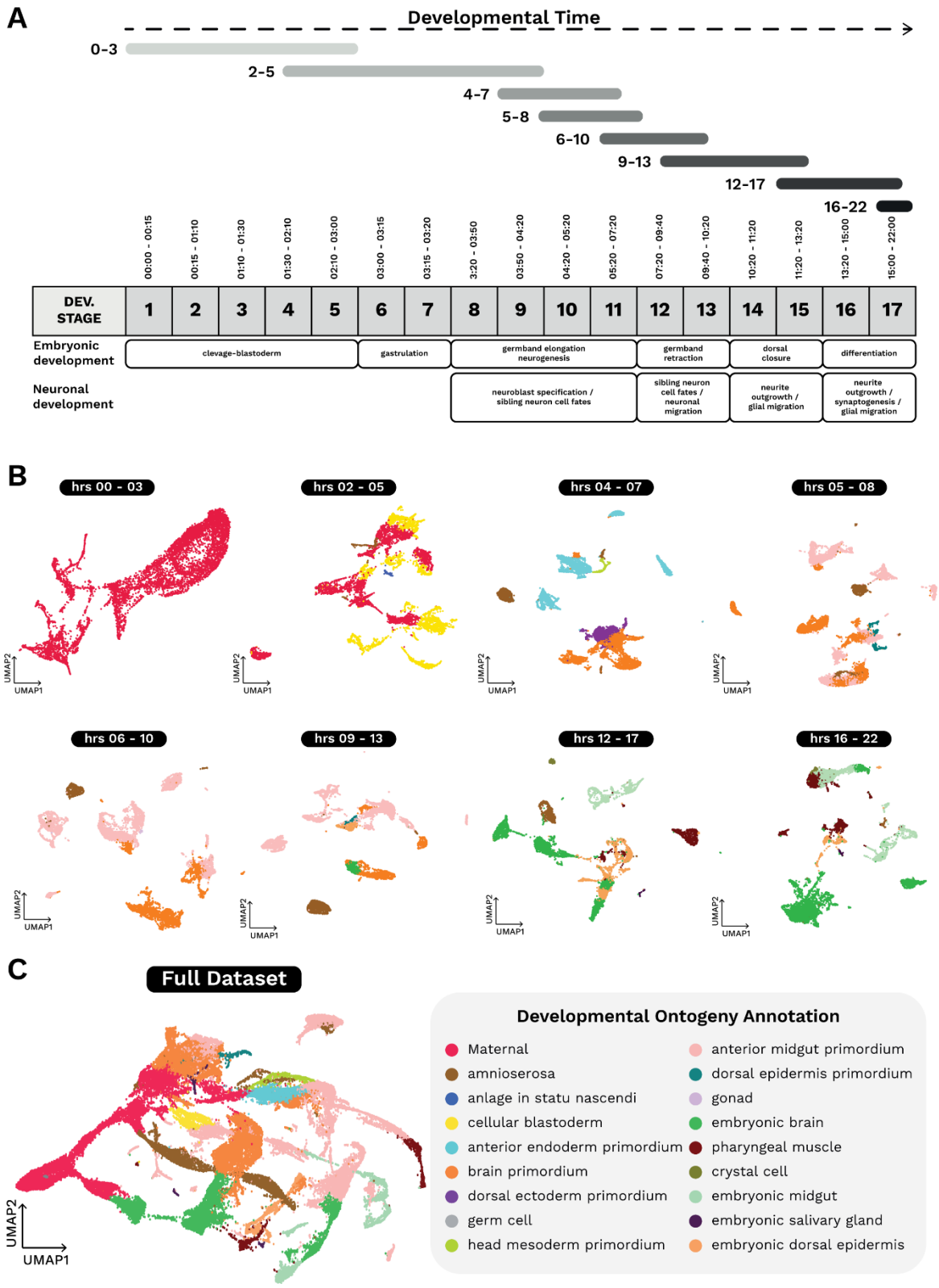

A
Developmental Time
0-3
2-5
4-7
5-8
6-10
9-13
12-17
16-22
00:00 - 00:15
00:15 - 01:10
01:10 - 01:30
01:30 - 02:10
02:10 - 03:00
03:00 - 03:15
03:15 - 03:20
3:20 - 03:50
03:50 - 04:20
04:20 - 05:20
05:20 - 07:20
07:20 - 09:40
09:40 - 10:20
10:20 - 11:20
11:20 - 13:20
13:20 - 15:00
15:00 - 22:00
DEV. STAGE
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Embryonic development
clevage-blastoderm
gastrulation
germband elongation neurogenesis
germband retraction
dorsal closure
differentiation
Neuronal development
neuroblast specification / sibling neuron cell fates
sibling neuron cell fates / neuronal migration
neurite outgrowth / glial migration
neurite outgrowth / synaptogenesis / glial migration
B
hrs 00 - 03
hrs 02 - 05
hrs 04 - 07
hrs 05 - 08
hrs 06 - 10
hrs 09 - 13
hrs 12 - 17
hrs 16 - 22
UMAP2
UMAP1
C
Full Dataset
Developmental Ontogeny Annotation
Maternal
amnioserosa
anlage in statu nascendi
cellular blastoderm
anterior endoderm primordium
brain primordium
dorsal ectoderm primordium
germ cell
head mesoderm primordium
anterior midgut primordium
dorsal epidermis primordium
gonad
embryonic brain
pharyngeal muscle
crystal cell
embryonic midgut
embryonic salivary gland
embryonic dorsal epidermis

Figure 1. Time-resolved sampling strategy and developmental ontogeny annotation of targeted scRNA-seq across *Drosophila* embryogenesis.

A. Timeline of eight partially overlapping collection windows, spanning embryogenesis, from 0–22 hrs after egg laying (AEL) aligned to canonical embryonic stages (Bownes stages 1–17, (Campos-Ortega and Hartenstein, 1985)) and key morphogenetic landmarks (cleavage/blastoderm, gastrulation, germband elongation, germband retraction, dorsal closure, and differentiation). Major neurogenic milestones are indicated below the timeline, including neuroblast specification, diversification of sibling neuron fates, neuronal migration, neurite outgrowth, glial migration, and synaptogenesis. **B.** UMAPs for each time window showing single-cell transcriptomes coloured by developmental ontology. **C.** UMAP of the combined dataset with cells coloured by developmental ontology derived from *in situ*–guided annotation. Representative labels include maternal, cellular blastoderm, amnioserosa, dorsal ectoderm primordium, brain primordium, embryonic brain, and embryonic midgut.

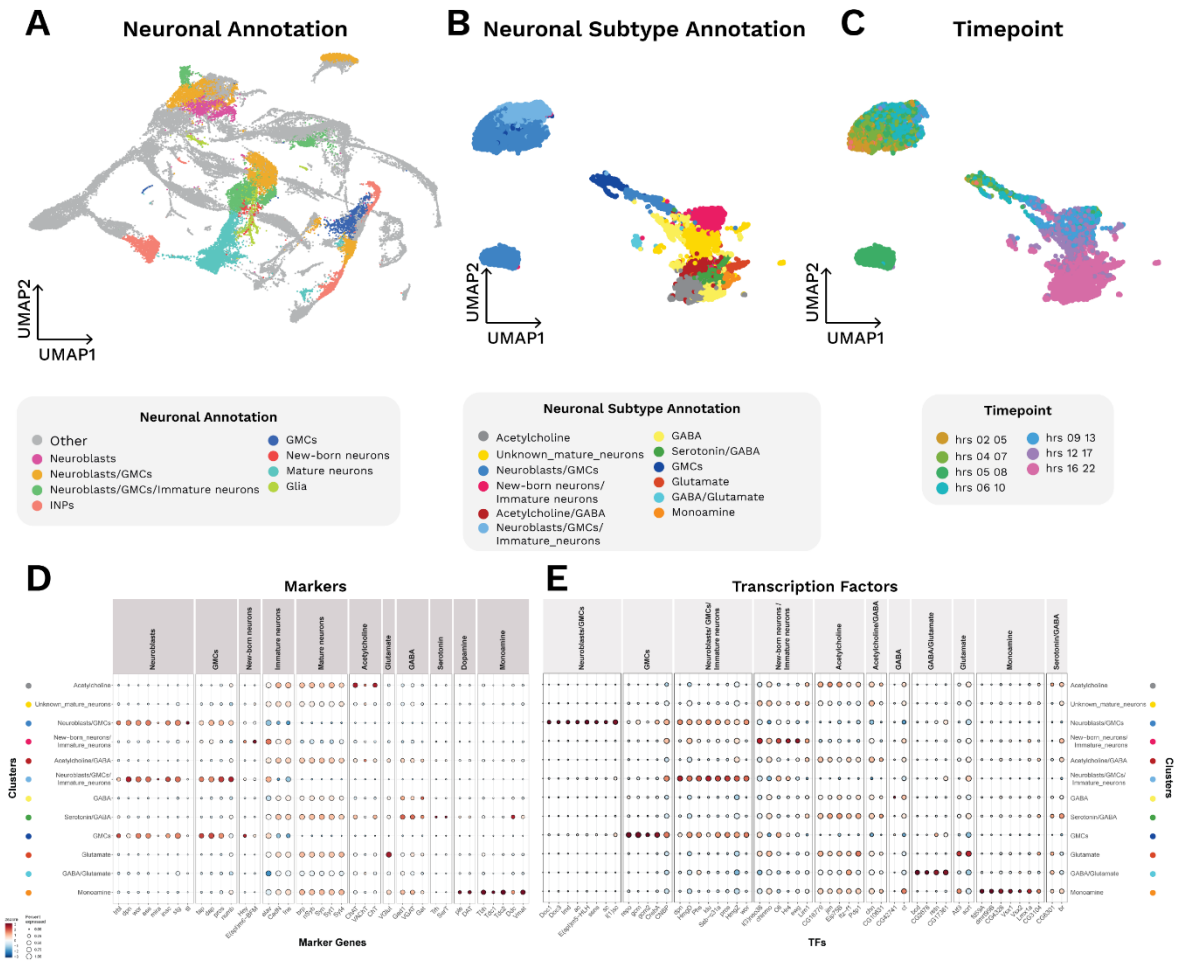


Figure 2. Time-resolved single-cell landscape of embryonic neurogenesis in *Drosophila*

A. Neuronal annotation. UMAP of the full embryonic dataset with neural lineages highlighted and non-neuronal cells in grey. Colours denote major neural states: neuroblasts, neuroblasts/GMCs, neuroblasts/GMCs/immature neurons, intermediate neural progenitors (INPs), GMCs, new-born/immature neurons, mature neurons, and glia. **B. Neuronal subtype annotation.** Subclustered neurogenic atlas reannotated to subtypes including transmitter-class/subtype as well as neuronal developmental intermediates. Classes include acetylcholine, GABA, glutamate, monoamine, mixed identities (e.g., acetylcholine/GABA; GABA/glutamate), unknown mature neurons, and progenitor states (neuroblasts/GMCs; neuroblasts/GMCs/immature neurons; GMCs; new-born/immature neurons). **C. Timepoint.** Subclustered dataset coloured by collection window, showing temporal separation across the UMAP. **D. Markers.** Dot plot of selected marker genes across the neuronal states/subtypes in the subclustered dataset. Dot size indicates the fraction of cells expressing the gene; colour reflects average expression. **E. Differential transcription factors.** Dot plot of differential TFs across neuronal states/subtypes, highlighting regulators associated with progenitors, early differentiation, transmitter specification, and mature neurons.

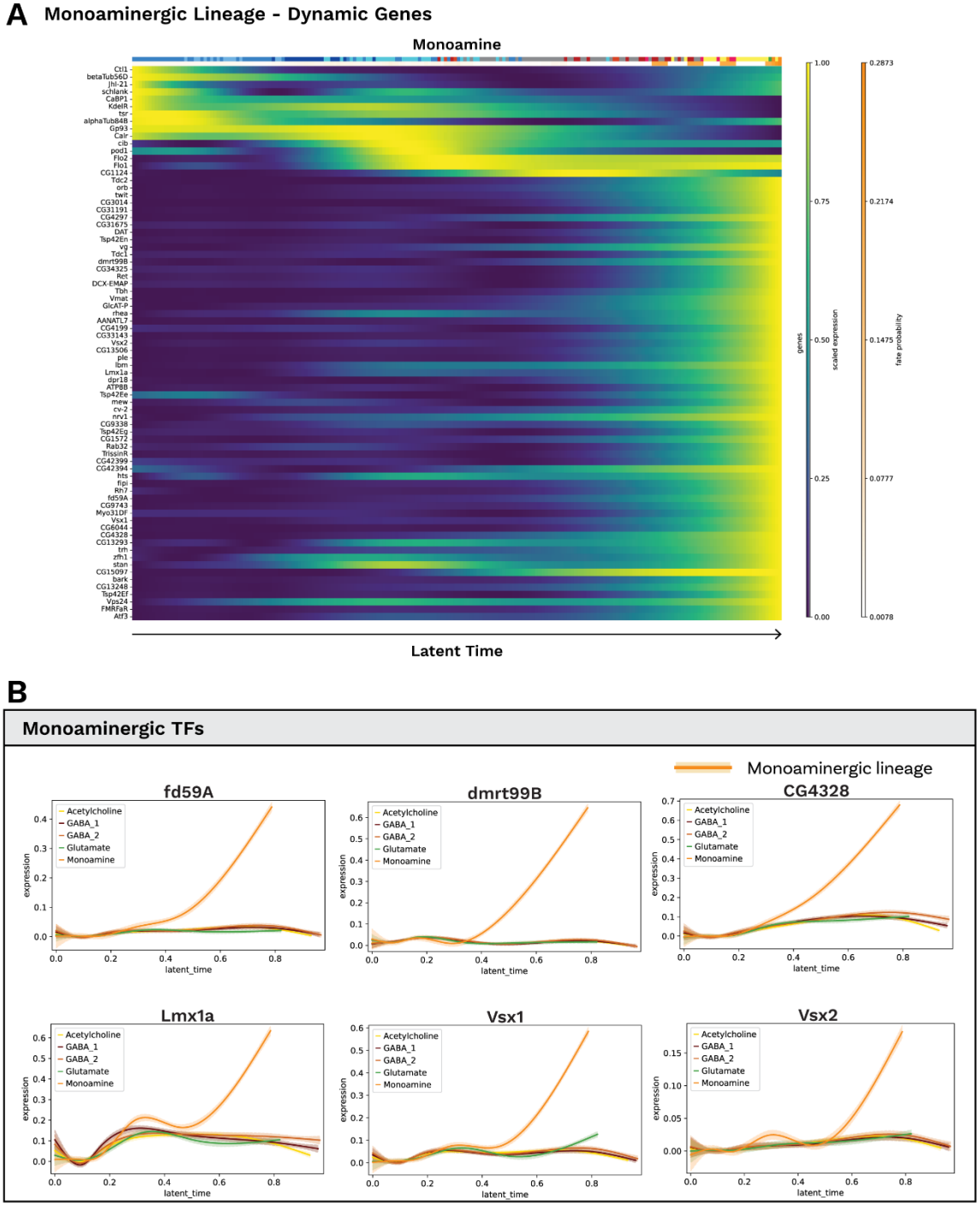
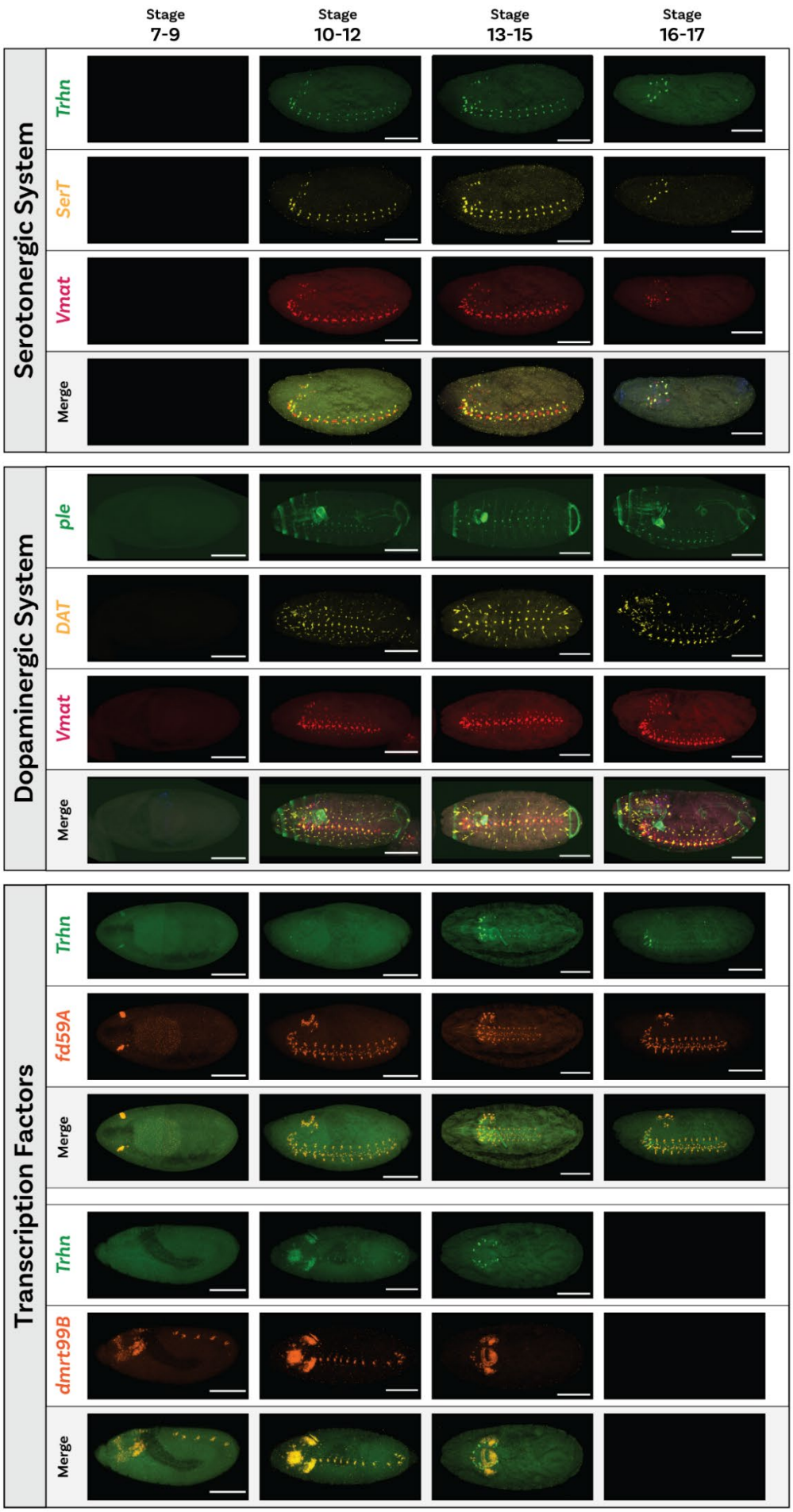


Figure 3. Monoaminergic lineage—dynamic gene programs and candidate transcription factors gene trends

A. Dynamic genes along latent time. Heatmap of CellRank-identified dynamical genes within the monoaminergic lineage. Cells are ordered by latent time (left→right; arrow) and expression is z-scaled per gene (yellow = higher expression). Many neurotransmitter-pathway genes and regulators turn on progressively toward late latent time as well as many of the candidate differentially expressed TFs identified

previously. **B. Monoaminergic transcription factors.** Smoothed expression trajectories across latent time for selected TFs (*fd59A*, *dmrt99B*, *CG4328*, *Lmx1a*, *Vsx1*, *Vsx2*). The monoaminergic lineage is highlighted in orange; other neuronal lineages (acetylcholine, GABA_1/2, glutamate) are shown for comparison. Curves show lineage-wise mean trends (y-axis, scaled expression) across latent time (x-axis), illustrating the preferential late-stage upregulation of these TFs in the monoaminergic lineage specifically.


Stage 7-9
Stage 10-12
Stage 13-15
Stage 16-17
Serotonergic System
Trhn
SerT
Vmat
Merge
Dopaminergic System
ple
DAT
Vmat
Merge
Transcription Factors
Trhn
fd59A
Merge
Trhn
dmrt99B
Merge

1005

Figure 4. HCR validates the temporal onset and spatial organisation of serotonergic and dopaminergic programs in the embryo

HCR staining across four developmental windows (columns: stages 7–9, 10–12, 13–15, 16–17). Each image shows a whole-mount embryo, lateral view with anterior to the left. Scale bars 100 µm. **Serotonergic system**. HCR for *Trhn* (*tryptophan hydroxylase*, green), *SerT* (*serotonin transporter*, yellow) and *Vmat* (*vesicular monoamine transporter*, magenta), plus merges. No signal is detected at stages 7–9. From stage 10–12 onwards *Trhn*, *SerT* and *Vmat* appear in segmentally repeated CNS domains; marking serotonergic neurons as they mature. **Dopaminergic system**. HCR for *ple* (*Tyrosine hydroxylase*, green), *DAT* (*dopamine transporter*, yellow) and *Vmat* (magenta), plus merges. *ple* is first detected by stage 10–12, with *DAT* and *Vmat* also appearing by stage 10-12. *ple*/*DAT* signals co-localise with *Vmat* in the CNS. **Transcription factors**. HCR for *fd59A* or *dmrt99B* (orange) together with *Trhn* (green), plus merges. TF expression is evident in the CNS earlier (stages 7-9) and broadens by stages 10-12, preceding with enzyme onset, consistent with a putative upstream regulatory role in monoaminergic fate.

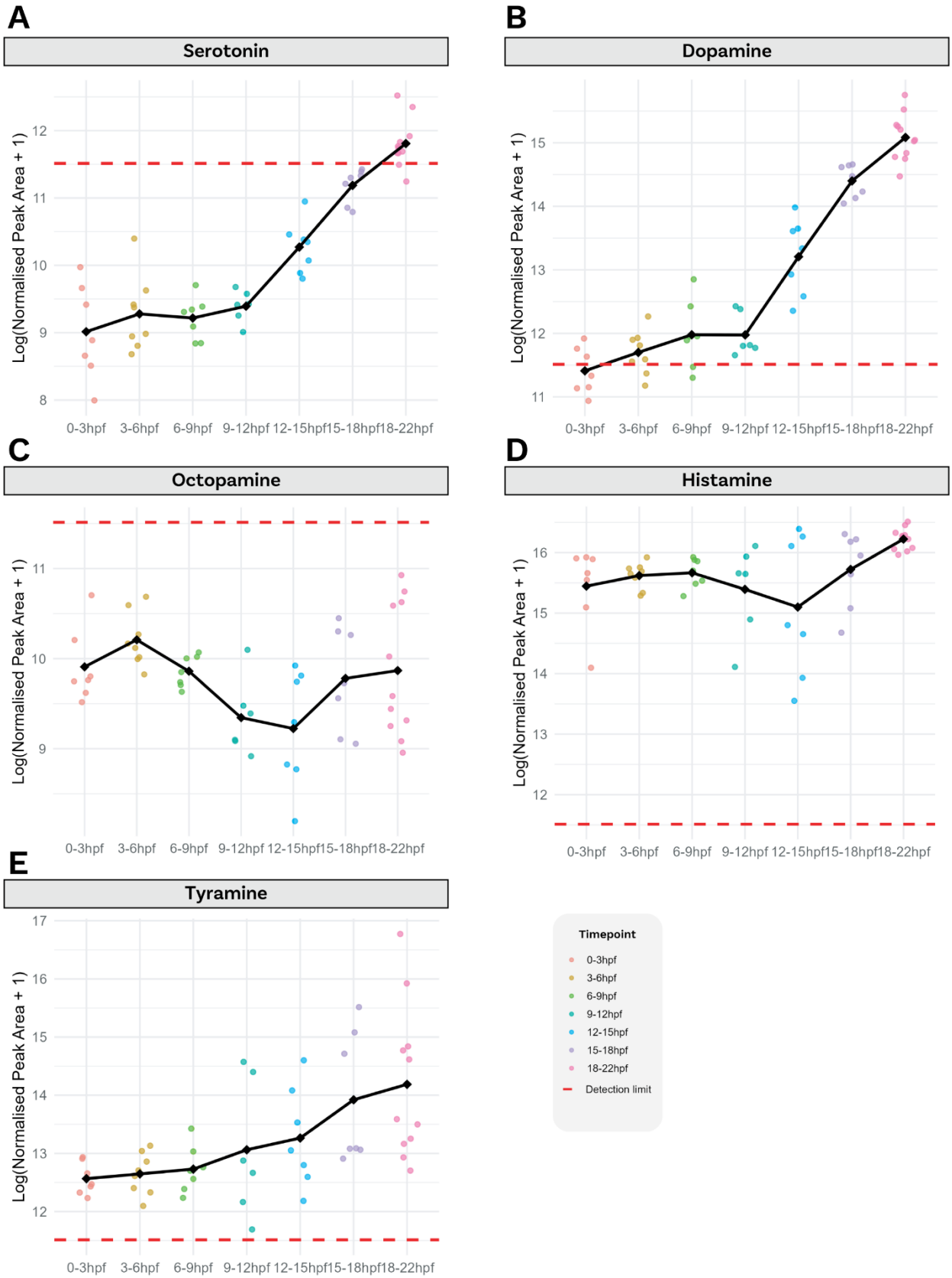
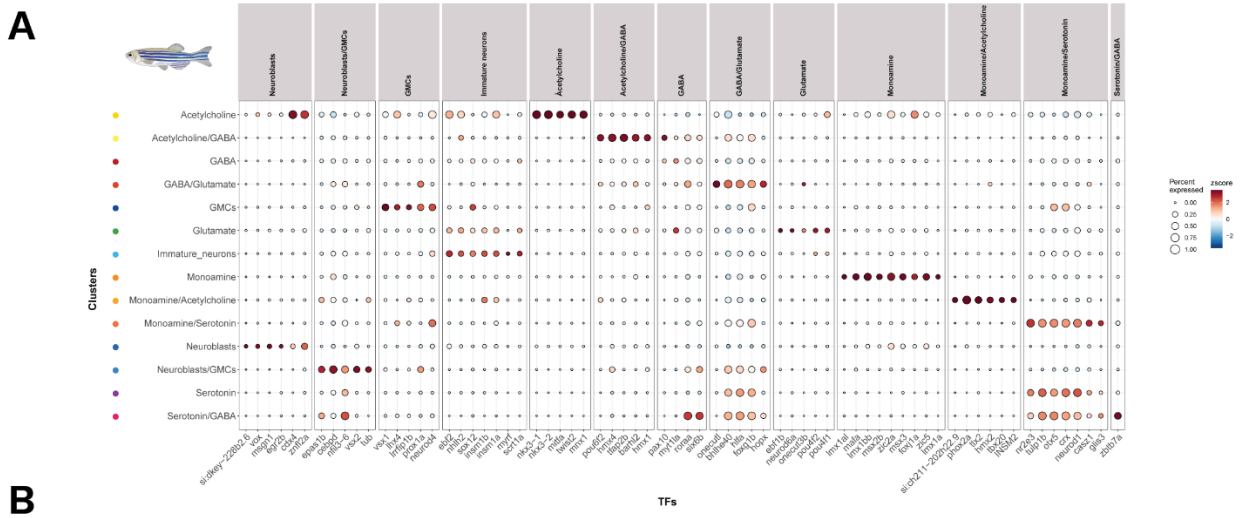
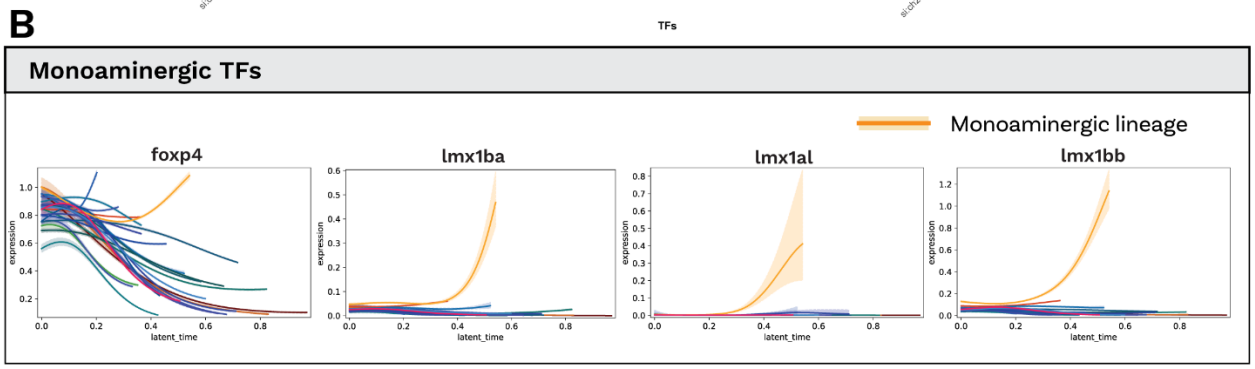
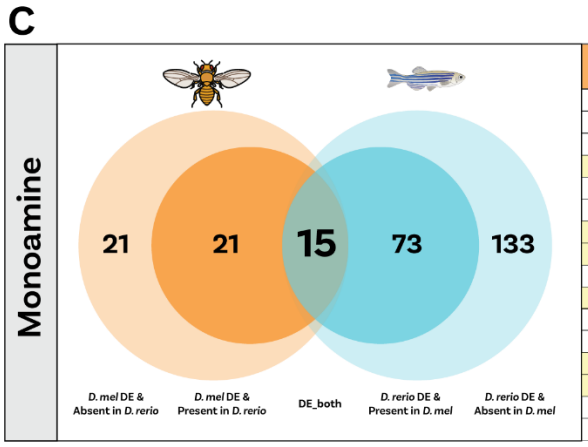

A
Serotonin
Log(Normalised Peak Area + 1)
12
11
10
9
8
0-3hpf
3-6hpf
6-9hpf
9-12hpf
12-15hpf
15-18hpf
18-22hpf
B
Dopamine
Log(Normalised Peak Area + 1)
15
14
13
12
11
0-3hpf
3-6hpf
6-9hpf
9-12hpf
12-15hpf
15-18hpf
18-22hpf
C
Octopamine
Log(Normalised Peak Area + 1)
11
10
9
0-3hpf
3-6hpf
6-9hpf
9-12hpf
12-15hpf
15-18hpf
18-22hpf
D
Histamine
Log(Normalised Peak Area + 1)
16
15
14
13
12
0-3hpf
3-6hpf
6-9hpf
9-12hpf
12-15hpf
15-18hpf
18-22hpf
E
Tyramine
Log(Normalised Peak Area + 1)
17
16
15
14
13
12
0-3hpf
3-6hpf
6-9hpf
9-12hpf
12-15hpf
15-18hpf
18-22hpf
Timepoint
0-3hpf
3-6hpf
6-9hpf
9-12hpf
12-15hpf
15-18hpf
18-22hpf
Detection limit

Figure 5. Targeted metabolomics across embryonic time reveals distinct onset of monoamine production

A–E. Concentrations of five monoamines measured by targeted LC–MS/MS across seven developmental windows (0–3, 3–6, 6–9, 9–12, 12–15, 15–18, 18–22 h AEL). Each dot is one biological replicate (≥6 per time point); colour encodes the time window. The black line connects the mean of replicates at each time. The y-axis shows $\log_{10}$(normalised peak area + 1). The red dashed line marks the empirical detection threshold for each analyte. **A. Serotonin** and **B. Dopamine:** low/near-baseline early, followed by a sharp rise beginning ~12–15 h, consistent with the transcriptional induction of biosynthetic enzymes. **C. Octopamine:** largely under the detection threshold. **D. Histamine** and **E. Tyramine**: detectable throughout development with gradual increases. Together, these profiles indicate that the functional availability of serotonin and dopamine lags their transcriptional commitment, whereas histamine and tyramine are present earlier and more continuously.

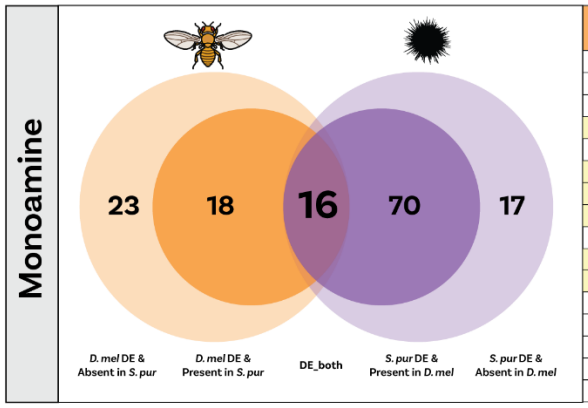






Conserved TF Orthogroups

D. melanogaster	*D. rerio*
cic	*cica*
Nfl	*nfia; nfixb*
cbt	*klf11b; klf11a*
CG4328; Lmx1a	*lmx1bb; lmx1al; lmx1ba*
FoxP	*foxp4*
trh	*NPAS3*
Eip78C	*nr1d1*
net	*atoh8*
tj	*mafa; mafb*
Pdp1	*hlfa*
luna	*klf6a*
Camta	*camta1a*
CG8301	*BX324216.3*
Hr3	*rorca; rorb*
shn	*hivep1*



Conserved TF Orthogroups

D. melanogaster	*S. purpuratus*
Vsx1; Vsx2	*LOC575706*
sim	*LOC577676*
Pur-alpha	*LOC579034;LOC105436940*
CG8301	*LOC100889018*
zfh1	*LOC594820*
CG4328; Lmx1a	*LOC585635*
Hr3	*LOC578618*
Eip78C	*LOC756187*
sr	*LOC100892559*
net	*LOC591662*
Pdp1	*Hlf*
Tusp	*LOC591879*
fd59A	*LOC751845*
Hr38	*LOC581160*
CG3104	*LOC592680*
Eip75B	*LOC592902; LOC580686*

Figure 6. Cross-species conservation of monoaminergic TF programs

A. Differentially expressed TFs in *D. rerio* neurogenic subtypes. Dot plot of differentially expressed TFs across zebrafish neurogenic subtypes. Columns are grouped by transmitter/state classes (Acetylcholine, Glutamate, GABA, Serotonin, Dopamine, Monoamine; plus, developmental states such as Neuroblasts, GMCs, Immature/Mature neurons). Dot size shows the fraction of cells in the subtype expressing the TF; colour shows the z-scored average expression. **B. Monoaminergic TF dynamics.** Smoothed latent-time expression trajectories for TFs associated with monoaminergic fate (examples shown: *foxp4*, *lmx1ba*, *lmx1a1*, *lmx1a*). Curves are plotted for the major neuronal lineages; the monoaminergic lineage is highlighted in orange with a light-orange confidence band, illustrating late-latent-time up-regulation of these TFs along the monoaminergic trajectory. **C. Orthogroup-level conservation analysis.** Venn diagram summarising Bilateria eggNOG orthogroups containing TFs that are differentially expressed in monoaminergic clusters in *D. melanogaster* and *D. rerio* and *D. melanogaster* and *S. purpuratus* (criteria as in Methods: adjusted $p < 0.05$; $|\log2FC| \geq 0.25$). Left/right circles denote species-specific DE orthogroups; the intersection denotes orthogroups DE in both species. The table lists representative conserved TF orthogroups with the corresponding gene names in each species. Together, these panels show that a core set of TF orthogroups is shared across insects and vertebrates while each lineage also exhibits species-specific components of the monoaminergic program.