Evolutionary dynamics of temporal transcription factor series in the insect optic lobe

Konstantina Filippopoulou[1], Elisavet Iliopoulou[1], Claire Julliot de La Morandière[1], Christy Lee[2], Marina Marcet-Houben[3,4,5], Toni Gabaldón[3,4,5,6], Jingyi Jessica Li[2,7], Nikolaos Konstantinides[1]

Affiliations

[1] Université Paris Cité, CNRS, Institut Jacques Monod, F-75013 Paris, France
[2] Department of Statistics and Data Science, University of California, Los Angeles, CA, USA
[3] Barcelona Supercomputing Centre (BSC-CNS). Plaça Eusebi Güell, 1-3 08034 Barcelona, Spain
[4] Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac, 10, 08028 Barcelona, Spain
[5] CIBER de Enfermedades Infecciosas, Instituto de Salud Carlos III, Madrid, Spain
[6] Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain
[7] Biostatistics Program, Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, WA, USA

Correspondence: nikos.konstantinides@ijm.fr

Abstract

The nervous system is composed of a wide diversity of neuronal cell types arranged into complex circuits that support a broad range of behaviors. Patterning of neural stem cells in time through the expression of series of temporal transcription factors is a key contributor to neuronal diversity. How do temporal series arise and diversify across species to support neuronal type evolution? Here, we reconstruct the evolutionary history of the visual brain temporal series in insects; we identify a conserved temporal ground plan, as well as species-specific variations. We find that temporal programs evolve through recurrent modifications of a shared scaffold. Finally, we show how such evolutionary changes in temporal patterning can translate into altered neuronal type identities. Together, our results reveal both the deep conservation and the evolutionary plasticity of temporal patterning programs, and establish temporal transcription factor series as a tractable substrate for the evolution of neuronal diversity.

Main text

The morphological diversity of the different neuronal types has been appreciated for more than a century, i.e. since the first time brains were observed under a microscope. Apart from diverse morphologies, neuronal types differ in other fundamental characteristics, such as connectivity, physiology, and molecular identity. In evolutionary neurobiology, comparative studies examine how these neuronal properties diverge across species, providing insight into how nervous systems adapt to distinct ecological pressures and behavioral demands(*1–4*). However, the ensemble of these properties that endows every neuron with a unique identity emerges during embryonic development, when the neuron is specified. Therefore, understanding neuronal evolution requires uncovering how developmental programs themselves have changed, since modifications in gene regulation, stem cell lineage progression and neuronal output can ultimately give rise to the divergent neuronal types observed across species.

How are different neuronal types specified? Research over the last decades has shown that the combination of spatial and temporal patterning of neuronal progenitors - i.e. their position within the tissue and their developmental stage, respectively - is the major mechanism underlying the generation of neuronal diversity in both invertebrates and vertebrates(*5–7*). Temporal patterning is driven by the sequential expression of transcription factors (temporal transcription factors - tTFs), which alter the ability of neural stem cells to produce distinct neuronal types(*8*, *9*).

Although temporal patterning was first functionally described in the vertebrate brain(*10*, *11*), the first tTF series was identified in the *Drosophila* ventral nerve cord(*12*, *13*), the counterpart of the mammalian spinal cord. Since then, multiple tTF series have been discovered across different neural stem cell populations in the *Drosophila* central nervous system(*8*, *14–17*). Interestingly, while the principle of temporal patterning is broadly conserved, the specific tTFs employed differ markedly across neuronal tissues within the same or different species (**Fig. 1A**). Notably, while the vertebrate spinal cord tTF series shares no homology with any of the fly temporal cascades, it appears to be conserved across vertebrate species and largely shared with the developing vertebrate brain(*9*).

At the same time, several isolated cases of conserved tTF series have emerged. The most compelling example is the expression of three *Drosophila* ventral nerve cord tTF orthologs - Hb/Ikzf1, Cas/Casz1, and Pdm/Pou2f2 - as temporal factors in the developing mouse retina(*18–20*) (**Fig. 1A**). Similarly, individual *Drosophila* tTFs are deployed temporally in various vertebrate neural stem cell populations even when the remainder of the cascade is not conserved(*8*, *21*). These scattered but intriguing parallels suggest that some components of temporal patterning may be evolutionarily ancient, while others might have diversified or been replaced. Thus, to address the evolution of neuronal development, we ultimately need to understand how temporal series in neural stem cells arise and diversify across species.

Of course, this question cannot be addressed by comparing animals as evolutionarily distant as flies and mice. Moreover, to study the evolution of temporal series, one needs a neuronal tissue model that is comprehensively described, cell-type resolved, and comparatively accessible, allowing the identification of both conserved and divergent mechanisms. The insect visual system - specifically the medulla region of the optic lobes - provides exactly such a model. This system is sufficiently conserved across insects to enable meaningful comparative analyses(*22*), yet it is experimentally tractable, and its temporal patterning program has been dissected in exceptional detail by our group and others(*8*, *14*): in the *Drosophila* medulla, a series of 14 tTFs (*hth, dmrt99B, erm, opa, L, ey, hbn, scro, slp1, slp2, BarHI, BarH2,* and *tll*) has been mapped, forming the longest and best-resolved tTF cascade described in any neural stem cell population, leading to the generation of over 250 distinct neuronal types (*23–26*). Notably, the progression of this series is ensured by the sequential application of temporal

gene regulatory logic - including repression of the predecessor, sequential activation, and modulatory inputs(*8*) - that together enforce the unidirectional flow of the cascade and couple stem-cell age to neuronal identity.

Here, we present the first systematic study of how a fundamental neurodevelopmental program evolves across an entire insect lineage spanning ~400 million years. We dissected developing optic lobes of eight insect species at their peak proliferative stages, performed single-cell mRNA sequencing, reconstructed neuroblast trajectories, and validated candidate tTFs by Hybridization Chain Reaction Fluorescence In Situ Hybridization (HCR-FISH) (**Fig. 1B**). By tracing the evolutionary history of optic lobe temporal series, we uncover a conserved core of tTFs present in the last common ancestor of hemimetabolous and holometabolous insects, alongside divergent factors that likely allowed insects to adapt their neuronal repertoire to defined ecological pressures. We further identify four principal evolutionary mechanisms shaping temporal series divergence and predict their impact on lineage-specific neuronal identities. Together, these results provide the first unified view of how temporal patterning programs evolve, revealing the rules by which neural stem cells diversify neuronal types across hundreds of millions of years.

The developing optic lobe in single-cell resolution

We first selected representatives from different insect orders that diverged from *D. melanogaster* between 50 and 420 million years ago (mya), spanning the insect phylogenetic tree and with high-quality genomes available: the fruit fly *Drosophila virilis* (Diptera, 50 mya), the house fly *Musca domestica* (Diptera, 120 mya), the mosquito *Aedes aegypti* (Diptera, 170 mya), the silkworm *Bombyx mori* (Lepidoptera, 290 mya), the beetle *Tribolium castaneum* (Coleoptera, 320 mya), the wasp *Nasonia vitripennis* (Hymenoptera, 350 mya), the cricket *Gryllus bimaculatus* (Orthoptera, 390 mya), and the mayfly *Cloeon dipterum* (Ephemeroptera, 420 mya)(*27*) (**Fig. 1C** and **Table S4**). Second, we identified the developmental stages at which neurogenesis occurs in the optic lobes of each species. Given the substantial differences in life cycles, we focused on stages showing the most extensive proliferation in the developing optic lobes (**Fig. 1D**). To this end, we performed EdU pulse-chase experiments (**Fig. 1D** and **Fig. S1A–D**), which revealed the following proliferative stages: *D. virilis* L3, *Musca* L3, *Aedes* L4, *Bombyx* L5, *Tribolium* prepupa, and *Nasonia* L3. For hemimetabolous insects lacking true larval stages, we assayed nymphal stages and selected *Gryllus* nymphal stages 1, 3, and 5, and *Cloeon* mid-nymphal stage, as periods of high optic lobe neuroblast proliferation (**Fig. S1E–F** and **Methods**). Finally, we developed cell dissociation protocols for each species' optic lobes to obtain sufficient numbers of single cells for downstream applications (**Methods**).

We performed single-cell mRNA sequencing using the 10X Genomics Single Cell 3′ Reagent Kit v3.1. To improve genome annotations, we combined these single-cell data with adult single-cell RNA-seq datasets from the same species and applied GeneExt(*28*) (**Methods** and **Fig. S2**). After mapping reads to the GeneExt-improved genomes and following initial quality control and preprocessing, we recovered a total of 304,485 cells across all species (**Methods** and **Fig. S3**).

In the insect optic lobes, neuroepithelial cells are converted into neuronal stem cells, or neuroblasts, which divide asymmetrically to self-renew and produce intermediate precursors, called ganglion mother cells (GMCs). GMCs then divide once to generate two neurons or glial cells(*29*). To identify the different progenitor and differentiated cell types, we examined the expression of conserved markers: *shotgun* (*shg*) for neuroepithelial structures and neuroblasts(*30*), *miranda* (*mira*), *deadpan* (*dpn*) and *worniu* (*wor*) for neuroblasts(*31*, *32*), *asense* (*ase*) for neuroblasts and GMCs(*33*), *found in neurons (fne)*, *embryonic lethal abnormal vision* (*elav*)(*34*), *neuronal Synaptobrevin* (*nSyb*), and *bruchpilot* (*brp*)(*35*) for neurons, and

reversed polarity (*repo*), *Excitatory amino acid transporter 1* (*Eaat1*), and *axotactin* (*axo*) for glia(*36*) (**Fig. 1E** and **Fig. S4**). As observed in the equivalent *D. melanogaster* dataset(*8*), the two-dimensional Uniform Manifold Approximation and Projection (UMAP) representations recapitulated the progression from progenitors to differentiated cells: *mira*- or *dpn*-positive neuroblasts (yellow) occupy central positions in the UMAP plots, leading into *nSyb*-, *brp*- or *fne*- positive neuronal trajectories (blue), while most *repo*-positive glia form relatively isolated clusters (red). The only exception is in hemimetabolous insects (Gryllus and Cloeon), where some neurons appear distant from neuroblasts; these isolated neurons likely represent embryonically born neurons that have already differentiated and become active in the hatchling.

Finally, because our dissections and dissociations did not allow us to separate the individual optic lobe neuropils (lamina, medulla, and lobula complex), we relied on conserved molecular markers to identify them in our datasets. Both the lamina and medulla originate from the Outer Proliferation Center (OPC), which generates medulla neuroblasts medially and lamina precursor cells laterally. To identify lamina neurons, we searched for conserved markers of this lineage - *dachsund* (*dac*), *eyes absent* (*eya*), *single-minded* (*sim*), *glial cells missing* (*gcm*), and *tailless* (*tll*)(*8*, *37*, *38*) (**Fig. S5**). Notably, most of these genes are reliable lamina markers in all insects, the most conserved ones being *eya*, *sim*, and *dac*. Neurons of the lobula plate (T4/T5) as well as T and C neurons (T2/T3 and C2/C3) originate from the Inner Proliferation Center (IPC), which is defined by the expression of *tll*, *dac*, and *acj6*(*39*, *40*) (**Fig. S5**). Using these genes, we were able to define lobula plate cell types in all insects, except for *Gryllus* (**Fig. 1F**). Having identified the major neuropils, we next investigated the neural stem cell populations present in our datasets. We examined the expression of conserved *Drosophila* stem cell markers such as *mira* (**Fig. 1E**), *dpn*, and *wor* (**Fig. S4**). To identify medulla neuroblasts and distinguish them from differentiating cells, we used the co-expression of *mira* and *dpn*, a well-established hallmark of medulla neuroblasts; the only exception was *Cloeon*, for which selection was based solely on *dpn* expression. As in *D. melanogaster*, medulla neuroblasts were positioned between the medulla and lamina clusters, consistent with the in *vivo* organization of the OPC (**Fig. 1F**).

In summary, we generated developmental atlases of eight insect visual systems during their major proliferative stages, annotated their constituent neuropils, and identified the medulla neuroblast populations that undergo the temporal series (**Fig. 1F**).

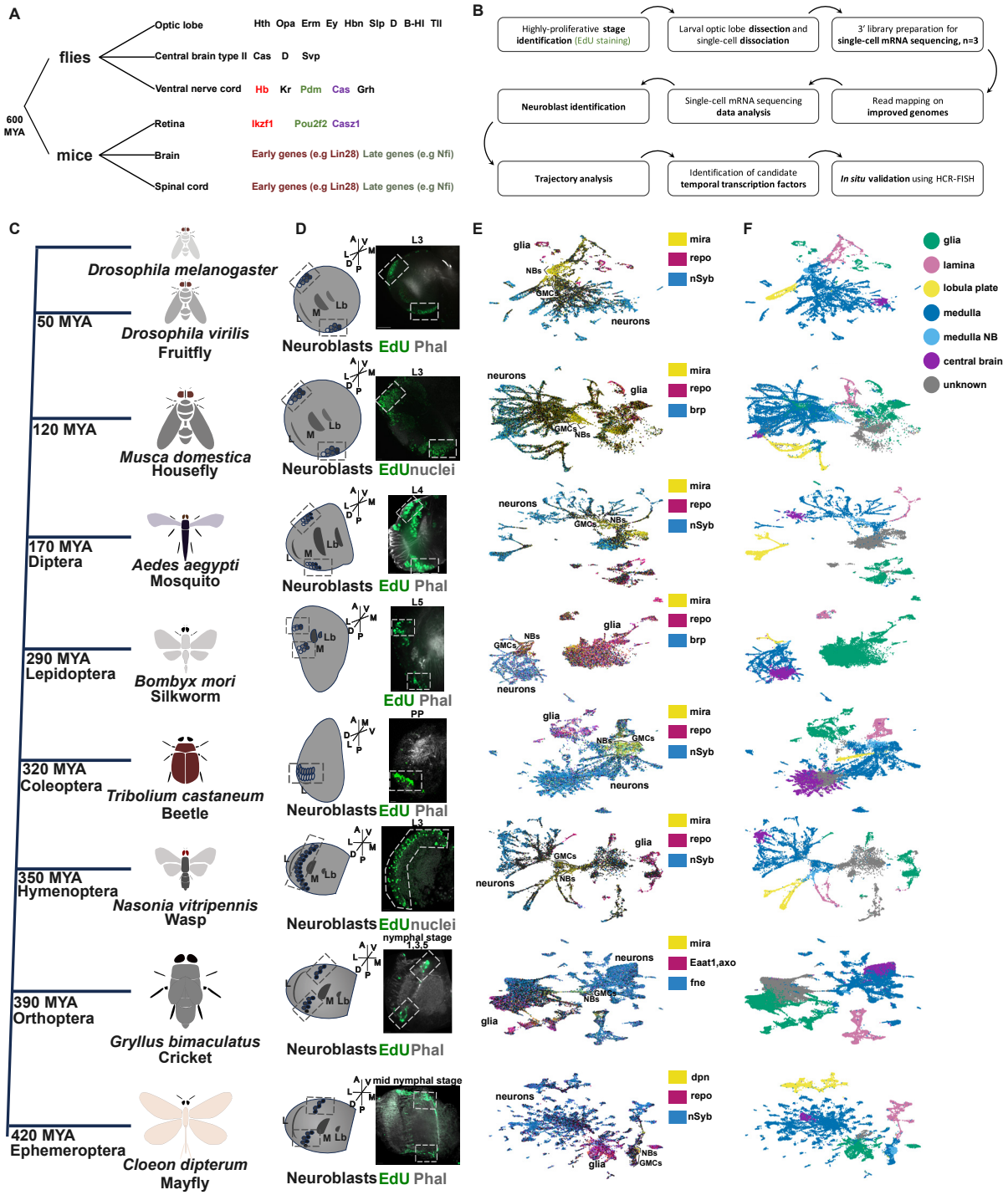


Fig. 1. Single-cell profiling of developing insect optic lobes. (A) Examples of tTF series from different neuronal structures in flies and mice. tTFs conserved between series are color-coded. (B) Overview of the experimental workflow, from the identification of highly proliferative developmental stages, through single-cell mRNA sequencing, to the computational identification of candidate tTFs and their *in situ* validation. (C) Phylogenetic relationships of the insect species analyzed in this study, with divergence times relative to *Drosophila melanogaster* indicated. (D) Schematic of the developing insect brains. Dotted boxes highlight the medulla neuroblast region. (E) EdU labeling of highly proliferative developmental stages selected for single-cell mRNA sequencing. Samples are counterstained with phalloidin to label neuropils or DAPI to label nuclei. Dotted boxes indicate proliferating medulla neuroblasts. (F) UMAP representations of single-cell mRNA sequencing data from developing optic lobes, showing expression of neuroblast markers (*mira* or *dpn*; yellow), glial markers (*repo* or *Eaat1*/*axo*; magenta), and neuronal markers (*nSyb, fne,* or *brp*; blue). (G) The same UMAP plots colored by cell type (glia, medulla neuroblasts) or brain region (lamina, medulla, lobula plate, central brain). Abbreviations: M,

medulla; Lb, lobula; L, larval stage; PP, prepupa; NBs, neuroblasts; GMCs, ganglion mother cells. Scale bar, 20 μm.

A conserved core of the temporal series

To compare the temporal transcription factor (tTF) series across insects, we first identified orthologous genes. To do so, we reconstructed phylomes(*41*) - comprehensive collections of gene-specific phylogenetic trees - for each of the eight insect genomes. This allowed us to classify one-to-one, one-to-many, many-to-one, and many-to-many orthologs, as well as species-specific genes (**Fig. S6** and **Methods**). Using these orthology relationships, we selected *Drosophila melanogaster* tTFs for different temporal windows, such as *hth* (early), *ey* or *opa* (mid), and *BarHI* or *scro* (late) and examined their expression patterns in the UMAP plots across species (**Fig. 2A**). Their expression pattern revealed that (a) neuroblasts from all insects undergo temporal patterning at the developmental stages we profiled, (b) as in *D. melanogaster*, neuroblasts exhibit a temporal organization in the UMAP space, and (c) the temporal series is likely conserved to a substantial extent across insects.

To investigate temporal regulation in these insects and identify medulla-neuroblast tTFs in an unbiased manner, we first computationally isolated the medulla neuroblast clusters and performed trajectory inference using Slingshot(*42*). We then sought to identify all transcription factors with dynamic expression along this trajectory. As a first step, we generated transcription factor lists for each insect species (**Table S1**). To do so, we included proteins that: (a) contained the terms "transcription factor" or "DNA-binding" in their annotation; (b) were assigned to known transcription factor families (e.g., containing "homeodomain" or "HLH"); (c) were one-to-one orthologs of *D. melanogaster* transcription factors; or (d) showed high sequence similarity to any *D. melanogaster* transcription factors (Methods). We deliberately used a permissive definition to prioritize sensitivity over specificity, aiming for comprehensive TF lists at the expense of potential false positives.

Next, we developed a computational pipeline to identify genes dynamically expressed along the inferred trajectory (**Fig. S7**). Using PseudotimeDE(*43*), we applied three filters to define dynamically expressed transcription factors: (a) an expression threshold ensuring the gene is expressed at some point along the trajectory; (b) a requirement that expression approaches zero at some point, ensuring temporality; and (c) a PseudotimeDE p-value to account for uncertainty in pseudotime estimation. Finally, we took into consideration their relative expression level, which tends to be high in validated temporal transcription factors, as previously demonstrated for *D. melanogaster*. Exceptionally for *Gryllus bimaculatus*, only the approximation to 0 was applied, since the levels of expression for the TFs with temporal expression were low in comparison to the other species. We then applied this standardized pipeline (**Fig. 2B**) to the medulla neuroblasts of all insects, including *D. melanogaster*, and identified tTF candidates in each species (**Fig. S8** and **Table S2**). In total, we found 14 candidate tTFs in *D. melanogaster*, 15 in *D. virilis*, 15 in *Musca*, 18 in *Aedes*, 17 in *Bombyx*, 16 in *Tribolium*, 15 in *Nasonia*, 10 in *Gryllus*, and 14 in *Cloeon*. Strikingly, despite differences in developmental strategies, eye size, and ecological reliance on vision, the number of temporal transcription factors is broadly similar across species. This suggests that the complexity of the temporal program is not simply scaled with organ size or sensory specialization, but instead reflects a conserved developmental strategy for generating neuronal diversity.

It was immediately apparent that most *D. melanogaster* tTFs were present in the majority of species examined. *D. virilis* neuroblasts expressed all *D. melanogaster* tTFs in the same order. *Musca* neuroblasts expressed 12 of 14 (lacking *ey* and *BarH2*), as did *Aedes* neuroblasts (lacking *slp2* and *BarH2*). *Bombyx* neuroblasts expressed 9 of 14 (missing *slp1/2, D*, and *BarH1/2*), *Tribolium* neuroblasts 11 of 14 (lacking *L*, *slp2*, and *BarH2*), and the *Nasonia* ones

10/14 (lacking *dmrt99B, L*, *slp2*, and *BarH2*). Gryllus expressed 7 of 14 (namely *hth*, *opa*, *ey*, *hbn*, *BarH*, *tll* and *L*), and *Cloeon* expressed 10 of 14 (lacking *slp1/2, scro,* and *tll*).

Notably, these comparisons reveal a conserved "core" of temporal transcription factors that can be traced back to the last common ancestor of almost all insects (**Fig. 2C**). In particular, 10 of the 14 *D. melanogaster* tTFs (*hth, dmrt99B, erm, opa, L, ey, hbn, scro, D,* and *BarH*) were likely expressed in the common ancestor of hemi- and holometabolous insects, and they appear to have been deployed in a similar temporal order for more than 400 million years. Furthermore, 11 of these factors were already expressed temporally in the last common ancestor of holometabolous insects, with the exception of the second paralog of the *slp* and *BarH* genes, as well as *scro*.

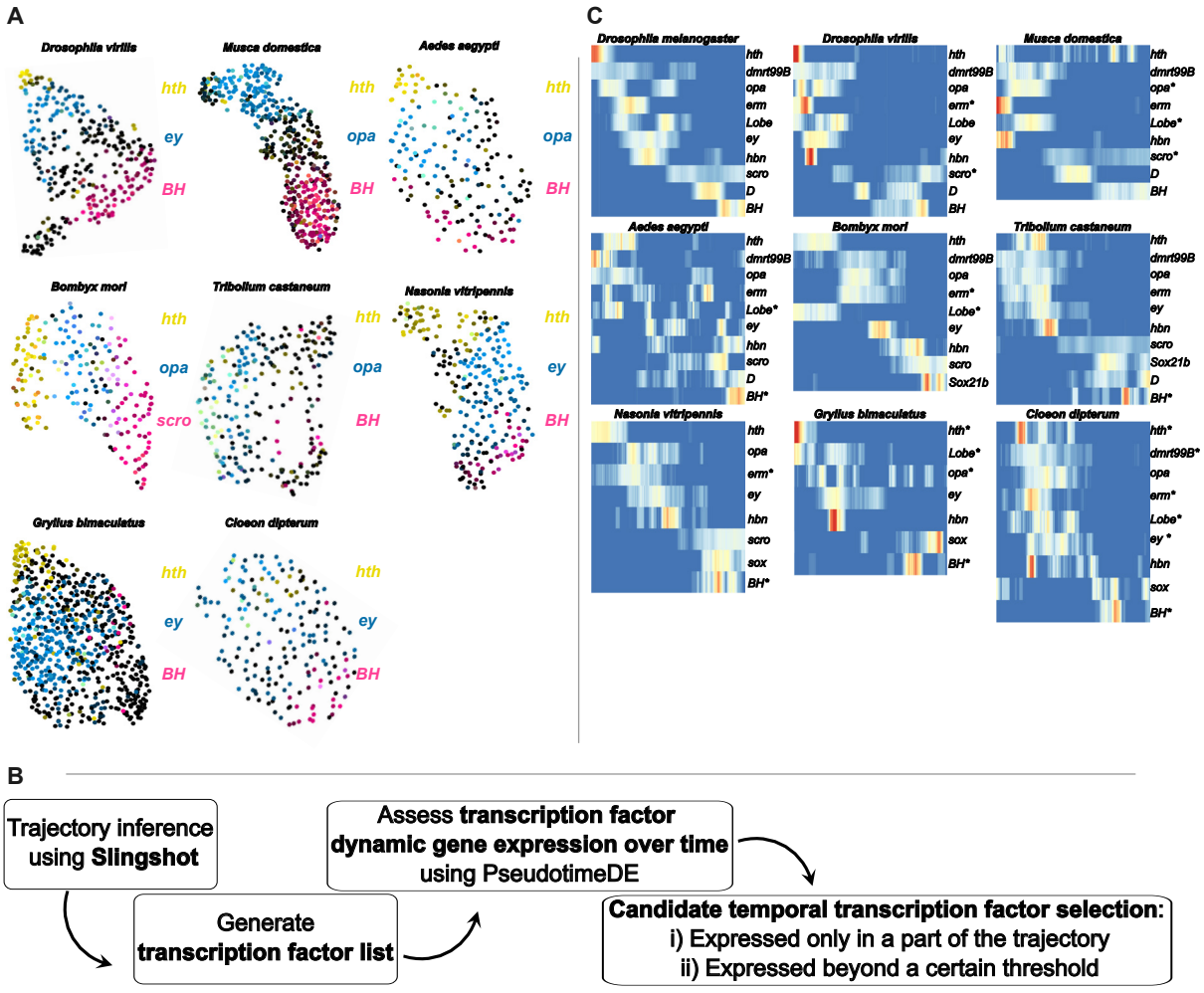


Fig. 2. Trajectory analysis of developing neuroblasts reveals a conserved core of the temporal series. (A) UMAP projections of neuroblasts from the different insect species, showing the expression of conserved temporal transcription factors corresponding to early (*hth*; yellow), mid (*ey* or *opa*; blue), and late (*BarH* or *scro*; magenta) temporal windows. (B) Schematic overview of the computational pipeline used for trajectory inference and unbiased identification of candidate temporal transcription factors. (C) Heatmap showing the expression of *Drosophila melanogaster* tTF homologs across all insects analyzed, highlighting the conserved core of the temporal series.

In vivo validation of conserved tTFs

Although our previous work had established trajectory inference as a reliable strategy for identifying temporal transcription factors, we sought direct *in vivo* confirmation of conserved and divergent candidates. To this end, we performed HCR-FISH in the developing optic lobes of all studied insects and have validated approximately 100 genes in six species, *D. virilis, M. domestica, A. aegypti, T. castaneum, N. vitripennis,* and *G. bimaculatus*. As a first step, we adapted and optimized the HCR-FISH protocol for each species, accounting for differences in tissue size, cuticle permeability, and fixation requirements (**Methods**).

To validate our computational predictions, we selected a subset of conserved tTFs - two to three per species - and examined their expression directly in medulla neuroblasts. In all cases, the HCR-FISH patterns recapitulated the temporal ordering inferred from the single-cell trajectories: early factors such as *hth*, mid-series factors such as *ey* or *hbn*, and late factors such as *BarH,* and *tll* were detected in sequential domains along the medulla neuroblast layer in *D. virilis, M. domestica, A. aegypti, T. castaneum, N. vitripennis,* and *G. bimaculatus* (**Fig. 3A-F**, respectively and **Fig. S9**). This congruence confirms that trajectory inference accurately captures the *in vivo* temporal progression of neuroblast gene expression and that the conserved core of the series is indeed deployed during optic lobe development across insects.

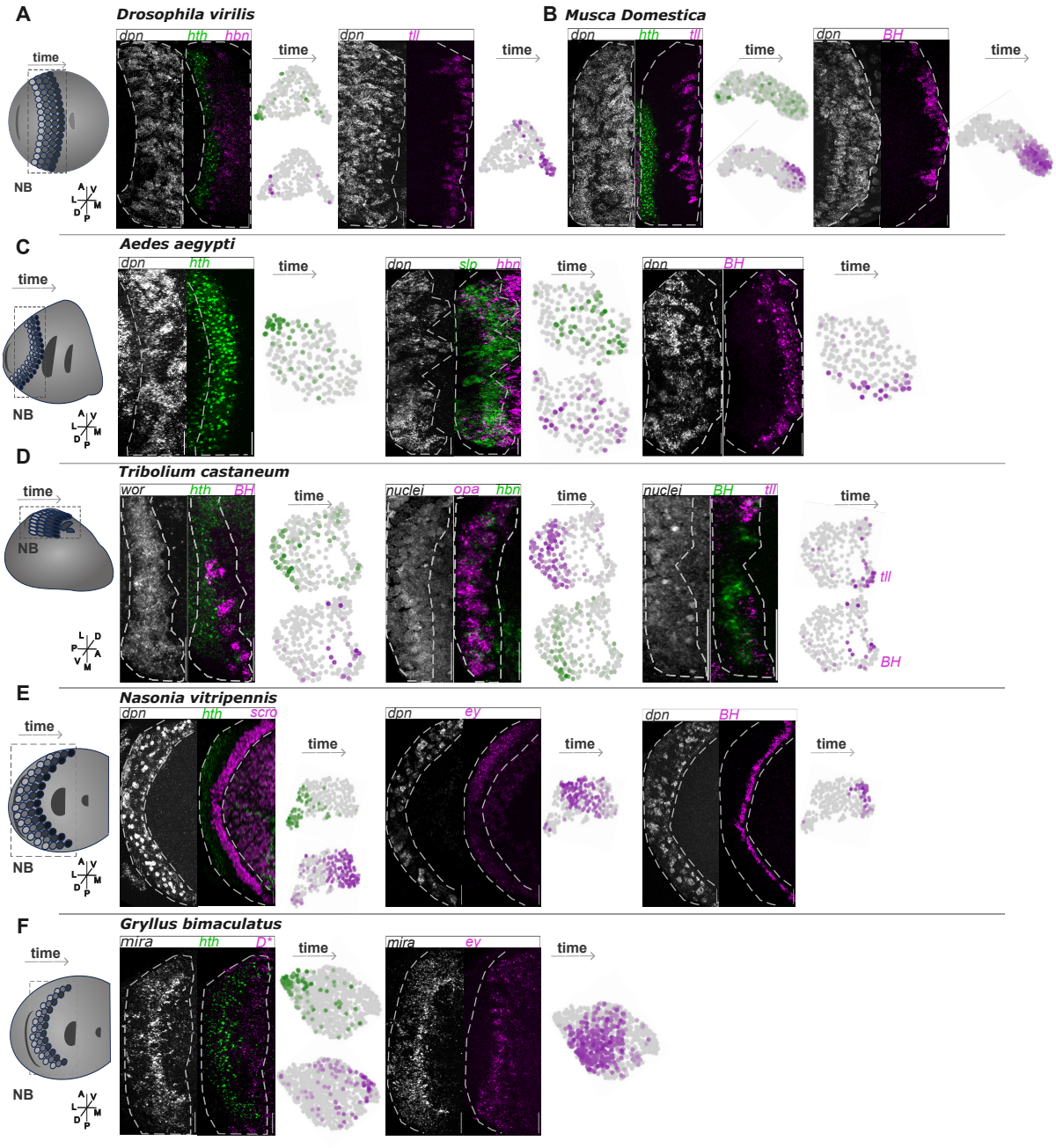
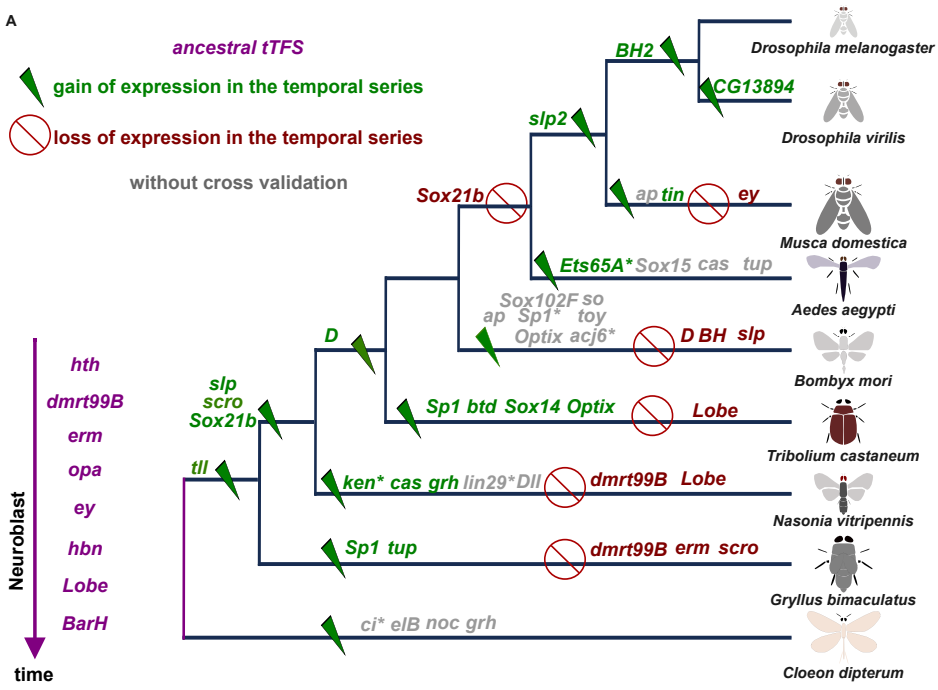


Fig. 3. Experimental validation of predicted temporal transcription factors. (A) In *D. virilis*, *dpn* expression marks neuroblasts. Consistent with trajectory inference and UMAP analysis, *hbn* is expressed in neuroblasts that are older than those expressing *hth* (left), whereas *tll* is restricted to the oldest neuroblasts (right). (B) In *Musca*, *dpn* similarly labels neuroblasts. *hth* is expressed in young neuroblasts, while *tll* and *BarH1* (left and right, respectively) are confined to the latest temporal windows, as predicted by the UMAP. (C) In *Aedes*, *hth* marks young neuroblasts (left), *slp* and *hbn* show overlapping expression in a mid-temporal window (middle), and *BarH1* is expressed in old neuroblasts (right), consistent between HCR-FISH and UMAP analyses. (D) In *Tribolium*, coordinated expression patterns are observed between HCR-FISH and UMAP plots: *hth* and *BarH1* label young and old neuroblasts, respectively (left); *opa* and *hbn* are expressed in early temporal windows (middle); and *BarH1* and *tll* mark late neuroblasts (right). (E) In *Nasonia*, *hth* is expressed in young neuroblasts and *scro* in old ones (left), *ey* marks a mid-temporal window (middle), and *BarH* is restricted to late neuroblasts (right), closely matching the computational predictions. (F) In *Gryllus*, *hth* is expressed in young neuroblasts, *D* in old neuroblasts (left), and *ey* in a mid-temporal window (right). Scale bar, 20µm.

Mechanisms of evolution of the temporal series

While a conserved core of the temporal series can be traced back to the last common ancestor of hemi- and holometabolous insects, we also identify numerous instances of temporal transcription factor recruitment and loss across lineages (**Fig. 4A**). Rather than rewiring the program, these changes predominantly occur as modifications of a shared, conserved scaffold, with most species retaining substantial portions of the ancestral series. Notably, we observe more gains than losses of tTF expression, suggesting that the addition of new temporal regulators may be more evolutionarily accessible than the elimination of existing ones. Importantly, these events do not appear to correlate with brain size or estimated neuron numbers, confirming that temporal series evolution is not simply driven by overall brain complexity. Consistent with this view, we did not identify any species-specific genes among the temporally expressed transcription factors; to confirm this, we systematically examined all species-specific genes in each insect and, indeed, found no species-specific transcription factor exhibiting temporal expression. We also asked whether certain stages of the temporal series - early, mid, or late neuroblasts - were more prone to change; we find that while tTF expression loss is almost equally distributed, gains are enriched in the late part of the temporal series, indicating that the tempo of diversification increases as the temporal series progresses with the latest windows being the most divergent (**Fig. 4B**).



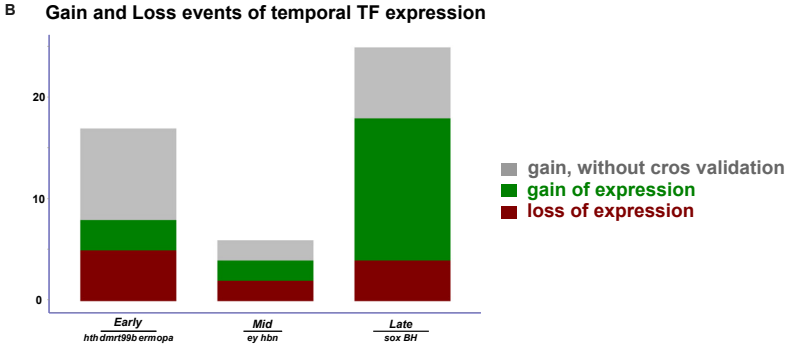


Fig. 4. The evolutionary trajectory of the temporal series. (A) Phylogenetic tree depicting the inferred evolution of the medulla neuroblast tTF series from a conserved core present in the last common ancestor of hemi- and holometabolous insects. Branches are annotated with inferred gains (green) and losses (red) of temporal expression, based on comparative single-cell and in situ analyses, as well as predicted gains lacking experimental validation (gray). Terminal branches correspond to the nine insect species analyzed in this study. (B) Barplot comparing the instances of confirmed gains (green), unvalidated gains (grey), and losses (red) of tTF expression in the different insect lineages at early-, mid-, and late-stage neuroblasts, showing an evident enrichment of diversification of the temporal series at late neuroblast stages.

Despite its apparent complexity, this phylogenetic reconstruction provides a conceptual map of temporal series evolution, which we then used to define general evolutionary mechanisms that act to diversify the series and, hence, neuronal diversity.

1. Diversification of conserved temporal windows

The first evolutionary mechanism we define is the diversification of an existing temporal window through the addition of new tTFs within an otherwise conserved temporal framework (**Fig. 5A**). Rather than altering the overall sequence of the temporal program, this mechanism expands the molecular composition of a given temporal window. We identify two distinct modes of diversification. First, a previously absent transcription factor can be recruited into an existing temporal window, as exemplified by the expression of *CG13894* within the *tll* temporal window in *D. virilis*. Second, diversification can occur through the deployment of paralogous genes within an established window, such as the expression of *slp2* alongside *slp1* in the Slp window, or *BarH2* alongside *BarH1* in the BarH window in *Drosophila* and *Musca*. Because these changes occur without disrupting the global order of the temporal series, diversification within existing windows may represent a comparatively conservative mode of temporal program evolution, allowing incremental modification of temporal identity and potentially facilitating later shifts in tTF usage.

2. Repurposing of transcription factors across developmental contexts

A second mechanism involves the co-option of transcription factors that, in *D. melanogaster*, are expressed in entirely different tissues or developmental contexts, but are deployed as temporal transcription factors in neuroblasts of other insects (**Fig. 5B**). The most striking example is *tinman (tin)*, which in *D. melanogaster* is restricted to the cardiac mesoderm(*44*), yet displays a clear and robust temporal expression pattern in medulla neuroblasts of *Musca*. Similarly, *ken*, which is not expressed in the *D. melanogaster* optic lobe and instead plays a key role in genital development(*45*), is expressed temporally in optic lobe neuroblasts of *Nasonia*. Moreover, *Optix*, a spatial transcription factor of the optic lobe neuroepithelium in *D. melanogaster*(*46*), shows temporally regulated expression in *Tribolium* neuroblasts (**Fig. 5B**). Interestingly, we identify two orthologs of the *D. melanogaster* ventral nerve cord embryonic neuroblast tTFs, *cas* and *grh,* being temporally expressed in *Nasonia* optic lobe neuroblasts (**Fig. S10A)**. These observations were unexpected, as the recruitment of transcription factors with such distinct roles implies substantial rewiring of temporal window identity. Nevertheless, these cases demonstrate that the temporal series is sufficiently flexible to allow the incorporation of factors from disparate regulatory contexts while maintaining a coherent developmental progression.

3. Positional redundancy between the same transcription family member

We also observe cases where paralogous genes from the same transcription factor family are co-expressed or substitute for one another within a temporal window (**Fig. 5C**). For example, members of the Sox family - *D*, *Sox14*, *Sox15*, and *Sox21b* - interchange their temporal expression in the late window across different species. In the hemimetabolous insects, *Gryllus* and *Cloeon*, a single Sox gene occupies this window. In *Nasonia*, *Sox21b* is added, while in

Tribolium this window is enriched by the expression of *D* and *Sox14* (**Fig. 5C**). In Diptera, *Sox21b* expression is lost, and *D* is the only Sox gene expressed, except in *Aedes* where *Sox15* is also expressed. This pattern suggests that functional redundancy within transcription factor families can facilitate smooth evolutionary transitions, allowing one paralog to replace or supplement another without disrupting temporal patterning.

4. Reuse of neuronal transcription factors as tTFs

Finally, we observed instances where transcription factors that, in *D. melanogaster*, are normally restricted to specific neuronal subtypes are repurposed as tTFs in other insects (**Fig. 5D**). For example, we confirmed the temporal expression of *Sp1* and *btd* in Tribolium neuroblasts, *Sp1* and *tup* in Gryllus (**Fig. 5D**), and *Ets65A* in *A. aegypti* (**Fig. S10B**). Notably, several of these genes - including *elB*, *noc*, *Ets65A*, *tup, Dll and Sox21b* - have been previously identified as candidate terminal selectors in *D. melanogaster* (*23*, *47*). This pattern suggests a relatively simple evolutionary route for generating new tTFs: shifting the regulatory program of a gene from the GMC or post-mitotic neuron back into the neuroblast lineage, or vice versa, allows it to acquire a temporal role without extensive rewiring of the existing temporal cascade.

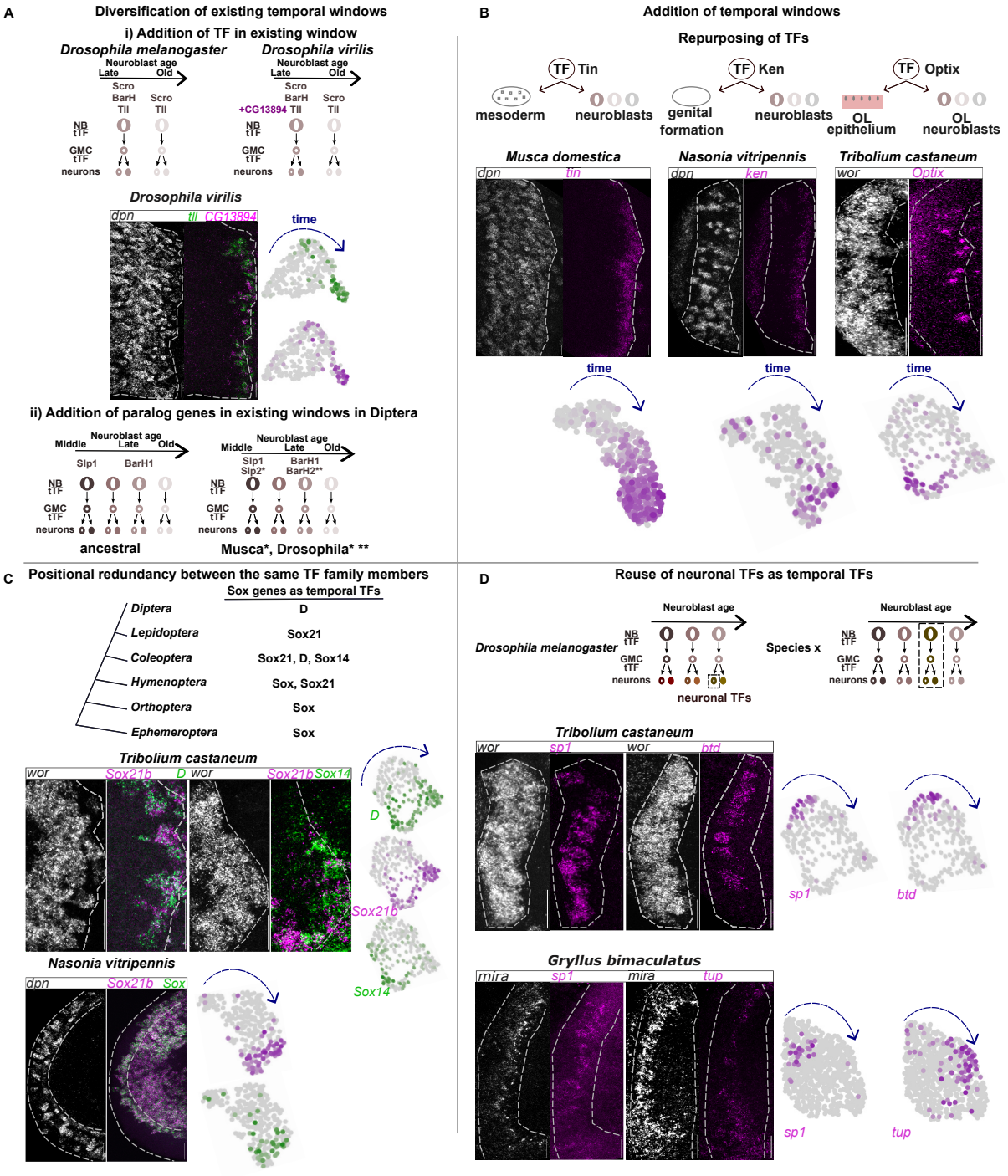


Fig. 5. Mechanisms underlying the evolution of temporal patterning in insect neuroblasts. (A) Diversification of existing temporal windows. (i) Recruitment of a new transcription factor into a preexisting window, exemplified by *CG13894* (magenta) within the *tll* (green) window in *D. virilis* (HCR-FISH and UMAP). (ii) Addition of paralogous genes within established windows, such as *slp2* in the *Slp* window in *Musca* and *BarH2* in the *BarH* window in *D. melanogaster.* (B) Repurposing of transcription factors from other developmental contexts. Temporal expression of *tin* in old neuroblasts of *Musca* (cardiac mesoderm in *D. melanogaster*), *ken* in *Nasonia* (genital development in *D. melanogaster*), and *Optix* in *Tribolium* (spatial factor in *D. melanogaster* neuroepithelium). (C) Positional redundancy among paralogous transcription factors. Overlapping temporal expression of Sox family members across lineages: *Sox21* and *D*, *Sox21b* and *Sox14* in Tribolium, and *Sox21b* and *Sox* in Nasonia. (D) Reuse of neuronal transcription factors as temporal factors. TFs normally restricted to post-mitotic neurons in *D. melanogaster* are deployed as tTFs in other insects: *Sp1* and *btd* in young *Tribolium* neuroblasts, *Sp1* and *tup* in young/mid-age *Gryllus* neuroblasts.

How does temporal patterning affect neuronal diversity?

We next asked how evolutionary differences in neuroblast temporal patterning may influence neuronal diversity. We note at the outset that the datasets presented here are not sufficient to definitively assign mature neuronal identities, as the neurons we profile are not fully specified at these developmental stages. In previous work in *D. melanogaster*, such questions required integration with later, pupal-stage datasets to annotate neuronal cell types (*23*). Nevertheless, several neuronal transcription factors begin to be expressed shortly after neuron birth(*8*), providing an opportunity to examine early signatures of neuronal diversification.

We focused this analysis on *Musca domestica*, which displays some of the most pronounced deviations from the *D. melanogaster* temporal series, including the co-option of a mesodermal transcription factor (*tin*), the absence of a canonical *D. melanogaster* tTF (*ey*), and the deployment of a paralog (*BarH2*) within an existing temporal window. Together, these differences provide an informative case study to explore potential downstream consequences of temporal series evolution (**Fig. 6**).

We first examined the effects of *tin* co-option (**Fig. 6A**). In *Musca*, *tin* is robustly expressed in neuroblasts and ganglion mother cells, whereas it is entirely absent from the optic lobe lineage in *D. melanogaster*, as expected. Using *tin* expression at the root of the inferred trajectories, we identified neurons likely originating from the *tin* temporal window. These GMCs and young neurons express *abrupt (ab)*, a transcription factor that in *D. melanogaster* is restricted to central brain neurons (and is similarly expressed in *Musca* central brain cells, marked by an asterisk). Strikingly, these *tin*-window neurons also express *Pox meso (Poxm)*, which is not expressed in *D. melanogaster* optic lobe neuroblasts or neurons and is instead involved in larval body wall muscle development(*48*). Together, these observations suggest that the *tin* temporal window in *Musca* generates neurons with transcriptional identities not present in *D. melanogaster*, consistent with the emergence of species-specific neuronal populations.

We next examined the consequences of the absence of *ey* as a temporal factor in *Musca* (**Fig. 6B and Fig. S11**). In *D. melanogaster*, Lim3^{+} neurons represent the NotchOFF progeny of the earliest neuroblast divisions and arise from several successive temporal windows (**Fig. 6B**): the Hth window (co-expressing *TfAP2* and *svp*; **Fig. S11**), the Hth/Opa window (co-expressing *run* and *tup*; **Fig. S11**), the Erm/Ey window (co-expressing *kn* and *tup*; **Fig. S11**), and the Ey/Hbn window (co-expressing *toy*; **Fig. 6B**)(*8*). In *Musca*, we identified all corresponding Lim3^{+} neuronal populations except those expressing the *ey/toy* homolog (*Pax6*). To determine whether this reflected the loss of a specific marker rather than the loss of an entire neuronal population, we searched for unaccounted Lim3^{+} neurons and found that all Lim3^{+} neurons derived from the Erm/Hbn temporal window in *Musca* are *tup*$^{+}$, whereas the equivalent trajectories in *Drosophila* contain both *tup*$^{+}$ and *tup*$^{-}$*toy*$^{+}$ neurons. We also looked for the NotchON neurons that come from the Ey/Hbn temporal window, which in *Drosophila* express *vvl*. While *vvl* is expressed in some Musca neuronal trajectories, these are markedly fewer than in *D. melanogaster* and they all express *erm* too (**Fig. S11**), suggesting that these neurons come from the Erm temporal window rather than the Ey one. Together, these results provide strong evidence for the absence of an Ey temporal window in Musca, accompanied by a corresponding loss of its neuronal progeny.

Finally, we examined the potential effects of the absence of *BarH2* expression in *Musca* (**Fig. 6C**). In *D. melanogaster*, neurons born during the *BarH* window express transcription factors such as *Ets65A*, *Fd59A*, and/or *Dll*(*8*). In Musca, we identified neurons expressing *Fd59A* and *Dll*, consistent with their origin from the BarH1 temporal window. Notably, however, *Fd59A* expression levels were markedly reduced, and *Ets65A* expression was entirely absent. This suggests that the addition of *BarH2* expression in *D. melanogaster* may modulate downstream transcriptional programs, either by amplifying the expression of pre-existing neuronal

transcription factors or by reshaping transcription factor combinations, thereby contributing to altered neuronal identity.

Although these observations are not conclusive in the absence of adult-stage neuronal data, they provide compelling evidence that evolutionary changes in the temporal transcription factor series can have tangible effects on the transcriptional identities of newly born neurons. Together, they support the idea that diversification of neuroblast temporal patterning is a plausible mechanism for generating neuronal diversity across insect species.

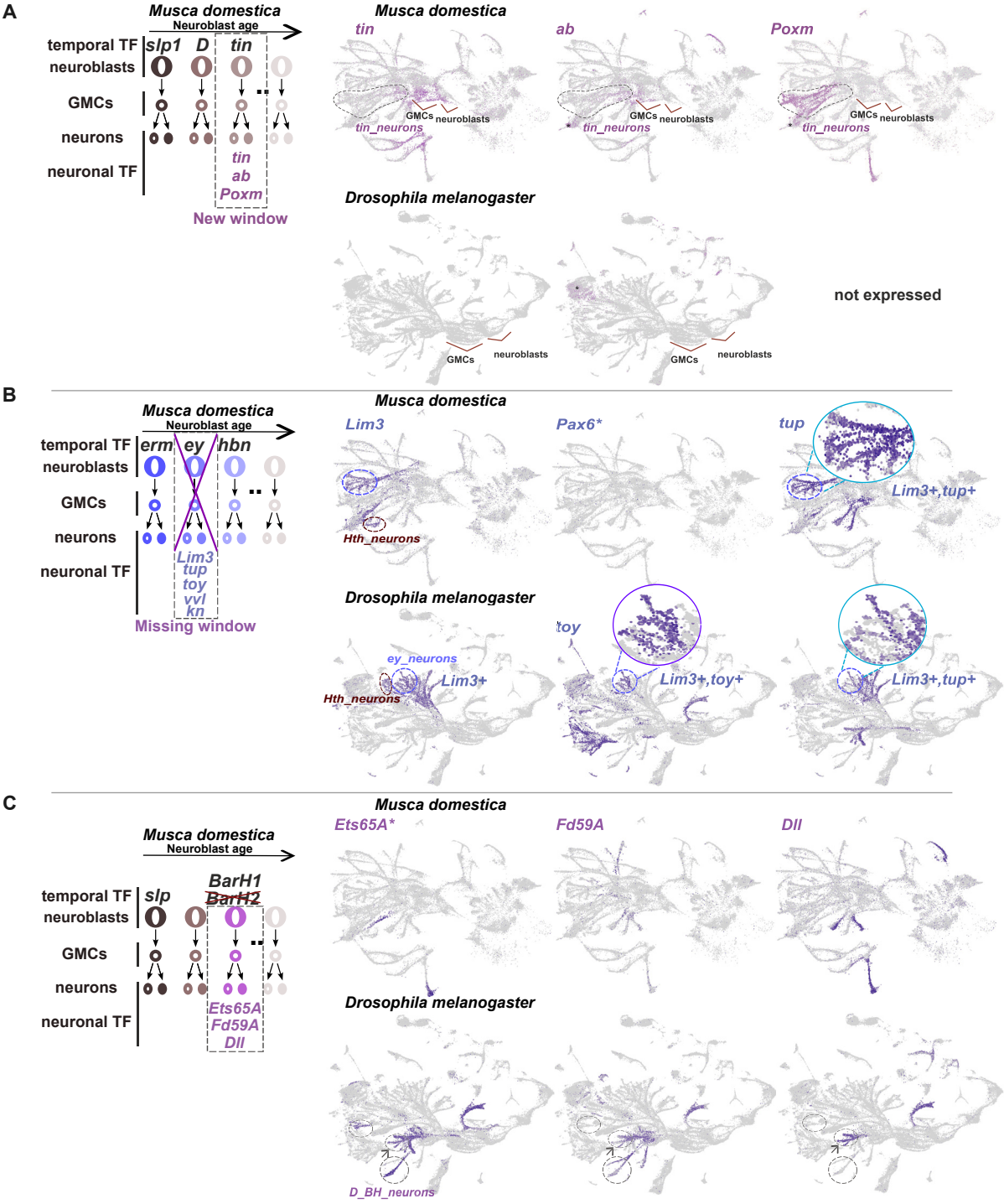


Fig. 6: Impact of tTF variation on early neuronal identity. (A) GMCs and neurons from the *tin* temporal window in Musca express *tin*, *ab*, and *Poxm*. UMAP plots showing the expression of *tin*, *ab*, and *Poxm* in the Musca and Drosophila developing brains. In *Musca, tin* is expressed in neuroblasts and GMCs from the same temporal window. The GMCs mark the beginning of the trajectories of neuronal progeny that later express *Poxm*. In *Drosophila, tin* and *Poxm* are not expressed at all, while *ab* is only expressed in few central brain neurons. (B) The absence of the *ey* temporal window leads to the absence of its neuronal progeny in *Musca*. In *Drosophila*, Lim3+ neurons come either from the Hth temporal widnow or from the Erm/Ey/Hbn windows. The neurons coming from the Ey/Hbn temporal window express *Lim3* and *toy* and are absent from Musca, while the ones coming from the Erm/Ey temporal widnow expressing *Lim3* and *tup* can be found in both *Musca* and *Drosophila*. (C) The absence of BarH2 expression in *Musca* correlates with asbence of *Ets65A* expression and marked reduction in *Fd59A* expression levels.

Discussion

In this study, we provide the first systematic analysis of how a neural stem cell temporal patterning program evolves across insects. By combining single-cell transcriptomics, trajectory inference, and in *vivo* validation, we show that medulla neuroblasts across deeply diverged insect lineages share a conserved temporal transcription factor core that was already present in the last common ancestor of hemi- and holometabolous insects. At the same time, we uncover extensive, yet structured, variation in temporal series composition across species. Rather than rewiring, temporal programs diversify through recurrent modifications of a shared scaffold: our comparative analysis identified four evolutionary mechanisms that may broadly shape the evolution of temporal patterning programs across species: diversification of existing temporal windows by adding new factors, positional substitution among paralogous transcription factors, co-option of transcription factors from other developmental contexts, and reuse of neuronal regulators as temporal transcription factors. Finally, neuronal transcription factor analyses suggest that such evolutionary changes in temporal patterning can translate into altered neuronal transcriptional identities. Together, our results reveal both the deep conservation and the evolutionary plasticity of temporal patterning programs, and establish temporal transcription factor series as a tractable substrate for the evolution of neuronal diversity.

While our study provides a broad view of temporal program evolution, some limitations remain. First, temporal transcription factors are often expressed at low levels and over narrow developmental windows, raising the possibility that additional regulators may have been missed; emerging single-cell technologies with increased sensitivity, as well as methods that better capture transcriptional dynamics, will be well suited to refine and extend the temporal series identified here. Second, our analysis necessarily depends on the quality of available genome assemblies and annotations, which vary across insect lineages. While we mitigated this by improving gene models using single-cell mRNA sequencing data, we also provide detailed pipelines that can readily be reapplied as genomic resources improve. These considerations highlight both the current scope and the future potential of comparative temporal analyses to uncover the principles guiding neuronal diversification across species.

While our initial expectation based on the existing data was that temporal series evolve quickly, our analyses instead revealed a highly conserved landscape. These conserved tTFs may represent a developmental constraint crucial for the precise progression and timing of neuroblast differentiation, with deviations from this program potentially leading to defects in neuronal identity, connectivity, or brain architecture. Alternatively, they may represent a core toolkit for neuroblast patterning, providing a modular framework that reliably generates temporal progression across species. Rather than being preserved solely due to constraint, these tTFs may constitute an evolutionary solution for specifying neuronal identity, offering a robust framework upon which neuronal diversity can be built, while remaining evolvable and permitting lineage-specific modifications.

Our findings further suggest that temporal series evolve not through the emergence of entirely new genes, but via recombination and repurposing of existing regulatory components within this conserved toolkit. This may explain the existence of a complex gene regulatory network driving the progression of the series: through redundancy and modularity, the temporal series can generate neuronal diversity without disrupting the overall developmental trajectory. Such built-in redundancy provides evolutionary flexibility, allowing species to adapt and diversify while maintaining the core temporal program. In this way, the temporal series gene regulatory network acts as a buffer, permitting changes to accumulate without destabilizing its fundamental structure, thereby supporting both stability and adaptability in nervous system evolution.

Our study indicates that insect optic lobe temporal series are highly conserved over hundreds of millions of years. While comparable data are not yet available for other insect neural stem cell populations, such as the ventral nerve cord or central brain, it is reasonable to hypothesize that similar constraints and evolutionary dynamics apply. In vertebrates, temporal transcription factor cascades in both the spinal cord and developing brain are likewise conserved across species suggesting a general principle. However the temporal series differ massively between vertebrates and invertebrates and between different neuronal tissues. One possibility is that temporal series may have evolved rapidly after the divergence of vertebrates, followed by lineage-specific stabilization. Within insects, tissue-specific adaptation and modular reuse of regulatory modules could explain differences between optic lobe, ventral nerve cord, and central brain, allowing distinct neuronal repertoires while preserving the core temporal patterning principle. Alternatively, temporal patterning may not arise *de novo* in each neural stem cell population but could be adapted from preexisting regulatory modules, with temporal and spatial programs interconvertible to some extent. Spatial cues could be co-opted to generate temporal progression, and vice versa, allowing conserved regulatory logic to be redeployed in novel contexts; the expression of some optic lobe tTFs along the anteroposterior axis of the developing *Drosophila* procephalic ectoderm supports this scenario(*49*). **Such flexibility could reconcile the deep conservation of temporal programs with the remarkable diversity of nervous system architectures across animals.**

Methods

Insect strains

Drosophila melanogaster: Canton-S; animals were reared at 25°C.
Drosophila virilis: 15010-1051.86; animals were reared at 25°C.
Musca domestica: M3 strain(*50*) (derivative of the wild-type multimarked strain aabys); animals were reared at 25°C.
Aedes aegypti: the inbred Paea strain(*51*); animals were received the days of experiments.
Bombyx mori: p50 strain (Silkworm Genetic Resource Database); animals were reared at 25°C.
Tribolium castaneum: wild type San Bernardino strain; animals were reared at 30°C.
Nasonia vitripennis: AsymCx wild-type strain; animals were reared at RT and 30°C .
Gryllus bimaculatus: wild-type strain with white-eyes(*52*); animals were reared at 28°C.
Cloeon dipterum: inbred wild type strain; animals were reared at 25°C.

Edu staining and selection of neurodevelopmental stages

EdU proliferation assays were performed to identify developmental stages in different insect species that most closely correspond to the late third larval stage (L3) of *Drosophila melanogaster*, when neuroblasts are highly proliferative and medulla neurons are actively generated. For each species, stages immediately preceding pupation and metamorphosis were targeted, except for *Gryllus bimaculatus* and *Cloeon dipterum*, where neuroblasts and differentiated neurons coexist in nymphs immediately after hatching. Stage selection was based on available literature descriptions of optic lobe morphology, including the presence of Outer Proliferation Centers and expansion of the medulla neuropil as indicators of ongoing neurogenesis. Accordingly, the following stages were analysed: wandering stage larvae- *D. virilis* (Late L3), *M. domestica* (Late L3), *A. aegypti* (L4), *B.* mori (L5), *N. vitripennis* (Late L3, white larvae, after gut cleaning, still moving), for *T. castaneum* (L5) and early prepupa(PP), since the neuropil is significantly expanded at this stage for beetles, as well as all the nymphal stages for *G. bimaculatus*(*53*) and early to mid-late nymph stages for *C. dipterum* (2, 3 and 4 weeks after hatching, respectively). Brains were dissected in appropriate insect media (Shneiders insect medium (Sigma Aldrich), except *B.mori* where Grace's insect medium (GIBCO), was used) and incubated *ex vivo*(*54*) with EdU at a concentration of 20uM using a 4-hour pulse, followed by a 4-hour chase to allow detection of active proliferation. After incubation, samples were chemically fixed (3,7% FA/PBS, *D.virilis, M. domestica* 20 min RT, *A. aegypti, B.mori, G, bimaculatus, C. dipterum* ON at 4°C, *T. castaneum* 1h15min RT, *N. vitripennis* 2h15min RT), rinsed 3x PBST 0.3% (1xPBS with Triton-X-100 0.3%) and washed 2x PBST 0.3% for 15 min each, processed using the Click-iT EdU Imaging Kit. Brains were rinsed with PBST 0.3% and were incubated for 2 days with primary antibodies at 4 °C. Phalloidin (Sigma Aldrich) and Acetylated-Tubulin were used at a concentration of 1/100 in PBST 0.03%. Brains stained with Phalloidin were washed with PBST 0.3% and mounted in Vectashield. For acetytaled tubulin staining, brains where incubated with secondary antibodies for 2 days at 4 °C. The secondary antibodies were washed, and the brains were mounted in Vectashield for observation. The ZEISS LSM 980 laser-scanning confocal microscope or CSU-X1 Zeiss Spinning was used to verify active proliferation in the medulla neuropil.

Single-cell dissociation protocols for the different insects and library preparation

Single-cell dissociation

Animals of all species were synchronized to the wandering stages mentioned above, except for *T. castaneum* that early prepupa was selected, *G.bimaculatus* stages 1,3,5 and *C.dipterum* mid-nymphal stage. The day of the single-cell dissociation and library preparation brains were dissected at the selected stage, but we covered a range of mid-late wandering animals to early prepupa for all species, in order to capture the maximum available information, except *G.bimaculatus and C.dipterum* that nymphs were similarly selected. All dissections were

carried out in ice-cold insect medium (Schneider's insect medium for all, except *B.mori* where Grace's was used), on pre-chilled Sylgard silicone plates using thoroughly cleaned with Ethanol 70% and double distilled water Dumont #5.5 forceps. Intact brains were gently transferred into glass staining dishes containing cold medium and kept on ice, taking precautions to prevent tissue damage and to avoid brains adhering to pipette tips, BSA-coated tips were used. Dissections did not exist 1h30min. Once dissections were completed optic lobes were separated from the rest of the brain using Vannas Spring Scissors with a 2-mm cutting edge (Fine Science Tools), or well sharpened forceps when possible.

To generate single-cell suspensions for scRNA-seq, optic lobes were pooled (numbers per preparation indicated below) and enzymatically digested with collagenase (GIBCO; 2 mg/mL in the respective inset media) alone or in combination with dispase (GIBCO; 2 mg/mL in the respective insect media); for selected *Gryllus bimaculatus* stages, collagenase was supplemented with TrypLE (Sigma Aldrich). Digestions were performed at 25–30 °C for species-specific durations, followed by mechanical dissociation to obtain single cells. After digestion, only plasticware and pipette tips compatible with 10x Genomics recommendations were used. Optic lobes were transferred into Eppendorf DNA LoBind tubes (Fisher Scientific) containing 150 µL DPBS + 0.04% BSA, and dissociated by vigorous pipetting (30-50 strokes per round), interspersed with 1 min rests on ice, until large tissue chunks were no longer visible; suspensions were inspected under a dissecting microscope and additional pipetting was performed only when remaining clumps were large. Small clumps were not forced into single cells, as these were typically neuropil fragments rather than somata, and reducing pipetting can improved data quality(*23*). The suspension was then filtered through a 20 µm strainer (pluriSelect), and the filter was washed with an additional 50 µL DPBS + 0.04% BSA, which was collected together with the flow-through (including by pipetting from the opposite side when needed). To minimize cell loss due to adhesion, the collection tubes and the 20 µm strainer were pre-coated with DPBS + 0.04% BSA for at least 20 min on ice. Cell suspensions were evaluated with a Luna FL cell counter (Logos Biosystems) for the ratio of single cells to aggregates to optimize enzymatic digestion time, and subsequently Luna FL used to estimate single-cell concentration , except for *D. virilis*, for which cell concentration was measured using a 0.02 mm deep cytometer. Conditions were as follows (number of pooled optic lobes (OLs); enzyme(s); temperature; duration): *D. virilis* (11 OLs; collagenase and dispase; 25 °C; 15 min), *M. domestica* (6 OLs; collagenase; 28 °C; 10 min), *A. aegypti* (10 OLs; collagenase and dispase; 28 °C; 20 min), *B. mori* (8 OLs; collagenase and dispase; 30 °C; 40 min), *N. vitripennis* (12 OLs; collagenase and dispase; 30 °C; 10 min; surrounding sheath removed prior to dissociation), *T. castaneum* (60 OLs; collagenase and dispase; 30 °C; 15 min), *C. dipterum* (9 OLs; collagenase and dispase; 28 °C; 70 min), and *G. bimaculatus* nymphal stages st1 (21 OLs; collagenase; 28 °C; 10 min), st3 (10 OLs; collagenase and TrypLE; 28 °C; 30 min), and st5 (5 OLs; collagenase and TrypLE; 28 °C; 35 min). Digestion was performed on optic lobes for all species, except *N. vitripennis*, where whole brains were digested and optic lobes were isolated afterwards for downstream processing.

Library preparation

Three independent single-cell mRNA-seq libraries were generated for each developmental stage from a single pooled single-cell suspension prepared by merging optic lobes from multiple individuals of the same species (i.e., the same suspension was split to generate the three libraries). In total, 30 libraries were produced (3 libraries per species, except *G. bimaculatus* where 9 libraries were prepared for three nymphal stages). Library preparation followed the 10x Genomics Chromium Next GEM Single Cell 3′ Reagent Kits v3.1 (Dual Index) protocol, which captures the 3′ end of polyadenylated mRNA transcripts. Libraries were

sequenced by Novogene on an Illumina NovaSeq platform using PE150 (paired-end 2×150 bp), targeting ~30,000 reads per cell.

GeneExt and CellRanger

GeneExt(*28*) was used to mitigate incomplete genome annotations in non-model species, where missing 3′ gene ends can reduce read/UMI assignment in 3′-biased single-cell RNA-seq datasets. GeneExt performs peak calling on aligned reads to detect transcriptional signal extending beyond annotated gene boundaries and extends nearby genes' 3′ regions accordingly, thereby improving UMI capture; it can also use unassigned intergenic peaks to propose putative novel genes. GeneExt was run on the initial BAM files generated with cellranger v7.1.0(*55*) (--expect-cells=15000) to produce an improved genome annotation file, and cellranger was then re-run for each library using the updated annotation to generate refined BAM files and gene expression matrices. To maximize evidence for extension calling, BAM files from all sequenced developmental and adult stages were merged per species using samtools v1.13(*56*) and processed with GeneExt using a subsample parameter (–subsamplebam) of 100 million reads to accelerate computation as well as the orphan parameter to report candidate novel genes; orphan predictions were subsequently filtered out and retained only for future exploratory analyses. The final, filtered improved annotation (integrating information across stages) was used for the second cellranger run to generate updated BAM files and final gene expression matrices with enhanced transcriptional coverage. Using this approach, GeneExt extended 3′ regions for 4260/15621 genes in *D. virilis* (median extension 562 bp), 4893/20747 in *M. domestica* (937 bp), 3994/19269 in *A. aegypti* (1427 bp), 5991/16687 in *B. mori* (3115 bp), 3094/14492 in *T. castaneum* (804 bp), 3876/15025 in *N. vitripennis* (1106 bp), 5629/17869 in *G. bimaculatus* (5423 bp), and 5403/16357 in *C. dipterum* (705 bp).

Single-cell mRNA sequencing data analysis

The libraries were analyzed using Seurat v5(*57*). For each dataset, a Seurat object was created from the Cell Ranger gene–barcode count matrix (UMI expression matrix), using all droplets classified as cells by Cell Ranger's cell-calling filters. Genes expressed in fewer than three cells were removed, and cells with fewer than 200 detected genes were excluded. Objects were further filtered by applying library-specific thresholds of genes per cell (nFeatures/cell) to optimize data quality. Filtering decisions were guided by inspection of nFeatures distributions visualized with histograms and violin plots (**Fig. S3**) for each library, and some low-complexity droplets were intentionally retained to maximize available information. DoubletFinder(*58*) was used to remove doublets in libraries with more than 20000 cells after filtering (*B. mori* and *M.domestica*). Standard Seurat preprocessing functions—including NormalizeData, FindVariableFeatures, and ScaleData—were applied with default parameters, followed by dimensionality reduction using PCA and UMAP (100–150 principal components depending on the dataset). The three libraries generated from the same pooled suspension for each species were integrated using the FindIntegrationAnchors and IntegrateData functions. After integration, residual groups of low-quality cells that showed poor integration and were characterized by reduced nFeatures compared to the main cell population were removed only for visualization purposes.

Phylome reconstructions

A phylome is the collection of phylogenetic trees for each gene in a genome. We reconstructed a phylome for each of the species included in the analysis and added *Daphnia pulex* to serve as outgroup. Proteomes were retrieved from public repositories and formatted using phylomeDB-compliant proteome codes as follows (species; proteome code; source; phylomeID): *Drosophila melanogaster* (DROME; FlyBase v6.49; 938), *Drosophila virilis*

(DROVI; NCBI assembly GCF_003285875.2; 939), *Musca domestica* (MUSDO; VectorBase-61; 940), *Aedes aegypti* (AEDAE; VectorBase-61; 941), *Bombyx mori* (BOMMO; Ensembl Metazoa release 55; 942), *Tribolium castaneum* (TRICA; Ensembl Metazoa release 55; 943), *Nasonia vitripennis* (NASVI; Ensembl Metazoa release 56; 944), *Gryllus bimaculatus* (GRYBI; *G. bimaculatus* genome resource; 945), *Cloeon dipterum* (proteome code 197152; NCBI assembly GCA_902829235.1; 946), and *Daphnia pulex* (DAPPPU; retrieved from phylomeDB; 947).

We used an automated pipeline that uses the same process to reconstruct gene trees as one would do manually(*59*). First the proteome database was reconstructed using the 10 species and formatting the codes to phylomeDB format. Then a phylome was reconstructed starting from each one of the species. For each species a blastp search was performed between each gene in their genome and the proteome database. Blast results were filtered using an e-value threshold of 1e-05 and an overlap threshold of 50%. The number of hits was limited to the 200 best hits for each protein. Then six different multiple sequence alignments were reconstructed using three programs (Muscle v3.8.1551(*60*), mafft v7.407(*61*) and kalign v2.04(*62*)) and aligning the sequences in forward and in reverse. From this group of alignments a consensus alignment was obtained using M-coffee from the T-coffee package v12.0(*63*). Alignments were then trimmed using trimAl v1.4.rev15 (consistency-score cut-off 0.1667, gap-score cut-off 0.9)(*64*). IQTREE v1.6.9(*65*) was then used to reconstruct a maximum likelihood phylogenetic tree. Model selection was limited to 5 models (DCmut, JTTDCMut, LG, WAG, VT) with freerate categories set to vary between 4 and 10. The best model according to the BIC criterion was used. 1000 rapid bootstraps were calculated. All trees and alignments were stored in phylomedb(*41*) with phylomeIDs ranging from 938 to 947 (http://phylomedb.org).

A species tree was reconstructed using a gene concatenation approach. To do that we first selected which genes in the phylome of *D. melanogaster* were found in single copy in all species. 547 such genes were selected. The alignments for those genes were then concatenated into a single multiple sequence alignment producing an alignment of 342,569 amino acid positions. IQTREE v1.6.9(*65*) was used to reconstruct the species tree using LG as the model. 1000 rapid bootstrap calculations were performed.

All trees across the phylomes were rooted using the species tree reconstructed above. To do so, for each tree the seed species was taken and the other species were numbered from closest to farthest according to the species tree. Then the root was placed at a node containing the farthest related species in the gene tree. Once the trees were rooted, all possible orthology and paralogy relations were extracted from each tree using the species overlap algorithm as implemented in ETE3(*66*). For each pair of sequences detected as orthologous in any of the trees we computed a score which divides the number of times the pair appeared as orthologs across all trees by the number of times the pair of sequences was found as either orthologs or paralogs (CS). If the CS was above 0.5 the pair of sequences were considered orthologous.

Identification of transcription factors in insect genomes

A transcription factor (TF) candidate list was generated by screening the genome annotation file of each species using a curated reference list of *D. melanogaster* TF names and TF-associated keywords (“homeobox”, “zinc finger”, “bHLH/HLH”, “leucine zipper”, “POU”, “LIM”, “forkhead”, “nuclear receptor”, “DNA-binding”, “transcription”) as functional cues. Matches were performed in a case-insensitive manner while accounting for minor formatting differences such as hyphenation, and genes with at least one match were retained as putative TF candidates, with the corresponding TF term recorded. Additionally, we queried the similarity-relationships list across the phylomes to identify proteins showing sequence similarity to any known *D. melanogaster* TF, further expanding the candidate lists. This inclusive strategy allowed recovery of maximal TF-related information, accepting the

possibility of false positives in order to capture as many potential candidate TFs as possible. In the resulting lists, *D. melanogaster* TF names were propagated to the corresponding orthologs in each species when orthology relationships were established.

Identification of temporally expressed transcription factors

Initially, we identified medulla neuroblast clusters in the integrated object based on expression of neuroblast markers such as *mira*, *dpn*, and *wor*, while lamina and lobula plate neuroblast pools were excluded using complementary markers. For each species, clusters of interest were extracted and developmental trajectories were inferred using Slingshot(*42*), which defines lineages from low-dimensional embeddings and orders cells in pseudotime. For each species, candidate temporal transcription factors (tTFs) were then investigated by testing for genes with pseudotime-associated expression changes using PseudotimeDE(*43*), focusing on the species-specific TF candidate lists. Because our goal was to detect temporal restriction (genes expressed in only part of the trajectory) rather than broadly dynamic expression across the entire trajectory, we implemented additional filtering steps to prioritize genes with windowed expression. Pseudotime values were rescaled to 0–1 and cells were binned into 20 equally spaced pseudotime intervals to generate a cell-by-interval assignment matrix. Candidate tTFs were then filtered using log-transformed counts with two criteria: (i) a reach-zero filter retaining genes whose mean expression dropped to near-background levels in at least one pseudotime interval (mean < 0.1), while requiring expression in at least 5% of all cells; and (ii) a minimum-expression filter requiring sufficient overall expression (mean non-zero log-expression above a defined threshold; values of 0.4 and 0.45 were tested) and at least 5% non-zero cells to remove lowly expressed/noisy genes. Genes passing both filters were retained and further inspected. Candidates with clearly temporally restricted patterns (i.e., not expressed throughout the trajectory) were prioritized as putative tTFs. In addition, we considered that bona fide tTFs are typically robustly expressed in neuroblasts relative to false-positive candidates , and the parametric p-value from pseudotimeDE was used as an additional selection criterion, printed at each plot of candidate tTF.

Hybridization Chain Reaction Fluorescence *In Situ* Hybridization (HCR-FISH)

To generate the final list of candidate temporal transcription factors, we validated expression patterns by HCR-FISH for 100 genes. Probe sets were either purchased from Molecular Instruments or designed with the HCR probe generator(*67*) (https://github.com/rwnull/insitu_probe_generator) and synthesized by IDT (**Table S3**). HCR-FISH was performed following published protocols with the addition of a blocking step(*68–70*). Briefly, brains from *D. melanogaster*, *D. virilis*, *M. domestica*, *A. aegypti*, *T. castaneum*, *N. vitripennis*, and *G. bimaculatus* were dissected in Schneider's insect medium, fixed in 4% FA/PBS for 20 min, rinsed 3x and washed 2x15min in PBST (PBS + 0.1% Triton X-100). While *Drosophila* brains were processed directly, brains from the other species were dehydrated in 100% methanol to improve tissue penetration and allow long-term storage, and were rehydrated prior to the hybridization step through a methanol series in PBST (100%, 75%, 50%, 25% in PBST 0.1%). On day 1, samples were pre-hybridized in 200 µL pre-warmed probe hybridization buffer for 30 min at 37°C in glass staining dishes, ensuring tissues were fully immersed. Probe solution was prepared by adding 1–4 pmol of each probe set to 200 µL of pre-warmed hybridization buffer, supplemented with denatured salmon sperm DNA (0.1 mg/mL; Sigma D7656; boiled for 10 min and cooled on ice for ≥5 min) and fresh BSA (2 mg/mL), and samples were incubated overnight at 37°C in the probe and blocking solution (sealed). On day 2, samples were washed four times for 15 min each in pre-warmed probe wash buffer at 37°C, followed by three 5-min washes in 5×SSCT (0.3% Triton) at room temperature. Samples were then pre-amplified for 30 min in amplification buffer at room temperature. Hairpins were

prepared by snap-heating 12 pmol each of h1 and h2 at 95°C for 90 s, cooling for 30 min at room temperature in the dark, diluting each in 100 µL amplification buffer, and combining to a final volume of 200 µL; samples were incubated overnight in hairpin solution at room temperature (protected from light). On day 3, hairpins were removed and samples were washed in 5×SSCT at room temperature (2 × 5 min, 2 × 30 min, and 1 × 5 min). The brains were mounted in Vectashield and imaged on a ZEISS LSM 980 laser-scanning confocal microscope, using a ×40 glycerol objective. Images were processed in Fiji.

References

1. R. J. V. Roberts, S. Pop, L. L. Prieto-Godino, Evolution of central neural circuits: state of the art and perspectives. Springer Nature [Preprint] (2022). https://doi.org/10.1038/s41583-022-00644-y.
2. N. Konstantinides, C. Desplan, Neuronal Circuit Evolution: From Development to Structure and Adaptive Significance. *Cold Spring Harb Perspect Biol*, a041493 (2024).
3. D. Lee, M. P. Shahandeh, L. Abuin, R. Benton, Comparative single-cell transcriptomic atlases of drosophilid brains suggest glial evolution during ecological adaptation. *PLoS Biol* 23 (2025).
4. I. A. Toker, L. Ripoll-Sánchez, L. T. Geiger, A. Sussfeld, K. S. Saini, I. Beets, P. E. Vértes, W. R. Schafer, E. Ben-David, O. Hobert, Divergence in neuronal signaling pathways despite conserved neuronal identity among Caenorhabditis species. *Current Biology* 35, 2927-2945.e7 (2025).
5. I. Holguera, C. Desplan, Neuronal specification in space and time. *Science* 362, 176–180 (2018).
6. A. Sagner, J. Briscoe, Establishing neuronal diversity in the spinal cord: a time and a place. *Development* 146 (2019).
7. S. Q. Sen, Generating neural diversity through spatial and temporal patterning. *Semin Cell Dev Biol* 142 (2023).
8. N. Konstantinides, I. Holguera, A. M. Rossi, A. Escobar, L. Dudragne, Y. C. Chen, T. N. Tran, A. M. M. Jaimes, M. N. Özel, F. Simon, Z. Shao, N. M. Tsankova, J. F. Fullard, U. Walldorf, P. Roussos, C. Desplan, A complete temporal transcription factor series in the fly visual system. *Nature* 604, 316–322 (2022).
9. A. Sagner, I. Zhang, T. Watson, J. Lazaro, M. Melchionda, J. Briscoe, A shared transcriptional code orchestrates temporal patterning of the central nervous system. *PLoS Biol* 19 (2021).
10. G. D. Frantz, S. K. McConnell, Restriction of late cerebral cortical progenitors to an upper-layer fate. *Neuron* 17, 55–61 (1996).
11. A. R. Desai, S. K. McConnell, Progressive restriction in fate potential by neural progenitors during cerebral cortical development. *Development* 127, 2863–2872 (2000).
12. T. Brody, W. F. Odenwald, Programmed transformations in neuroblast gene expression during Drosophila CNS lineage development. *Dev Biol* 226, 34–44 (2000).
13. T. Isshiki, B. Pearson, S. Holbrook, C. Q. Doe, Drosophila neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. *Cell* 106, 511–521 (2001).
14. H. Zhu, S. D. Zhao, A. Ray, Y. Zhang, X. Li, A comprehensive temporal patterning gene network in Drosophila medulla neuroblasts revealed by single-cell RNA sequencing. *Nat Commun* 13 (2022).
15. C. Bertet, X. Li, T. Erclik, M. Cavey, B. Wells, C. Desplan, Temporal patterning of neuroblasts controls Notch-mediated cell survival through regulation of Hid or Reaper. *Cell* 158, 1173–1186 (2014).
16. O. A. Bayraktar, C. Q. Doe, Combinatorial temporal patterning in progenitors expands neural diversity. *Nature* 498, 449–455 (2013).
17. M. H. Syed, B. Mark, C. Q. Doe, Steroid hormone induction of temporal gene expression in drosophila brain neuroblasts generates neuronal and glial diversity. *Elife* 6 (2017).
18. J. Elliott, C. Jolicoeur, V. Ramamurthy, M. Cayouette, Ikaros confers early temporal competence to mouse retinal progenitor cells. *Neuron* 60, 26–39 (2008).
19. P. Mattar, J. Ericson, S. Blackshaw, M. Cayouette, A conserved regulatory logic controls temporal identity in mouse neural progenitors. *Neuron* 85, 497–504 (2015).

20. A. Javed, P. Mattar, S. Lu, K. Kruczek, M. Kloc, A. Gonzalez-Cordero, R. Bremner, R. R. Ali, M. Cayouette, Pou2f1 and Pou2f2 cooperate to control the timing of cone photoreceptor production in the developing mouse retina. *Development* 147 (2020).
21. K. Toma, T. Kumamoto, C. Hanashima, The timing of upper-layer neurogenesis is conferred by sequential derepression and negative feedback from deep-layer neurons. *J Neurosci* 34, 13259–13276 (2014).
22. I. Sinakevitch, J. K. Douglass, G. Scholtz, R. Loesel, N. J. Strausfeld, Conserved and convergent organization in the optic lobes of insects and isopods, with reference to other crustacean taxa. *J Comp Neurol* 467, 150–172 (2003).
23. M. N. Özel, F. Simon, S. Jafari, I. Holguera, Y. C. Chen, N. Benhra, R. N. El-Danaf, K. Kapuralin, J. A. Malin, N. Konstantinides, C. Desplan, Neuronal diversity and convergence in a visual system developmental atlas. *Nature* 589, 88 (2021).
24. Y. Z. Kurmangaliyev, J. Yoo, J. Valdes-Aleman, P. Sanfilippo, S. L. Zipursky, Transcriptional Programs of Circuit Assembly in the Drosophila Visual System. *Neuron*, doi: 10.1016/j.neuron.2020.10.006 (2020).
25. K. F. Fischbach, A. P. Dittrich, The optic lobe of Drosophila melanogaster. I. A Golgi analysis of wild-type structure. *Cell Tissue Res* 258, 441–445 (1989).
26. R. Coyne, M. Treese, Y.-C. Chen, C. Lake, O. B. Tabi, R. Rajesh, H. Hassan, H. Li, C. Desplan, M. N. Özel, Regulatory logic of neuronal identity specification in Drosophila. *bioRxiv*, 2025.09.01.673531 (2025).
27. B. Misof, S. Liu, K. Meusemann, R. S. Peters, A. Donath, C. Mayer, P. B. Frandsen, J. Ware, T. Flouri, R. G. Beutel, O. Niehuis, M. Petersen, F. Izquierdo-Carrasco, T. Wappler, J. Rust, A. J. Aberer, U. Aspöck, H. Aspöck, D. Bartel, A. Blanke, S. Berger, A. Böhm, T. R. Buckley, B. Calcott, J. Chen, F. Friedrich, M. Fukui, M. Fujita, C. Greve, P. Grobe, S. Gu, Y. Huang, L. S. Jermiin, A. Y. Kawahara, L. Krogmann, M. Kubiak, R. Lanfear, H. Letsch, Y. Li, Z. Li, J. Li, H. Lu, R. Machida, Y. Mashimo, P. Kapli, D. D. McKenna, G. Meng, Y. Nakagaki, J. L. Navarrete-Heredia, M. Ott, Y. Ou, G. Pass, L. Podsiadlowski, H. Pohl, B. M. Von Reumont, K. Schütte, K. Sekiya, S. Shimizu, A. Slipinski, A. Stamatakis, W. Song, X. Su, N. U. Szucsich, M. Tan, X. Tan, M. Tang, J. Tang, G. Timelthaler, S. Tomizuka, M. Trautwein, X. Tong, T. Uchifune, M. G. Walzl, B. M. Wiegmann, J. Wilbrandt, B. Wipfler, T. K. F. Wong, Q. Wu, G. Wu, Y. Xie, S. Yang, Q. Yang, D. K. Yeates, K. Yoshizawa, Q. Zhang, R. Zhang, W. Zhang, Y. Zhang, J. Zhao, C. Zhou, L. Zhou, T. Ziesmann, S. Zou, Y. Li, X. Xu, Y. Zhang, H. Yang, J. Wang, J. Wang, K. M. Kjer, X. Zhou, Phylogenomics resolves the timing and pattern of insect evolution. *Science (1979)* 346, 763–767 (2014).
28. G. Zolotarov, X. Grau-Bové, A. Sebé-Pedrós, GeneExt: a gene model extension tool for enhanced single-cell RNA-seq analysis. *bioRxiv*, 2023.12.05.570120 (2024).
29. N. Nériec, C. Desplan, N. Neriec, C. Desplan, From the Eye to the Brain: Development of the Drosophila Visual System. *Curr Top Dev Biol* 116, 247–271 (2016).
30. U. Tepass, E. Gruszynski-DeFeo, T. A. Haag, L. Omatyar, T. Török, V. Hartenstein, shotgun encodes Drosophila E-cadherin and is preferentially required during cell rearrangement in the neurectoderm and other morphogenetically active epithelia. *Genes Dev* 10, 672–685 (1996).
31. S. L. Lai, M. R. Miller, K. J. Robinson, C. Q. Doe, The Snail family member Worniu is continuously required in neuroblasts to prevent Elav-induced premature differentiation. *Dev Cell* 23, 849–857 (2012).
32. C. C. F. Homem, J. A. Knoblich, Drosophila neuroblasts: a model for stem cell biology. *Development* 139, 4297–4310 (2012).
33. M. Weng, H. Komori, C. Y. Lee, Identification of neural stem cells in the Drosophila larval brain. *Methods Mol Biol* 879, 39 (2012).

34. S. Robinow, K. White, Characterization and spatial distribution of the ELAV protein during Drosophila melanogaster development. *J Neurobiol* 22, 443–461 (1991).
35. D. A. Wagh, T. M. Rasse, E. Asan, A. Hofbauer, I. Schwenkert, H. Dürrbeck, S. Buchner, M. C. Dabauvalle, M. Schmidt, G. Qin, C. Wichmann, R. Kittel, S. J. Sigrist, E. Buchner, Bruchpilot, a protein with homology to ELKS/CAST, is required for structural integrity and function of synaptic active zones in Drosophila. *Neuron* 49, 833–844 (2006).
36. W. C. Xiong, H. Okano, N. H. Patel, J. A. Blendy, C. Montell, repo encodes a glial-specific homeo domain protein required in the Drosophila nervous system. *Genes Dev* 8, 981–994 (1994).
37. C. Piñeiro, C. S. Lopes, F. Casares, A conserved transcriptional network regulates lamina development in the Drosophila visual system. *Development* 141, 2838–2847 (2014).
38. O. Guillermin, B. Perruchoud, S. G. Sprecher, B. Egger, Characterization of tailless functions during Drosophila optic lobe formation. *Dev Biol* 405, 202–213 (2015).
39. H. Apitz, I. Salecker, Spatio-temporal relays control layer identity of direction-selective neuron subtypes in Drosophila. *Nature Communications 2018 9:1* 9, 1–16 (2018).
40. T. Schilling, A. H. Ali, A. Leonhardt, A. Borst, J. Pujol-Martí, Transcriptional control of morphological properties of direction-selective T4/T5 neurons in Drosophila. *Development* 146 (2019).
41. Di. Fuentes, M. Molina, U. Chorostecki, S. Capella-Gutiérrez, M. Marcet-Houben, T. Gabaldón, PhylomeDB V5: an expanding repository for genome-wide catalogues of annotated gene phylogenies. *Nucleic Acids Res* 50, D1062–D1068 (2022).
42. K. Street, D. Risso, R. B. Fletcher, D. Das, J. Ngai, N. Yosef, E. Purdom, S. Dudoit, Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* 19, 477 (2018).
43. D. Song, J. J. Li, PseudotimeDE: inference of differential gene expression along cell pseudotime with well-calibrated p-values from single-cell RNA sequencing data. *Genome Biol* 22 (2021).
44. R. Bodmer, The gene tinman is required for specification of the heart and visceral muscles in Drosophila. *Development* 118, 719–729 (1993).
45. R. P. Kühnlein, C. K. Chen, R. Schuh, A transcription unit at the ken and barbie gene locus encodes a novel Drosophila zinc finger protein. *Mech Dev* 79, 161–164 (1998).
46. T. Erclik, X. Li, M. Courgeon, C. Bertet, Z. Chen, R. Baumert, J. Ng, C. Koo, U. Arain, R. Behnia, A. del Valle Rodriguez, L. Senderowicz, N. Negre, K. P. White, C. Desplan, Integration of temporal and spatial patterning generates neural diversity. *Nature* 541, 365–370 (2017).
47. M. N. Özel, C. S. Gibbs, I. Holguera, M. Soliman, R. Bonneau, C. Desplan, Coordinated control of neuronal differentiation and wiring by sustained transcription factors. *Science* 378 (2022).
48. H. Duan, C. Zhang, J. Chen, H. Sink, E. Frei, M. Noll, A key role of Pox meso in somatic myogenesis of Drosophila. *Development* 134, 3985–3997 (2007).
49. K. Filippopoulou, N. Konstantinides, Evolution of patterning. *FEBS J* (2023).
50. X. Li, S. Visser, J. H. Son, E. Geuverink, E. N. Kıvanç, Y. Wu, S. Schmeing, M. Pippel, S. Y. Anvar, M. A. Schenkel, F. Marec, M. D. Robinson, R. P. Meisel, E. A. Wimmer, L. van de Zande, D. Bopp, L. W. Beukeboom, Divergent evolution of male-determining loci on proto-Y chromosomes of the housefly. *Nat Commun* 15 (2024).
51. M. Vazeille-Falcoz, L. Mousson, F. Rodhain, E. Chungue, A. B. Failloux, Variation in oral susceptibility to dengue type 2 virus of populations of Aedes aegypti from the islands of Tahiti and Moorea, French Polynesia. *Am J Trop Med Hyg* 60, 292–299 (1999).

52. T. Mito, S. Noji, The Two-Spotted Cricket Gryllus bimaculatus: An Emerging Model for Developmental and Regeneration Studies. *CSH Protoc* 2008 (2008).
53. J. Neiro, Postembryonic development of the two-spotted field cricket (Gryllus bimaculatus): a staging system. *bioRxiv*, 2021.02.24.432775 (2021).
54. S. Pop, C. L. Chen, C. J. Sproston, S. Kondo, P. Ramdya, D. W. Williams, Extensive and diverse patterns of cell death sculpt neural networks in insects. *Elife* 9, 1–31 (2020).
55. G. X. Y. Zheng, J. M. Terry, P. Belgrader, P. Ryvkin, Z. W. Bent, R. Wilson, S. B. Ziraldo, T. D. Wheeler, G. P. McDermott, J. Zhu, M. T. Gregory, J. Shuga, L. Montesclaros, J. G. Underwood, D. A. Masquelier, S. Y. Nishimura, M. Schnall-Levin, P. W. Wyatt, C. M. Hindson, R. Bharadwaj, A. Wong, K. D. Ness, L. W. Beppu, H. J. Deeg, C. McFarland, K. R. Loeb, W. J. Valente, N. G. Ericson, E. A. Stevens, J. P. Radich, T. S. Mikkelsen, B. J. Hindson, J. H. Bielas, Massively parallel digital transcriptional profiling of single cells. *Nature Communications 2017 8:1* 8, 14049- (2017).
56. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, Twelve years of SAMtools and BCFtools. *Gigascience* 10 (2021).
57. Y. Hao, T. Stuart, M. H. Kowalski, S. Choudhary, P. Hoffman, A. Hartman, A. Srivastava, G. Molla, S. Madad, C. Fernandez-Granda, R. Satija, Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* 42, 293–304 (2024).
58. C. S. McGinnis, L. M. Murrow, Z. J. Gartner, DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* 8, 329-337.e4 (2019).
59. J. Huerta-Cepas, S. Capella-Gutierrez, L. P. Pryszcz, I. Denisov, D. Kormes, M. Marcet-Houben, T. Gabaldón, PhylomeDB v3.0: an expanding repository of genome-wide collections of trees, alignments and phylogeny-based orthology and paralogy predictions. *Nucleic Acids Res* 39 (2011).
60. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32, 1792–1797 (2004).
61. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30, 772–780 (2013).
62. T. Lassmann, E. L. L. Sonnhammer, Kalign--an accurate and fast multiple sequence alignment algorithm. *BMC Bioinformatics* 6 (2005).
63. I. M. Wallace, O. O'Sullivan, D. G. Higgins, C. Notredame, M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res* 34, 1692–1699 (2006).
64. S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973 (2009).
65. L. T. Nguyen, H. A. Schmidt, A. Von Haeseler, B. Q. Minh, IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32, 268–274 (2015).
66. J. Huerta-Cepas, F. Serra, P. Bork, ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol Biol Evol* 33, 1635–1638 (2016).
67. E. Kuehn, D. S. Clausen, R. W. Null, B. M. Metzger, A. D. Willis, B. D. Özpolat, Segment number threshold determines juvenile onset of germline cluster expansion in Platynereis dumerilii. *J Exp Zool B Mol Dev Evol* 338, 225–240 (2022).
68. H. M. T. Choi, M. Schwarzkopf, M. E. Fornace, A. Acharya, G. Artavanis, J. Stegmaier, A. Cunha, N. A. Pierce, Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* 145 (2018).

69. A. Ferreira, B. Sieriebriennikov, H. Whitbeck, HCR RNA-FISH protocol for the whole-mount brains ofDrosophila and other insects v1. *protocols.io*, doi: 10.17504/PROTOCOLS.IO.BZH5P386 (2021).
70. J. C. Duckhorn, I. Junker, Y. Ding, T. R. Shirangi, Combined in situ hybridization chain reaction and immunostaining to visualize gene expression in whole-mount Drosophila central nervous systems. *bioRxiv*, 2021.08.02.454831 (2021).

Acknowledgements

We thank all colleagues who kindly provided animals and expertise for insect rearing: Leo Beukeboom, Anna Rensink, Daniel Bopp, Claudia Brunner-Barios, Siegfried Roth, Matthias Pechmann, Eloise Muller, Isabel Almudi, Gregor Bucher, Marie Vazeille and Anna Bella Failloux, and Manon Monier. We also thank members of the Konstantinides lab for valuable discussions, especially Rebekah Ricquebourg for help with animal rearing and Isabel Holguera for expertise in mounting and imaging. We also acknowledge the ImagoSeine core facility at the Institut Jacques Monod, a member of France-BioImaging (ANR-10-INBS-04) and certified by GIS-IBiSA. This work is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 949500) and the HORIZON-WIDERA-2023-ACCESS-02 grant no. 101159925 - SCENTINEL. TG group acknowledges support from the Spanish Ministry of Science and Innovation (grant numbers PID2021-126067NB-I00, CPP2021-008552, PCI2022-135066-2, PLEC2023-010225, and PDC2022-133266-I00), cofounded by ERDF "A way of making Europe", as well as support from the Catalan Research Agency (AGAUR) (grant number SGR01551); Gordon and Betty Moore Foundation (grant number GBMF9742); "La Caixa" foundation (grant number LCF/PR/HR21/00737), Fundació La Marató de TV3 (202328-31), AECC (PRYGN234923GABA), and Instituto de Salud Carlos III (CIBERINFEC CB21/13/00061- ISCIII-SGEFI/ERDF).

Author contributions

N.K. and K.F. designed the project. K.F. established and maintained all animal cultures in the laboratory. K.F. designed protocols and performed experiments, imaging, and single-cell data generation and downstream computational analysis. K.F., C.L., and J.J.L. developed the code for temporal transcription factor selection. E.I. improved genome annotations and processed single-cell sequencing reads. K.F. and C.J.L.M. performed HCR-FISH and imaging. M.M.-H. and T.G. performed phylogenomic analyses. K.F. and N.K. interpreted the data and wrote the manuscript. All authors edited and approved the final manuscript.

Competing interests statement

The authors declare no competing interests.

Supplementary figure legends

Fig. S1. EdU incorporation identifies the peak neuroblast proliferative stage. (A) EdU staining in *A. aegypti* early L4 stage. (B) EdU staining in *B. mori* late L4 stage. (C) EdU staining in *T. castaneum* L4 and L5 stages. (D) EdU staining in *N. vitripennis* early L3 stage. (E) EdU staining in *G. bimaculatus* stages 3, 4, 5, and 7. (F) EdU staining in *C. dipterum* early and mid-late nymphs.

Fig. S2. GeneExt pipeline and results. (A) GeneExt pipeline that was followed for the improvement of genome annotation in the eight insect species. (B) Violin plots of gene extension lengths and numbers of extended genes across eight insect species.

Fig. S3. Quality control and filtering of insect single-cell mRNA sequencing libraries. For each of the thirty libraries, a histogram showing the distribution of detected features per cell (left) and violin plots of the number of detected features per cell before (middle) and after (right) filtering are shown. The red dotted line indicates the library-specific filtering threshold.

Fig. S4. Cell-class identification across insect species by marker gene expression. UMAP projections are shown for *D. virilis* (A), *M. domestica* (B), *A. aegypti* (C), *B. mori* (D), *T. castaneum* (E), *N. vitripennis* (F), *G. bimaculatus* (G), and *C. dipterum* (H), coloured by the expression of selected marker genes used to identify major neural cell classes, including neuroepithelial cells, neuroblasts, ganglion mother cells (GMCs), and glial cells. These markers enable the assignment of cells to distinct cell classes within each species.

Fig. S5. Neuropil identification across insect species by marker gene expression. UMAP projections are shown for *D. virilis* (A), *M. domestica* (B), *A. aegypti* (C), *B. mori* (D), *T. castaneum* (E), *N. vitripennis* (F), *G. bimaculatus* (G), and *C. dipterum* (H), coloured by the expression of selected marker genes used to identify different neuropils. These markers enable the assignment of cells to the lamina and lobula plate.

Fig. S6. Phylome reconstruction and ortholog identification. (A) Percentage of genes in the species of the x-axis that have at least one ortholog in the species of the y-axis. As the relationship between orthologs may not be one-to-one and the proteome sizes vary the graph is not symmetrical. (B) Number of *D. melanogaster* genes that have one-to-one (dark blue), one-to-many (cyan), many-to-one (green), many-to-many (pink) orthologs or are *D. melanogaster*-specific (grey) for each of the species used the in the phylome reconstuction.

Fig. S7. Pipeline for the identification of candidate tTFs. (A) For each gene, expression was plotted along pseudotime. Two filters were applied to ensure that expression was temporally regulated (i.e., present in specific segments of the trajectory but absent in others). Genes passing these criteria were further evaluated based on expression levels and PseudotimeDE p-values, and selected candidates were subsequently tested by HCR-FISH. (B) HCR-FISH analysis of *elb* (purple) in the developing optic lobes of *D. virilis*. Neuroblasts are visualized by *dpn* expression (white). No temporal expression of *elb* was detected; therefore, this candidate tTF was discarded.

Fig. S8. Expression of all candidate tTFs along the neuroblast trajectory. Dynamic gene expression along the neuroblast pseudotime trajectory (left) and UMAP visualization of gene expression in neuroblasts (right) are shown for all candidate tTFs in all eight species. Genes highlighted in red or green were tested by HCR-FISH; among these, genes highlighted in red

represent species-specific tTFs. Genes highlighted in light or dark grey were not cross-validated by HCR-FISH; those in dark grey correspond to *D. melanogaster* tTF orthologs. Genes below the horizontal dark line were eliminated from the candidate tTF list either because cross-validation did not support their temporal expression or because they were expressed at low levels.

Fig. S9. Experimental validation of predicted temporal transcription factors. (A) In *D. melanogaster, hth* (left, green), *tll* (left, magenta), and *scro* (right, magenta) HCR-FISH in neuroblasts (*dpn,* white) *agrees* with the antibody stainings (*8*). (B) HCR-FISH expression of *hth* (left, green), *hbn* (left, magenta), and *tll* (right, magenta) in *D. virilis* neuroblasts (*dpn*, white). (C) HCR-FISH expression of *hth* (left, green), *tll* (left, magenta), *erm* (middle, magenta) and *Bar-H* (right, magenta) in *M. domestica* neuroblasts (*dpn*, white). (D) HCR-FISH expression of *hth* (left, green), *slp* (middle-left, green), *hbn* (middle-left, magenta), *Bar-H* (middle-right, magenta), and *tll* (right, magenta) in *A. aegypti* neuroblasts (*dpn*, white). (E) HCR-FISH expression of *hth* (left, magenta), *D* (left, green), and *ey* (right, magenta) in *G. bimaculatus* neuroblasts (*mira*, white). (F) HCR-FISH expression of *opa* (left, green), *scro* (left, magenta), *Bar-H* (middle, magenta), and *tll* (right, magenta) in *N. vitripennis* neuroblasts (*dpn*, white). (G) HCR-FISH expression of *dmrt99B* (left, green), *tll* (left, magenta), *hth* (middle-left, magenta), *Bar-H* (middle-left, green), *opa* (middle, magenta), *slp* (middle, green), *hbn* (middle-right, green), *opa* (middle-right, magenta), *Bar-H* (right, green), and *tll* (right, magenta) in *T. castaneum* neuroblasts (*wor*, white).

Fig. S10. Experimental validation of new temporal transcription factors. (A) HCR-FISH shows *Ets65A* (magenta) expression in old *A. aegypti* neuroblasts (*dpn*, white). (B) HCR-FISH shows *cas* (left, magenta) and *grh* (right, magenta) expression in mid and old *N. vitripennis* neuroblasts (*dpn*, white), respectively.

Fig. S11. The absence of the *ey* temporal window leads to the absence of its neuronal progeny in *Musca*. UMAP projections showing the expression of *ey*, *erm*, *vvl*, *kn*, *TfAP-2*, and *svp* in developing visual brains of *M. domestica* (top) and *D. melanogaster* (bottom). In *D. melanogaster*, these genes mark neuronal types that come from different neuroblast temporal windows (Hth, Hth/Opa, Erm/Ey, and Ey/Hbn windows). In *M. domestica*, all corresponding neuronal populations are detected except those expressing the *ey/toy* homolog (*pax6*), consistent with the absence of an Ey temporal window. NotchON neurons expressing *vvl* are markedly fewer in *M. domestica* than in *D. melanogaster* and co-express *erm*, indicating they derive from the Erm temporal window rather than the Ey window.

Supplementary Tables

Table S1: Transcription factor lists for each insect species.

Table S2: Candidate temporal transcription factor lists for each insect species.

Table S3: HCR-FISH probe sets used for validation of candidate temporal transcription factors

Table S4: Insect strains, genomes, and reagents