Synaptic gene expression is coordinated across development by conserved chromatin regulators

Authors and affiliations:
James A. Kentro[1,2,3], Gunjan Singh[1,2], Tuan M. Pham[1,2,4], Justin Currie[1,2,3], Saniya Khullar[5,6], Audrey T. Medeiros[2], Erica Larschan[1,3,*], Kate M. O'Connor-Giles[2,3,*]

1. Department of Molecular Biology, Cell Biology, & Biochemistry, Brown University, Providence, RI, USA
2. Department of Neuroscience, Brown University, Providence, RI, USA
3. Carney Institute for Brain Science, Brown University, Providence, RI, USA
4. Center for Computational Molecular Biology, Brown University, Providence, RI, USA
5. Waisman Center, University of Wisconsin-Madison, Madison, WI, USA
6. Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI, USA

* Co-corresponding authors: erica_larschan@brown.edu, kate_oconnor-giles@brown.edu

The authors declare no competing interests.

Abstract

Chemical synapses are the primary sites of communication in the nervous system. Synaptogenesis is a complex transcellular process involving hundreds of proteins that must be expressed at the same time. How this spatiotemporal coordination is achieved remains an open question. We find that synaptic genes are broadly and specifically coordinated at the level of transcription across species. Through genomic and functional studies in *Drosophila*, we demonstrate corresponding coordination of chromatin accessibility and identify chromatin regulators DEAF1 and CLAMP as broad repressors of synaptic gene expression. Disruption of either factor causes increased synaptic gene expression across neuronal subtypes and ectopic synapse formation. We further find that DEAF1, which is implicated in syndromic intellectual disability, is both necessary and sufficient to repress synapse formation. Our findings reveal the critical importance of broad temporally coordinated negative regulation of synaptic gene expression in sculpting neural circuit formation and identify two key repressors.

Introduction

The nervous system is remarkable for its cellular diversity and functional complexity. In the human brain, billions of neurons of distinct subtypes form trillions of synapses that connect neurons into functional circuits that interpret the environment and control behavior. There has been significant progress in understanding how the great diversity of neurons is generated[1–4]. Yet, how neurons acquire shared traits has been less studied.

The chemical synapses that underlie neuronal communication are hallmarks of mature neurons. Synapses are complex intercellular structures designed for the rapid transmission of information in the form of Ca^{2+}-dependent neurotransmitter release and uptake. Building a functional synapse necessitates the simultaneous expression of hundreds of proteins, including neurotransmitter transporters and receptors, regulators of synaptic vesicle release, and scaffolding and adhesion molecules, as well as the proteins that regulate their function and transport to synapses[5,6]. Adding to the challenge, the expression of synaptic genes, which we define broadly here as all genes that encode proteins required for the formation of functional synapses, must not only be coordinated within a given cell, but across synaptically connected neurons. How the spatio-temporal coordination of synaptic gene expression is achieved remains an open question. While this coordination can and likely does occur at multiple levels, including transcription, translation, and protein trafficking, our analysis of *Drosophila* time-series transcriptomic data indicates that hundreds of synaptic genes are co-regulated at the level of transcription. We show that this coordinated regulation of synaptic genes is also conserved in mouse.

Here, we investigate the mechanisms and regulators that temporally coordinate the expression of synaptic genes across neurons. Consistent with prior studies implicating chromatin regulation in neuronal maturation[7,8], we find that chromatin accessibility is dynamically and coordinately regulated across large numbers of synaptic genes. Using motif analysis, chromatin binding data, and a new machine learning approach, we identify a gene regulatory network that includes transcription factors DEAF1, which is linked to intellectual disability in humans, and CLAMP, a pioneer factor that regulates chromatin accessibility, as candidate negative regulators of coordinated synaptic gene expression. DEAF1 and CLAMP exhibit enriched co-accessibility at synaptic genes and co-bind chromatin at a subset of their binding sites. Through *in vivo* studies, we find that DEAF1 and CLAMP function as direct repressors for a majority of synaptic genes that are directly bound by both factors. Finally, we find that loss of DEAF1 or CLAMP specifically results in ectopic synapse formation. Conversely, overexpression of DEAF1 blocks synapse formation, indicating an instructive role. Overall, we demonstrate that synaptic gene transcription is broadly co-regulated across species and define the critical importance of negative regulators of gene expression in nervous system development.

Results

Broad temporal coordination of synaptic gene transcription

Synaptogenesis requires the simultaneous expression of hundreds of genes required for synapse formation and function. Coordinated expression could occur at the level of transcription or translation. To determine whether synaptic genes are broadly transcriptionally co-regulated, we examined the developmental transcriptional expression profiles of well-characterized *Drosophila* genes that encode proteins critical for synapse formation and function, including calcium channels, neurotransmitter transporters and receptors, active zone scaffolding molecules, and synaptic cell adhesion molecules among many others[9–13]. We observed similar expression patterns across embryonic, larval, and pupal development, suggesting coordinated regulation at the level of transcription (Fig. 1A). This shared expression pattern reveals coordinated upregulation of synaptic gene expression during mid-embryogenesis at ~12 hours, immediately prior to widespread synaptogenesis in the developing embryonic nervous system. Expression peaks in late stage embryogenesis, then drops during larval development to very low levels in third instar larvae when primary synaptogenesis is complete. Synaptic genes are again coordinately upregulated during early pupal stages, with expression peaking shortly before eclosion. As in embryos, this peak coincides with peak synaptogenesis during the development of the adult nervous system. Therefore, well-characterized synaptic genes exhibit similar expression patterns, suggesting co-regulation at the transcriptional level.

To identify additional co-regulated genes, we generated a composite profile using 20 well-characterized synaptic genes (Fig. S1A) and compared this ‘synaptic’ footprint to every gene in the *Drosophila* genome, calculating a Spearman correlation score for each. Examination of this ranked list revealed that genes with high correlation scores are all neuronally enriched with the majority directly involved in synapse development and function based on enrichment for synaptic signaling gene ontology (GO) terms (Fig. 1B). Synaptic GO terms become less significant among genes with correlation scores below 0.80, a group that also includes few well-characterized synaptic genes (Fig. S1B). Therefore, in our study of coordinated synaptic gene expression we focused on the 485 genes with a correlation score above 0.80 (Table S1).

The majority of these genes have well-characterized roles in synapse formation and/or function and include most of the genes that encode the factors that regulate the synthesis, transport, and reception of glutamate, acetylcholine, gamma-aminobutyric acid (GABA), serotonin, dopamine, and octopamine/tyramine along with those that encode core pan-synaptic proteins. 168 of the 485 co-regulated genes have not been characterized in *Drosophila* and are designated as computed genes (CGs). These include genes encoding homologs of ion channels, solute carriers, enzymes, and cytoskeletal proteins linked to synapses in other species as well as genes linked to neurodevelopmental disorders, and recent studies in *Drosophila* have identified roles in synapse formation or function for several, including TRMT9B/fid[14] and PDZD8[15–18]. GO analysis of human orthologs of these CGs indicate enrichment for ion transport and other synaptic pathways. Consistently, all 485 genes are expressed in the nervous system, with the majority nervous system-specific or highly enriched. Their highly correlated temporal transcriptional profile suggests a regulatory mechanism that coordinates the expression of pan-neuronal and cell-type specific synaptic genes across neuronal subtypes.

In pupating insects, many tissues undergo morphogenesis requiring two distinct phases of gene expression. We therefore examined the specificity of the synaptic gene temporal transcriptional expression profile by comparing genes involved in earlier stages of neurodevelopment, specifically neurite outgrowth and axon development. As with the synaptic genes, we created a composite profile of 20 well-characterized genes necessary for neurite development and axon guidance (Fig. S1C) and calculated a Spearman correlation score for every neuronal gene. The 485 most correlated genes exhibit no overlap with the synaptic gene list and GO analysis reveals enrichment for neuron projection development, axon guidance, and broader developmental parent categories (Fig. S1D, Table S2). Like synaptic genes, neurite/axon development genes exhibit two distinct peaks of expression, but notably these peaks precede peaks in synaptic gene expression in both embryos and pupae (Fig. 1A). These temporally separated expression patterns demonstrate the ordered expression of broad cohorts of functionally related genes during nervous system development.

To further investigate the specificity of the synaptic gene temporal transcriptional profile, we performed an unbiased k-means clustering analysis of all genes expressed in the nervous system based on temporal expression. We identified 12 clusters with distinct temporal expression patterns, one of which closely matches our synaptic temporal expression profile (Fig. S1E, cluster 8). Cluster 8 overlaps with 324/485 coordinately regulated synaptic genes (padj = $0x10^{00}$). The next most similar profile (cluster 1) overlaps with 85 of our synaptic genes (padj = $6x10^{-32}$), while all other clusters have non-significant or zero overlap (Fig. S1F). We performed Gene Ontology (GO) analysis on the 12 clusters and found that cluster 8 was enriched by synaptic gene ontology categories (Fig. S1G). In comparison, clusters 1 and 2 are enriched for axon development and more general neurodevelopmental GO categories and have peaks of expression earlier in development (Fig. S1E). This analysis shows a logical progression through neuronal maturation with clusters enriched for GO categories from early to late neuronal development, consistent with observations in mammals[19,20].

Finally, to independently validate the broad temporally coordinated transcription of synaptic genes, we analyzed time-series single-cell RNA-seq (scRNA-seq) data sets[21–23]. Specifically, we clustered whole embryo data to obtain embryonic nervous system cells at six timepoints for analysis along with five larval whole brain and 9 pupal visual system timepoints. We then calculated the average synaptic gene expression level in each cell and plotted average expression across timepoints. We observe a temporal expression pattern that matches our bulk analysis (Fig. 1C-D).

We hypothesized that coordinated expression of large numbers of synaptic genes is a conserved feature of nervous system development across species. To test this hypothesis, we examined time-series RNA-seq data from the developing mouse brain beginning at embryonic day 10.5 [24]. To generate an analogous set of synaptic genes, we created an expression footprint for stages E10.5 through E12.5 in whole brain and E13.5 through P0 in forebrain and hindbrain (separately) using the orthologs of the same 20 synaptic genes we used in flies (see methods, Fig. 1E). We averaged these patterns and calculated correlation scores for all expressed genes.

We identified 742 genes with correlation scores above 0.95 and performed GO analysis. The most significantly represented categories are involved in synaptic development and transmission (Fig. 1F), suggesting that coordinated temporal transcriptional regulation of synaptic gene expression is a conserved feature of nervous system development.

Chromatin accessibility is dynamically and coordinately regulated across synaptic genes

Synaptic genes are widely dispersed across chromosomes (Fig. S2A), raising the question of how their expression is coordinated. Hypothesizing that dynamic coordination of chromatin accessibility might play a role, we analyzed existing DNase-seq time series data from isolated neuronal and mesodermal embryonic tissue[25]. Promoter regions upstream of synaptic genes are closed relative to a random set of nervous system expressed genes in early stages (Fig. 2A). Chromatin becomes substantially more accessible at synaptic genes relative to random neuronal genes beginning at 10-12 hours and immediately preceding the coordinated upregulation of synaptic gene transcription (Fig. 2B). This increase in chromatin accessibility is specific to the nervous system (Fig. S2C). As before, we also examined chromatin accessibility at 485 genes involved in the earlier process of neurite development. In contrast to the chromatin at synaptic genes, chromatin at these genes is already accessible by 4-6 hours of embryogenesis. Therefore, chromatin accessibility at synaptic genes precedes their transcription in a tissue and stage specific manner consistent with findings for other groups of coordinately regulated genes[25,26]**.** Using time series proteomic data available for a subset of synaptic proteins[27], we observe a coordinated increase in synaptic protein abundance immediately following chromatin opening and transcription initiation, suggesting that this coordination influences protein production (Fig. S2B).

DEAF1 motifs and binding are enriched at dynamically regulated accessible regions near synaptic genes

To define transcription factors (TFs) that could coordinately regulate synaptic gene expression, we identified newly accessible regions in 10-12 hour neurons within a 3 kilobase window around each synaptic gene and identified enriched TF binding motifs. These analyses identified 27 enriched motifs, with the DEAF1 binding motif most significantly enriched (Fig. 2C). DEAF1 is a highly conserved neuronally enriched TF[28–30] that is associated with developmental delay and intellectual disability[31,32].

Multiple TFs are often capable of binding to the same motifs with varying levels of affinity and many sites remain unbound, so we next examined available chromatin binding data for over 450 TFs to identify TFs actively bound at newly accessible regions near synaptic genes[33,34]. We compared binding peaks at newly accessible regions within 3kb of synaptic genes to those near three random control sets of neuronally expressed genes. Of all the TFs, DEAF1 showed the largest difference in average profile peak height at newly accessible regions near synaptic vs. random genes (Fig. 2D). DEAF1 also binds at the most newly accessible regions around synaptic genes (~50%) and the largest percentage of synaptic genes (75%, Table S3-4). Thus, DEAF1 is a strong candidate regulator of synaptic gene expression.

DEAF1 and CLAMP are predicted to co-repress synaptic gene expression

In parallel, we took an independent approach to identify candidate regulators of coordinated synaptic gene expression. We hypothesized that by correlating the expression of TFs with the expression of synaptic genes across multiple stages of neuronal development, we could identify TFs that may function together to co-regulate synaptic target genes. To identify correlated TF/synaptic gene expression, we used a recently developed machine learning tool called NetREm[35]. NetREm analyzes single-cell RNA-seq (scRNA-seq) expression data and incorporates protein-protein interaction (PPI) data as a network constraint among the TF predictors along with binding motif accessibility at target genes (TGs) from single-cell ATAC-seq (scATAC-seq) data to generate weighted correlation scores between TF-TF pairs. By linking multiple TF-TF pairs, NetREm identifies broader networks that cooperatively or antagonistically co-regulate the expression of TGs.

We first clustered wild-type cells from the recently published scRNA-seq and scATAC-seq third instar larval brain atlas[36]. We annotated clusters broadly to capture substantial differences in synaptic gene expression during neuronal development: neural progenitors, where synaptic genes are inactive; immature neurons, which express only low levels of some synaptic genes; mature neurons, which broadly express synaptic genes; and glia, which were excluded from further analysis (Fig. 3A-B). As input for NetREm analysis, we included the 15 TFs bound at the most newly accessible regions near synaptic genes (Table S3), all unique TFs identified by the motif analysis above (see Fig. 2C), and a second independent motif analysis (see methods, Fig. S3A) for a total of 149 TFs. With this input, NetREm provided weighted coordination scores for each TF-TF pair. Across neuronal clusters, ten TFs predicted to regulate at least 280 synaptic target genes were identified, including DEAF1, which was predicted to regulate 294 TGs (64.3%; Fig. S3B).

Since DEAF1 has the most enriched motif (Fig. 2C) and chromatin binding (Fig. 2D) at newly accessible sites near synaptic genes and is among the top candidates again, we next used the NetREm analysis to predict a DEAF1 co-regulatory network (Fig. 3C). Four TFs have consistently high coordination scores (above |0.90|): computed gene CG4360, Chromatin-linked adaptor for MSL proteins (CLAMP), Medea (Med), and NFAT nuclear factor (NFAT). Four TFs have experimentally verified physical interactions with DEAF1: CLAMP, NFAT, Tramtrak (Ttk), and abdominal A (Abd-A), ranked 1, 2, 3, and 4, respectively. CLAMP and NFAT each have strong coordination scores and known physical interactions with DEAF1. In mature neurons, DEAF1 and NFAT are predicted to act antagonistically in the regulation of synaptic genes, while DEAF1 and CLAMP are predicted to act cooperatively. Notably, DEAF1 and CLAMP are predicted to act together to repress synaptic gene expression during this critical period.

To independently validate the relationship between DEAF1, CLAMP and synaptic genes, we analyzed a scRNA-seq and scATAC-seq atlas spanning embryogenesis[21]. Following clustering and annotation (Fig. 3D), we compared *deaf1*, *clamp*, and synaptic gene expression from embryonic hours 12 to 16 (Fig. 3E). We found that while the expression of both *deaf1* and *clamp* correlate both positively and negatively with a large number of neuronal genes, both are specifically anti-correlated with the synaptic gene expression, consistent with the NetREm prediction that the two factors function together as repressors of synaptic gene expression.

We next used the corresponding scATAC-seq data to determine if the accessibility of DEAF1 and CLAMP binding motifs is coordinated at synaptic genes. We grouped regions into sets of consistently co-accessible regions and identified Differentially Accessible Regions (DARs) across cell-types. The DEAF1 binding motif was accessible at 308 (63.5%) synaptic genes, the CLAMP binding motif was accessible at 230 (47.4%), and 174 synaptic genes (35.9%) had accessible motifs for both DEAF1 and CLAMP (Fig 3G). Of these 174 synaptic genes, 87 have spatially neighboring co-accessible motifs (Table S5). Genes with co-accessible motifs include pan-neuronal synaptic active zone components and sub-type specific components of excitatory glutamatergic and cholinergic and inhibitory GABAergic systems (Fig. 3H-I, SJ-K). These data suggest that DEAF1 and CLAMP bind simultaneously to dynamically regulated regions at a subset of synaptic genes.

DEAF1 and CLAMP colocalize at synaptic genes

CLAMP directs many different types of regulators to the correct genomic locations to drive context-specific functions, including directing dosage compensation[37] and splicing complexes[38]. Therefore, we hypothesized that colocalization of DEAF1 and CLAMP at a subset of sites may be necessary to specify the repression of many of our synaptic genes. DEAF1 and CLAMP have each been previously found to bind to hundreds of sites on polytene chromosomes[39–42], so we co-stained polytene chromosomes to visualize their colocalization across the genome in a tissue in which synaptic genes are repressed (Fig. 4A). We observe DEAF1 and CLAMP colocalization at a subset of sites across all chromosomes, consistent with their combined binding defining a subset of co-regulated genes.

However, the resolution of polytene chromosomes does not allow mapping of specific target genes. Therefore, we again turned to DNase-seq accessibility data[25] in 10-12 hour neurons, when chromatin at synaptic genes first becomes accessible, to determine if co-accessibility of DEAF1 and CLAMP motifs is enriched at synaptic genes. Measured separately, the accessibility of DEAF1 or CLAMP motifs is approximately 2-fold more enriched at synaptic genes than at random neuronally expressed genes (padj = $3x10^{-8}$ and $3x10^{-7}$, respectively). Notably, accessible peaks containing both motifs (co-accessibility) are 4-fold higher (padj = $1x10^{-41}$) at synaptic genes than other neuronally expressed genes (Fig. 4B). This 2-fold enrichment of co-accessibile vs. individually accessibile DEAF1 and CLAMP motifs at synaptic genes further supports the prediction that DEAF1 and CLAMP coordinately regulate synaptic gene expression.

We next investigated DEAF1 and CLAMP ChIP-seq data[34,43] for co-binding and observed significant overlap at synaptic genes (Fig. 4B-C). Specifically, analysis of ChIP-seq data reveals DEAF1 binding at 75% of coordinately regulated synaptic genes (padj = $2x10^{-11}$) and CLAMP binding at ~40% (padj = $2x10^{-11}$). At 85% of synaptic genes bound by CLAMP, we also observe binding by DEAF1 (~34% of synaptic genes, padj = $2x10^{-23}$). Together these data suggest that DEAF1 and CLAMP co-regulate synaptic gene expression through co-binding at dynamically accessible sites.

Loss of DEAF1 or CLAMP derepresses synaptic gene expression

Our modeling predicts that DEAF1 and CLAMP function together to repress synaptic gene expression. If our model is correct, loss of either factor would be expected to dysregulate temporally coordinated synaptic gene expression. To test our model *in vivo*, we focused on the third instar larval stage, when synaptic gene expression is normally repressed. We knocked down *deaf1* or *clamp* using *elav*-Gal4, which drives expression in all neurons and salivary glands, to express RNAi constructs, confirmed loss of each factor on chromatin (Fig. S4), and conducted RNA-seq. Knock down of *deaf1* resulted in the differential expression of 2862 genes relative to RNAi controls, 1558 upregulated and 1304 downregulated. GO analysis of all significantly upregulated genes following *deaf1* knockdown identified the most significant enrichment for neurodevelopment terms including synaptic organization and synaptic signaling. Consistent with broad DEAF1-mediated repression of synaptic genes, every gene annotated to the "maintenance of presynaptic active zone structure" or "regulation of calcium ion transmembrane transporter activity" terms is significantly upregulated (Fig 5A). In contrast, GO enrichment analysis of genes downregulated upon *deaf1* knockdown yielded broad metabolic process terms, including peptide biosynthesis and lipid oxidation (Fig. S5A).

An analysis of the 485 coordinately regulated synaptic genes, reveals that 207 are upregulated, while only 23 are downregulated in the absence of *deaf1* (Fig. 5B, S5B). Among the upregulated genes are those that encode: 1) active zone components (including nSyb, brp, Rim, Syt1, Ca-β); 2) synaptic cell adhesion molecules (including multiple dpr, side, beat, Nlg, CadN proteins); 3) neurotransmitter transporters and receptors from diverse neuronal subtypes, including vesicular glutamate and monoamine transporters (VGlut and Vmat), Choline acetyltransferase (ChAT), and receptors for acetylcholine, GABA, glutamate, dopamine, and the serotonin receptors 5-HT1A and 5-HT1B. Interestingly, DEAF1 has been implicated as a key regulator of 5-HT1A expression in human cell lines[44,45], indicating conserved roles for DEAF1 between *Drosophila* and humans. These data suggest that DEAF1 is a broad negative regulator of synaptic gene expression.

Knockdown of *clamp* in all neurons resulted in the differential expression of 1566 genes relative to RNAi controls, 861 upregulated and 703 downregulated. Similar to *deaf1* knockdown, of 485 synaptic genes, 106 were upregulated and only 27 downregulated (Fig. 5B, S5B). GO analysis of all upregulated genes shows enrichment of terms for synaptic signaling, calcium transport and homeostasis, and neurotransmitter transport (Fig. 5C), while downregulated genes are enriched for protein catabolism and intracellular transport GO terms (Fig. S5C). We also observe upregulated GO term enrichment for organic acid metabolism and glucose metabolism, suggesting that CLAMP also has a broader role in negatively regulating gene expression in the nervous system. Synaptic genes upregulated upon *clamp* knockdown include 63 genes also upregulated upon *deaf1* knockdown, including genes encoding both cell-type specific synaptic proteins such as Vmat, ChAT, and receptors for acetylcholine, glutamate, and dopamine, presynaptic remodeler as well as common synaptic proteins such as Syt, which promotes calcium-dependent neurotransmitter release. We observed a similar upregulation of synaptic genes in *clamp* null third instar larval brains[46]. Additionally, we overlapped the up and downregulated genes with previously examined ChIP-seq data[34,43] and found direct binding

enrichment by DEAF1 and CLAMP at upregulated, but not downregulated genes (Fig. S5D). Together, these results demonstrate that DEAF1 and CLAMP directly repress synaptic gene expression across diverse neuronal subtypes (Fig. 5D, S5D).

DEAF1 and CLAMP restrain synapse formation

Synapse formation and refinement are critical developmental processes. In wild-type animals synaptic gene transcription peaks during embryonic nervous system development, then falls to very low levels following this window of robust synaptogenesis (see Fig. 1A). This observation suggests that both coordinated activation and broadly coordinated repression of synaptic genes may be critical for proper nervous system development.

To determine the role of DEAF1/CLAMP-dependent synaptic gene repression on synaptogenesis , we turned to the well-characterized, highly stereotyped third instar larval neuromuscular junction. Here, glutamatergic motor neurons innervate single muscle cells and form a stereotyped number of synaptic boutons, each containing ~10 individual synapses. Pan-neuronal knockdown of *deaf1* caused a 56% increase in synaptic bouton numbers and significantly increased active zone numbers compared to control (Fig. 6A-E). We also observe a 54% increase in synapse formation in a heteroallelic combination of *deaf1* alleles (*deaf1*$^{k3/S10B}$) that produces rare escapers that survive to third instar larvae (Fig. S6A-B). Similarly, pan-neuronal knockdown of *clamp* resulted in a 28% increase in synaptic bouton formation, and significantly increased active zone numbers (Fig. 6F-I). In *clamp*^{KG00473} hypomorphs we observe a 24% increase in synaptic boutons (Fig. S6C-D), similar to previous observations[47]. Thus, transcriptional repression of synaptic gene expression by DEAF1 and CLAMP is required to constrain synaptic growth during nervous system development.

Our findings suggest that activation and repression of synaptic gene transcription is dynamically controlled through the temporally regulated actions of TFs acting in opposite directions. This further suggests that exuberant repression of synaptic genes might inhibit synapse formation. To determine whether DEAF1 misexpression is sufficient to repress synapse formation, we overexpressed *deaf1* in developing neurons and assessed synapse formation at the NMJ. We observed a reduction in bouton size and 19% decrease in bouton formation (Fig. 6F-H). Strikingly, synapse formation was reduced by 52% (Fig. 6I). Thus, the dynamic regulation of synaptic gene activation and repression plays a critical role in controlling synapse formation during nervous system development. Together, *in vivo* loss- and gain-of-function experiments validate our computational predictions and reveal a previously unknown role for DEAF1 and CLAMP in repressing synaptic gene expression and synaptogenesis.

Discussion

We combined analysis of single-cell transcriptomics and chromatin accessibility with lineage-resolved DNase time-series data to identify regulators of coordinated synaptic gene expression. Our findings identify two key repressors and reveal the critical importance of broad negative regulation of synaptic gene expression across neuronal subtypes in sculpting neural circuit formation.

Synapse number is tightly controlled through the regulation of synapse formation and maintenance during development. Altered synapse number is a common feature of neurodevelopmental disorders, highlighting the importance of proper regulation. Synapse formation requires the coordinated function of hundreds of proteins, suggesting that synaptic gene expression must be tightly regulated. Indeed, we find that large numbers of synaptic genes, including those encoding pan-synaptic and subtype specific synapse components, are co-regulated at the level of transcription (Fig. 1). These same synaptic genes are coordinately turned on in mammals, consistent with previous observations of temporal waves of gene expression programs during nervous system development[19,20]. We observe a widespread increase in chromatin accessibility upstream of the majority of synaptic genes over a short window immediately preceding the coordinated transcription of synaptic genes, indicating a key role for dynamic chromatin regulation (Fig. 2). In addition to coordinated activation, we observe temporally coordinated repression of synaptic genes following developmental windows of peak synaptogenesis. This observation is consistent with a large literature revealing dynamic changes in chromatin accessibility throughout development and neuronal maturation[19,48,49].

Through motif analysis, we identified DEAF1 as a key TF associated with dynamically accessible regions. DEAF1 interacts with an array of chromatin regulators that both positively and negatively regulate chromatin accessibility in different contexts[32,50,51]. Through analysis of single-cell transcriptomic and chromatin accessibility data with a new network analysis tool that incorporates protein-protein interaction data, we identified CLAMP as a DEAF1 co-regulator (Fig. 3). Notably, CLAMP motifs are also enriched at dynamically regulated sites upstream of synaptic genes. Expression of both DEAF1 and CLAMP are strongly negatively correlated with synaptic gene expression, thus they are predicted to repress synaptic gene expression. Consistently, we observe DEAF1 and CLAMP colocalizing on polytene chromosomes and co-binding at a significant fraction of synaptic genes (Fig. 4).

While multiple TFs required for the activation of gene expression programs in neurodevelopment have been identified[1,47,52–55], negative regulators have been more difficult to identify. To determine if DEAF1 and CLAMP are negative regulators of synaptic gene expression, we conducted RNA-seq in the absence of each factor at a developmental time point when synaptic gene expression is normally repressed. In both cases, we observed upregulation of synaptic gene expression. Furthermore, both factors are directly bound at upregulated genes but not downregulated genes, demonstrating a direct function that is specific to synaptic gene repression. Both pan-synaptic and subtype specific synaptic genes are upregulated, indicating a broad role in repressing synaptic gene expression across neuronal subtypes. Interestingly, a recent study of dynamically accessible enhancers in nervous system development found that

only enhancers positively correlated with gene expression were cell specific[56]. Together with our findings, this suggests a regulatory logic in which negative regulators of synaptic gene expression act more broadly than positive regulators, which may be neuronal subtype specific.

A key question to be addressed in the future is how activation and repression of synaptic genes are coordinated. It is possible that positive regulators outcompete DEAF1 and CLAMP during activation. We identified several candidates TFs that may be involved in transcriptional activation of synaptic genes, including: (1) Trl (GAGA factor), a GA-binding TF similar to CLAMP that has been very well-studied[57–59] as a transcriptional activator and competes with CLAMP for GA-rich binding sites involved in dosage compensation[41] and histone gene regulation[39] and (2) Onecut, which was recently identified as a positive transcriptional regulator of pan-neuronal genes in *C. elegans*[60,61]. Alternatively, regulation of DEAF1 and/or CLAMP themselves could modulate the transition between negative and positive regulation of synaptic gene expression. DEAF1 can both positively and negatively regulate 5-HT_{1A} transcription, with GSK3-dependent phosphorylation promoting its repressive role[44,45]. Therefore, it is possible that DEAF1 remains at its binding sites during activation and its function is altered by phosphorylation state, a model that will be important to test in the future.

Finally, we demonstrate that repression of synaptic gene expression plays an instructive role in synapse formation. Targets of repression include pan-synaptic genes as well as cell-type specific neurotransmitter transporters and receptors indicating a broad role across cell types. The neurotransmitters released and interpreted by neurons are a central part of their identity and it has long been thought that neuronal identity is conferred by cell-specific combinations of TFs termed terminal selectors[1]. One interesting model that reconciles these two findings is that repression of synaptic gene expression before and after periods of synaptogenesis occurs broadly across cell types, while activation of synaptic gene expression relies on the action of cell-specific terminal selectors. This is consistent with studies in *C. elegans* demonstrating that pan-neuronal gene expression is regulated by a diversity of redundant factors[55]. Recent findings in the mouse cortex provide compelling support. There, a similar analysis of dynamic gene expression and chromatin accessibility in single cells revealed that positive regulators of neuronal maturation tend to function in a cell-type manner while negative regulators act more broadly[56]. Loss of either DEAF1 or CLAMP results in ectopic synapse formation, demonstrating the critical role of synaptic gene repression in proper circuit development. Conversely, neuronal overexpression of DEAF1 blocks synapse formation, indicating the importance of dynamic modulation of positive and negative regulatory factors across development. Both dominant and recessive pathogenic *DEAF1* alleles cause syndromic intellectual disability [31,62,63], underscoring the importance of decoding the regulatory logic of synaptic gene expression.

Materials and methods

Drosophila stocks:

The following fly lines used in this study were obtained from the Bloomington Drosophila Stock Center (BDSC, NIH P40OD018537): *w1118* (RRID:BDSC_5905), *UAS-Deaf1-RNAi* (RRID:BDSC_32512), *Deaf1*^{k3} (RRID:BDSC_26279), *Deaf1*^{S10B} (RRID:BDSC_91626), *Clamp*^{KG00473} (RRID:BDSC_14352), *UAS-GFP-RNAi* (RRID:BDSC_9330), *pros-Gal4* (RRID:BDSC_80572), *elav*^{C155}*-Gal4* (RRID:BDSC_458). Additional fly lines used in this study were obtained from The Zurich ORFeome Project (FlyORF)[64]: UAS-*Deaf1-3xHA* (RRID:Flybase_FBst0501296). *Drosophila melanogaster* stocks were raised on molasses food (Lab Express, Type R) in a 25°C incubator with controlled humidity and 12h light/dark cycle.

Expression pattern analysis

Transcriptomic data was downloaded via FlyBase[9,65] from Graveley et al., 2011, *Nature*. Data was separated into tissue specific or developmental timepoints. Developmental timepoints embryo 0-2h through pupal p15 were retained for finding the Spearman correlation with the average profile of 20 synaptic genes using cor() in the R stats package[66]. Tissue specific data was used to remove any genes not found to be expressed in CNS of P8 pupa or L3 larva. Random control gene sets were collected from genes below 0.80 Spearman correlation and expressed in CNS tissue with sample_n() in the dplyr package[67].

Clustering analysis

Kmeans clustering was performed with the R stats package. The number of clusters was determined from the gap statistic calculated with packages cluster[68] and factoextra[69]. Gene ontology analysis was performed with the R package gprofiler2[70] and sorted into parent categories with rrvgo[71].

Single-cell pattern analysis

Using Seurat[72], all nervous system annotated clusters were subset from 2 hour embryonic datasets[21], as Seurat objects, then merged. Larval brain, CNS, and VNC datasets[22] from each timepoint were loaded as count matrices, Seurat objects were created and filtered with min.cells = 3, min.features = 200, percent.mt < 5, then merged. Likewise for pupal visual system datasets[23]. Each merged object was normalized, 5000 variable genes were used for principle component analysis (PCA) and 50 PCs used to identify nearest neighbors and RunUMAP with clustering resolutions of 0.47 for embryo and 0.6 for larva and pupa. The R package scCustomize[73] was used with AddModuleScore to average the expression of all synaptic genes in our list for each cell, the score of each cell was averaged per timepoint.

Mouse pattern analysis

We selected all orthologs that ranked as high or moderate according to the DRSC Integrative Ortholog Prediction Tool[74], finding at least two orthologous mouse genes for each of the 20 *Drosophila* genes. We used these orthologs to generate a mouse ‘synaptic’ footprint for stages E10.5 through E12.5 in whole brain and E13.5 through P0 in forebrain and hindbrain. These data were then analyzed like the *Drosophila* analysis above.

DNase-seq analysis
DNase-seq bigwig files were downloaded from the Furlong laboratory website. Average profiles of accessibility at synaptic genes were made using deepTools computeMatrix with parameters 'computeMatrix reference-point --referencePoint TSS -a 1000 -b 1000 --binSize 100 --skipZeros --smartLabels --sortRegions descend' and plotProfile '--plotType se --yMin 0 --yMax 350''.

Protein expression data analysis
Protein expression data from Casas-Vila et al., 2017, Genome Research, supplemental table 1. The imputed.log2.LFQ.intensity data for each protein were averaged across the four samples at each timepoint. The data were then subset to synaptic proteins, scores were shifted by the minimum average imputed score and the score of each gene normalized to its maximum expression value.

Motif discovery
DNase-seq bigwig files were converted to bedgraph format with bigWigToBedGraph[75]. Peaks were called with MACS2[76] bdgpeakcall with a cutoff above the standard deviation and parameters '--max-gap 10 --min-length 150'. The neuronal 10-12 hours peak file was subset using BEDTools[77] 'subtract -A' to remove accessibility peaks intersecting those present in neuronal 8-10 hours. Likewise neuronal 10-12 hours peaks were removed from neuronal 8-10 hours peaks to use as background control regions in motif analysis. Peaks newly accessible in the neuronal 10-12 hours data were subset to those within 3kb of synaptic genes or within 3kb of three random control gene sets using bedtools 'window -w 3000 -u'. Homer[78] findMotifGenome.pl was used for motif analysis with the parameters 'dm3 -len 8,10,12 -size given -p 16 -dumpFasta'. The newly accessible neuronal 10-12 hours peak regions were used as input with the peaks unique to neuronal 8-10 hours data as background controls. A second analysis was done without inputting a background region set, defaulting to randomized control sequences. Additional analyses were performed with MEME Suite/5.5.5[79] XSTREME using default settings with Combined Drosophila motif databases and expected motif site distribution set to any number of repetitions. Also, MEME and TomTom were used with MEME parameters '-dna -objfun classic -mod anr -revcomp -nmotifs 21 -maxw 15' and TomTom with default parameters searching the Combined Drosophila motif databases plus JASPAR2018_CORE_insects databases.

Chromatin binding analysis
DNA binding data was downloaded from ENCODE or NCBI GEO databases. From the ENCODE database, files matching DNA binding assay type, *Drosophila melanogaster* organism, and available file types bed narrowPeak and bigWig were downloaded for analysis. For individual datasets from NCBI GEO, preprocessed bigwig, bedgraph, or peak bed files were downloaded for analysis. Average profiles were made from the bigWig files with deepTools computeMatrix reference-point either centered on newly accessible regions within 3kb of synaptic genes or random control gene sets or centered on the TSS of the same gene sets, with parameters '-a 3000 -b 3000 --skipZeros --smartLabels --sortRegions descend' and plotHeatmap with default settings. DNA binding peaks annotated to accessible regions within

3kb each gene set were identified with BEDTools[77] 'intersect -u -sortout', unique intersections were counted and percent of total accessible regions in each set was calculated.

Single-cell analysis

Single-cell data processing for NetREm input

Third instar larval brain scRNA-seq and scATAC-seq data[36] were processed with Seurat[72] and Signac[80], objects were created from counts matrix then filtered, normalized, and clustered using the parameters detailed in Mohana et al.[36]. Genes with highest average expression in each scRNA cluster (gene activity scores for scATAC), were used with enriched marker genes from FindAllMarkers(seurat_obj, only.pos = TRUE, min.pct = 0.25, logfc.threshold = 0.25, test.use = "wilcox") to broadly annotate the 16 RNA and 19 ATAC clusters, which were then merged by annotation.

NetREm

NetREm is a tool that uses TF and target gene (TG) expression to score TFs by how well they predict the TG expression levels. We applied NetREm according to the methods described in the publication[35], with the following specifications. We randomly selected 80% of the samples for training (M cells) and the remaining 20% for testing. We selected $\beta = 100$ for our network prior constraint to stress identifying groups of functionally coordinating TFs for regulating the TG, and emphasizing known direct and/or indirect PPIs among the TFs. We trained the model on the training gene expression data. The sparsity parameter α is selected using LassoCV with 5-fold cross-validation to optimize the mean square error on training data. We do not fit any y-intercept.

For NetREm TF input, we included 15 TFs with the highest percentage of DNA binding intersections in accessible peaks at synaptic genes plus all 141 TFs identified from motif analysis, with seven overlapping for a total of 149 TFs. Of these, 135 were identified in the 3rd instar larval brain scRNA-seq expression data. For target genes, we included all 485 genes from our synaptic genes list of which 457 were identified in the larval data. For PPI Network data, we used the full network of combined scores from *Drosophila melanogaster* STRINGdb version 12. STRINGdb v12 was subset to include 2,046 unique TF-TF edges of known direct and/or indirect PPIs, comprising 133 of the 135 default TFs examined in this analysis, E(spl)mbeta-HLH and E(spl)mgamma-HLH were not identified in the PPI data and given artificial edge weights of 0.01. For target genes accessibility regions, we identified all chromatin accessibility peaks in the 3rd instar larval brain scATAC-seq data that were found within 3kb of synaptic genes and subset them according to the four cell annotations. We performed motif analysis on these accessible regions using motifmatchr[81] matchMotifs function with the JASPAR2024 motif database and BSgenome.Dmelanogaster.UCSC.dm6 genome and output scores for each peak. We retained the top 50% TFs for each peak based on the highest scores. Only 96 of the 135 TFs were identified by JASPAR2024; thus, we queried additional motif databases using query(MotifDb[82], andStrings=c("Dmelanogaster"), orStrings=c("jaspar2024", "Dmelanogaster"), notStrings="ctcfl"). These additional motifs are drawn from Cisbp_1.02, FlyFactorSurvey, JASPAR_2014, JASPAR_CORE, Jaspar2016, Jaspar2018, and Jaspar2022. To identify the motifs in the accessibility sequences, we used monaLisa findMotifHits function with parameters

‘query=pwm_list, subject=seqs, min.score=”75”, method=”matchPWM”, BPPARAM=BiocParallel::SerialParam()’. We found 112 of the 135 TFs using MonaLisa (min. score percentile 75), with 91 TFs in common with the JASPAR2024 database for a total of 117 TFs with some accessible motif. For genes in cell types without identified accessible regions or motifs, we use the 135 default TFs, these include 131 in glia, 74 in neural progenitors, 140 in immature neurons, and 97 in mature neurons. We annotate coordination known PPI links as potentially physical interactions if they are found in the STRING v12 database for direct PPIs; otherwise, we surmise that these known PPIs may likely be via indirect mechanisms.

Embryo single-cell expression analysis

Embryo scRNA-seq data[21] was initially processed as described above in *Single-cell pattern analysis*, then annotated broadly based on genes with highest average expression per cluster and enriched marker genes. These data were then subset to only neural progenitors, immature neurons, and mature neurons from 06 to 18h timepoints. To compensate for dropout common to all single-cell experiments, we used cell-to-cell similarities based on Pearson correlations of normalized gene expression patterns, performed Network Enhancement[83,84] to denoise the resulting matrix, and multiplied by the original counts matrix. We evaluated the imputed results with cluster analysis, and preservation of marker gene specificity for developmental stages (Fig. S3C-I). Mature neuron cell numbers increase and neural progenitors decrease across the span of 10 to 18 hours. Therefore, to more precisely compare expression trends, we subset the time-points from 12 to 16 hours because they contain similar numbers of cells from each of the three annotations.

Imputation of scRNA-seq was done following the protocol recommended by Cheng et al.[84] using NE algorithm (...) with adaptations to impute scRNA-seq. The raw counts were first normalized by $log(\frac{10^5 * C_{ij}}{\sum_i C_{ij}} + 1)$ where C_{ij} is the count value of gene *i* in cell *j*. Then we calculated Pearson correlation between all cells to obtain cell-to-cell similarity matrix. We then apply network enhancement to improve the similarity matrix and the cell-to-cell correlations. The denoised matrix was divided by the maximum value of each cell for normalization while the self-weight was set to 1.5 times the maximum of its neighbors for each cell. The imputed counts were constructed by multiplying the original input matrix by the denoised weighting matrix.

Gene-to-gene Pearson correlation matrix was then obtained from the imputed count matrix.

scATAC-seq preprocessing with pyCisTopic

scATAC-seq embryo data[21] was subset from all nervous system annotated clusters in each 2h time point. These data were merged, log normalized, and processed with Signac[80] using default RunTFIDF, FindTopFeatures(min.cutoff = 'q1'), and RunSVD. Nearest neighbors were calculated from dimensions 2:50 with lsi reduction, and clustered with resolution 0.3. Any clusters with only 2 cells were merged with their visually closest cluster in the UMAP space. The resulting 16 clusters were annotated broadly from marker and average gene expression, then merged according to annotation. These data were then subset to only neural progenitors,

immature neurons, and mature neurons from 6 to 18h timepoints. Preprocessing was performed using pyCisTopic following the steps outlined in et al.[85]. Specifically, MACS2[76] was used for peak-calling with the following parameters: input_format - “BEDPE”, shift - 73, ext_size - 146, keep_dup - all, q_value - 0.05. This was done separately for each cell type using pseudobulk fragment bed files. Consensus peaks were identified for each cell type and across all cell types using an iterative approach in which overlapping peaks are filtered out based on significance. High quality cells were selected based on the following metrics: total number of unique fragments (>1x10^3.5), TSS enrichment score (> 2), and fraction of reads occurring in called peaks, FRiP (> 0.2) . Predicted doublets were removed using scrublet with expected_doublet_rate set to 0.1 and a doublet score threshold of 0.25. 31,389 high-quality cells remained after these filtering steps. For topic modeling we used the serial Latent Dirichlet Allocation (LDA) algorithm with a collapsed Gibbs sampler. This was done using the run_cgs_models() function from pyCisTopic with the following parameters: n_iter - 200, random_state - 285, alpha - 50, eta - 0.1. We ran topic modeling for each number of topics from 95 to 140 in intervals of 5, and selected 135 as the optimal number of topics using pyCisTopic’s evaluate_models() function which is based on previously published metrics[86–88]. The topic-region distribution was binarized using the binarize_topics() function from pyCisTopic with the method set to ‘otsu’, which converts topics into distinct region sets. Dropout imputation was performed using the impute_accessibility() function with the scale_factor set to 10^6. This step determines the probability of a region being accessible in a given cell by multiplying the region-topic (probability of a region belonging to a topic) and topic-cell (contribution of a topic within a cell) distributions. The imputed accessibility matrix was then used to identify Differentially Accessible Regions (DARs) through a Wilcoxon rank-sum test with the find_diff_features() function. Regions with a logFC > 1.5 and a Benjamini-Hochberg adjusted p-value < 0.05 were selected as being differentially accessible. After topic modeling and DAR selection, 43,613 regions were selected for downstream analysis.

Motif matching using motifmatchr

Regions within 3kb of a synaptic gene were identified using BEDtools ‘window -w 3000’, which yielded 3,128 regions. We generated a motif database using the MotifDb[82] R package as follows: query(MotifDb, andStrings = c("Dmelanogaster"), orStrings = c("jaspar2024", "Dmelanogaster"), notStrings = "ctcfl"). This call draws Position Frequency Matrices (PFMs) from Cisbp_1.02, FlyFactorSurvey, JASPAR_2014, JASPAR_CORE, Jaspar2016, Jaspar2018, and Jaspar2022, generating a database of 1437 motifs. PFMs were converted to objects of class PWMatrix from the TFBSTools package using the convert_motifs() function from the universalmotif package. Motif matching was performed using the matchMotifs() function from motifmatchr[81] with ‘genome’ set to ‘dm6’ and ‘out’ set to ‘positions’. This generates the coordinates of all hits for each motif in our database.

Polytene chromosome immunostaining, and analysis

Polytene chromosome squashes were prepared using the protocol described in Rieder et al.[39]. For immunostaining we used rabbit anti- CLAMP (1:500, SDIX) and guinea-pig anti- Deaf1 (1:500, gift from William McGinnis[89]) antibodies. For detection, we used all Alexa Fluor secondary antibodies against rabbit and guinea pig at a concentration of 1:500. All samples

were finally washed in PTX, stained with DAPI and mounted in ProLong Diamond (Thermo Fisher Scientific), and we visualized slides at 40× on a Zeiss Axioimager M1 Epifluorescence upright microscope with the AxioVision version 4.8.2 software.

Polytene images were analyzed using NIS Elements. Histone loci were identified using Clamp staining and using Profile Analysis Tool a 70um line was drawn from the center of the histone locus down the chromosome[90]. Line profile fluorescence intensity was measured including a neighborhood of ~1um on either side of the line for each channel.

RNA isolation, library preparation, and sequencing

Larvae were sexed and their brains (optic lobes and VNC) were dissected out in PBS in rapid succession. The larval brains (N=12) were then immediately placed in RNase-free 1.5ml tubes (containing 100µl of TRIzol lysis reagent, Invitrogen, USA) on ice. There were eight biological replicates (tubes) for each genotype, with each replicate/tube containing 12 brains. The larval brains in each tube were homogenized with a pestle. QIAGEN RNeasy Plus Universal Mini Kit was used for the RNA extraction and ZYMO RNA Clean & Concentrator-25 kit was used for further purification. The RNA was eluted and confirmed for purity and concentration via 260/280 and 260/230 ratios measured using a Thermo Scientific NanoDrop One^{C}. Illumina libraries were prepared and sequenced at Novogene Corporation for 150bp Poly-A selected mRNA and paired-end sequenced using an Illumina NovaSeq 6000, targeting at least 20 million reads per sample.

RNA-seq analysis

Raw sequencing data were run for quality check using FastQC/0.11.9. Adapter content was trimmed with Trim-galore/0.6.10 and Cutadapt/4.6. Hisat2/2.2.1[91] was used for paired-end read alignment to the *Drosophila* reference genome dmel-r6.42, and Samtools/1.16.1[92] flagstat and idxstats for alignment check. From Subread/2.0.2, featureCounts[93] was used to count reads at gene features with the following parameters: -p --countReadPairs -t gene -g gene_id -C -T 8. DESeq2/1.40.2[94] was used for differential gene expression analysis with ashr[95] to quantify false discovery rates (FDR). We consider genes with padj (FDR) values below 0.05 to be differentially expressed. GO analysis was performed with gProfiler2/0.2.2[70]. Jaccard index significance scores were calculated with the R package GeneOverlap/1.36.0[96].

Phenotypic analysis of synaptogenesis

Wandering third-instar larvae were sexed and dissected in ice cold Ca2+-free saline and fixed for 6 minutes in Bouin's Fixative (Electron Microscopy Sciences, Cat# 15990). Dissections were washed several times with 1X PBS and then permeabilized with 1X PBS containing 0.1% Triton-X (Fisher Scientific, Cat# BP151-100). Dissected larvae were then blocked either for 1 hour at room temperature or overnight at 4°C in 1X PBS containing 0.1% Triton-X and 1% BSA (Sigma-Aldrich, Cat# A7906). Dissections were incubated in primary antibodies overnight at 4°C and secondary antibodies for 2–4 hours at room temperature or overnight at 4°C. Dissected larvae were mounted in Vectashield (Vector Laboratories, Cat# H-1000-10). The following antibodies were used at the indicated concentrations: mouse anti-Bruchpilot (Brp) at 1:100 (DSHB, Cat# nc82, RRID: AB_2314866), Alexa Fluor 488 goat anti-mouse at 1:500 (Invitrogen,

Cat# A-11029, RRID: AB_2534088) and anti-Horseradish Peroxidase (HRP) conjugated to Alexa Fluor 647 at 1:500 (Jackson ImmunoResearch Labs, Cat# 123-605-021, RRID: AB_2338967). Larval NMJ images were taken on a Nikon A1R HD confocal microscope with a 40x oil immersion objective. Boutons were quantified at the highly stereotyped larval NMJ4 muscle segments, A2-A4. A bouton was defined as a bulbous swelling highlighted by the neuronal membrane marker HRP and co-stained with the active zone marker Brp. Quantification of bouton number was performed masked to genotype/condition using the “multi-point” tool in FIJI/ImageJ (NIH, RRID: SCR_003070).

To count individual active zones, Nikon Elements GA3 software was used to preprocess images with Guassian and Rolling ball filters and identify active zones using the Bright Spots module. ROIs were drawn using HRP and used as a mask to count Brp puncta. Quantification was performed blinded to experimental conditions.

Violin plots made with R package ggplot2[97] and ggstatsplot[98]. P-values were calculated with Welch's t-test for unequal variances, for parametric distributions, or Mann-Whitney U-test, for non-parametric distributions, or Kruskal-Wallis test with Dunn test and Benjamini-Hochberg p-value adjustment, for nonparametric multiple pairwise comparisons.

Data and materials availability

Data and materials generated in this study are available upon request from the corresponding authors. Transcriptomic data have been deposited in NCBI Gene Expression Omnibus (GEO) Database, under the accession number GSE280405.

Declaration of interests

The authors declare no competing interests.

Acknowledgements

We thank the Developmental Studies Hybridoma Bank and Bloomington Drosophila Stock Center (NIH P40OD018537) for antibodies and fly stocks. FlyBase[65] provided access to critical reagent and genomic information and databases. We are grateful to members of the Larschan and O’Connor-Giles labs for thoughtful discussions and comments on the manuscript. This work was supported by grants to KOG and EL from the National Institute of Neurological Disorders and Stroke, National Institutes of Health to KMOG (R21NS125864) and the Brown University Carney Institute for Brain Science, and from the NIH to the Brown MCB Program training grant T32GM136566.

References

1. Hobert, O., and Kratsios, P. (2019). Neuronal identity control by terminal selectors in worms, flies, and chordates. Curr Opin Neurobiol *56*, 97–105.
2. El-Danaf, R.N., Rajesh, R., and Desplan, C. (2023). Temporal regulation of neural diversity in Drosophila and vertebrates. Semin Cell Dev Biol *142*, 13–22.
3. Hamid, A., Gutierrez, A., Munroe, J., and Syed, M.H. (2023). The Drivers of Diversity: Integrated genetic and hormonal cues regulate neural diversity. Semin Cell Dev Biol *142*, 23–35.
4. Xing, Y., Zan, C., and Liu, L. (2023). Recent advances in understanding neuronal diversity and neural circuit complexity across different brain regions using single-cell sequencing. Front Neural Circuits *17*, 1007755.
5. Batool, S., Raza, H., Zaidi, J., Riaz, S., Hasan, S., and Syed, N.I. (2019). Synapse formation: from cellular and molecular mechanisms to neurodevelopmental and neurodegenerative disorders. J. Neurophysiol. *121*, 1381–1397. https://doi.org/10.1152/jn.00833.2018.
6. Wilhelm, B.G., Mandad, S., Truckenbrodt, S., Krohnert, K., Schafer, C., Rammner, B., Koo, S.J., Classen, G.A., Krauss, M., Haucke, V., et al. (2014). Composition of isolated synaptic boutons reveals the amounts of vesicle trafficking proteins. Science *344*, 1023–1028.
7. Gallegos, D.A., Chan, U., Chen, L.F., and West, A.E. (2018). Chromatin Regulation of Neuronal Maturation and Plasticity. Trends Neurosci *41*, 311–324.
8. Kishi, Y., and Gotoh, Y. (2018). Regulation of Chromatin Structure During Neural Development. Front Neurosci *12*, 874.
9. Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landolin, J.M., Yang, L., Artieri, C.G., van Baren, M.J., Boley, N., Booth, B.W., et al. (2011). The developmental transcriptome of Drosophila melanogaster. Nature *471*, 473–479.
10. Yoshihara, M., Ensminger, A.W., and Littleton, J.T. (2001). Neurobiology and the Drosophila genome. Funct Integr Genomics *1*, 235–240.
11. Emperador-Melero, J., and Kaeser, P.S. (2020). Assembly of the presynaptic active zone. Curr Opin Neurobiol *63*, 95–103.
12. Sanes, J.R., and Zipursky, S.L. (2020). Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits. Cell *181*, 1434–1435.
13. Duhart, J.C., and Mosca, T.J. (2022). Genetic regulation of central synapse formation and organization in Drosophila melanogaster. Genetics *221*.
14. Honjo, K., Mauthner, S.E., Wang, Y., Skene, J.H.P., and Tracey, W.D. (2016). Nociceptor-Enriched Genes Required for Normal Thermal Nociception. Cell Rep. *16*, 295–303. https://doi.org/10.1016/j.celrep.2016.06.003.
15. Bruckner, J.J., Gratz, S.J., Slind, J.K., Geske, R.R., Cummings, A.M., Galindo, S.E., Donohue, L.K., and O'Connor-Giles, K.M. (2012). Fife, a Drosophila Piccolo-RIM homolog, promotes active zone organization and neurotransmitter release. J Neurosci *32*, 17048–17058.
16. Bruckner, J.J., Zhan, H., Gratz, S.J., Rao, M., Ukken, F., Zilberg, G., and O'Connor-Giles, K.M. (2017). Fife organizes synaptic vesicles and calcium channels for high-probability neurotransmitter release. J Cell Biol *216*, 231–246.
17. Hogan, C.A., Gratz, S.J., Dumouchel, J.L., Thakur, R.S., Delgado, A., Lentini, J.M., Madhwani, K.R., Fu, D., and O'Connor-Giles, K.M. (2023). Expanded tRNA methyltransferase family member TRMT9B regulates synaptic growth and function. EMBO Rep *24*, e56808.
18. Thakur, R.S., and O'Connor-Giles, K.M. (2023). PDZD8 promotes autophagy at ER-Lysosome contact sites to regulate synaptogenesis. bioRxiv.
19. Frank, C.L., Liu, F., Wijayatunge, R., Song, L., Biegler, M.T., Yang, M.G., Vockley, C.M., Safi, A., Gersbach, C.A., Crawford, G.E., et al. (2015). Regulation of chromatin accessibility

and Zic binding at enhancers in the developing cerebellum. Nat. Neurosci. *18*, 647–656. https://doi.org/10.1038/nn.3995.
20. Bakken, T.E., Miller, J.A., Ding, S.-L., Sunkin, S.M., Smith, K.A., Ng, L., Szafer, A., Dalley, R.A., Royall, J.J., Lemon, T., et al. (2016). A comprehensive transcriptional map of primate brain development. Nature *535*, 367–375. https://doi.org/10.1038/nature18637.
21. Calderon, D., Blecher-Gonen, R., Huang, X., Secchia, S., Kentro, J., Daza, R.M., Martin, B., Dulja, A., Schaub, C., Trapnell, C., et al. (2022). The continuum of *Drosophila* embryonic development at single-cell resolution. Science *377*, eabn5800.
22. Corrales, M., Cocanougher, B.T., Kohn, A.B., Wittenbach, J.D., Long, X.S., Lemire, A., Cardona, A., Singer, R.H., Moroz, L.L., and Zlatic, M. (2022). A single-cell transcriptomic atlas of complete insect nervous systems across multiple life stages. Neural Dev *17*, 8.
23. Kurmangaliyev, Y.Z., Yoo, J., Valdes-Aleman, J., Sanfilippo, P., and Zipursky, S.L. (2020). Transcriptional Programs of Circuit Assembly in the Drosophila Visual System. Neuron *108*, 1045-1057 e6.
24. Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., Liechti, A., Ascencao, K., Rummel, C., Ovchinnikova, S., et al. (2019). Gene expression across mammalian organ development. Nature *571*, 505–509.
25. Reddington, J.P., Garfield, D.A., Sigalova, O.M., Karabacak Calviello, A., Marco-Ferreres, R., Girardot, C., Viales, R.R., Degner, J.F., Ohler, U., and Furlong, E.E.M. (2020). Lineage-Resolved Enhancer and Promoter Usage during a Time Course of Embryogenesis. Dev Cell *55*, 648-664.e9.
26. Secchia, S., Forneris, M., Heinen, T., Stegle, O., and Furlong, E.E.M. (2022). Simultaneous cellular and molecular phenotyping of embryonic mutants using single-cell regulatory trajectories. Dev. Cell *57*, 496-511.e8. https://doi.org/10.1016/j.devcel.2022.01.016.
27. Casas-Vila, N., Bluhm, A., Sayols, S., Dinges, N., Dejung, M., Altenhein, T., Kappei, D., Altenhein, B., Roignant, J.-Y., and Butter, F. (2017). The developmental proteome of Drosophila melanogaster. Genome Res *27*, 1273–1285.
28. Huggenvik, J.I., Michelson, R.J., Collard, M.W., Ziemba, A.J., Gurley, P., and Mowen, K.A. (1998). Characterization of a nuclear deformed epidermal autoregulatory factor-1 (DEAF-1)-related (NUDR) transcriptional regulator protein. Mol Endocrinol *12*, 1619–1639.
29. Michelson, R.J., Collard, M.W., Ziemba, A.J., Persinger, J., Bartholomew, B., and Huggenvik, J.I. (1999). Nuclear DEAF-1-related (NUDR) Protein Contains a Novel DNA Binding Domain and Represses Transcription of the Heterogeneous Nuclear Ribonucleoprotein A2/B1 Promoter*. J Biol Chem *274*, 30510–30519.
30. Hammonds, A.S., Bristow, C.A., Fisher, W.W., Weiszmann, R., Wu, S., Hartenstein, V., Kellis, M., Yu, B., Frise, E., and Celniker, S.E. (2013). Spatial expression of transcription factors in Drosophila embryonic organ development. Genome Biol *14*, R140.
31. Nabais Sa, M.J., Jensik, P.J., McGee, S.R., Parker, M.J., Lahiri, N., McNeil, E.P., Kroes, H.Y., Hagerman, R.J., Harrison, R.E., Montgomery, T., et al. (2019). De novo and biallelic DEAF1 variants cause a phenotypic spectrum. Genet Med *21*, 2059–2069.
32. McGee, S.R., Rajamanickam, S., Adhikari, S., Falayi, O.C., Wilson, T.A., Shayota, B.J., Cooley Coleman, J.A., Skinner, C., Caylor, R.C., Stevenson, R.E., et al. (2023). Expansion and mechanistic insights into de novo DEAF1 variants in DEAF1-associated neurodevelopmental disorders. Hum Mol Genet *32*, 386–401.
33. The modENCODE Consortium, Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., et al. (2010). Identification of Functional Elements and Regulatory Circuits by Drosophila modENCODE. Science *330*, 1787–1797. https://doi.org/10.1126/science.1198374.
34. Fisher, W.W., Hammonds, A.S., Weiszmann, R., Booth, B.W., Gevirtzman, L., Patton, J.E.J., Kubo, C.A., Waterston, R.H., and Celniker, S.E. (2023). A modERN resource: identification of Drosophila transcription factor candidate target genes using RNAi. Genetics *223*.

35. Khullar, S., Huang, X., Ramesh, R., Svaren, J., and Wang, D. (2024). NetREm Network Regression Embeddings reveal cell-type transcription factor coordination for gene regulation. bioRxiv, 2023.10.25.563769.
36. Mohana, G., Dorier, J., Li, X., Mouginot, M., Smith, R.C., Malek, H., Leleu, M., Rodriguez, D., Khadka, J., Rosa, P., et al. (2023). Chromosome-level organization of the regulatory genome in the Drosophila nervous system. Cell *186*, 3826-3844 e26.
37. Soruco, M.M., Chery, J., Bishop, E.P., Siggers, T., Tolstorukov, M.Y., Leydon, A.R., Sugden, A.U., Goebel, K., Feng, J., Xia, P., et al. (2013). The CLAMP protein links the MSL complex to the X chromosome during Drosophila dosage compensation. Genes Dev *27*, 1551–1556.
38. Ray, M., Conard, A.M., Urban, J., Mahableshwarkar, P., Aguilera, J., Huang, A., Vaidyanathan, S., and Larschan, E. (2023). Sex-specific splicing occurs genome-wide during early Drosophila embryogenesis. eLife *12*, e87865. https://doi.org/10.7554/eLife.87865.
39. Rieder, L.E., Koreski, K.P., Boltz, K.A., Kuzu, G., Urban, J.A., Bowman, S.K., Zeidman, A., Jordan, W.T., 3rd, Tolstorukov, M.Y., Marzluff, W.F., et al. (2017). Histone locus regulation by the Drosophila dosage compensation adaptor protein CLAMP. Genes Dev *31*, 1494–1508.
40. Urban, J.A., Doherty, C.A., Jordan, W.T., 3rd, Bliss, J.E., Feng, J., Soruco, M.M., Rieder, L.E., Tsiarli, M.A., and Larschan, E.N. (2017). The essential Drosophila CLAMP protein differentially regulates non-coding roX RNAs in male and females. Chromosome Res *25*, 101–113.
41. Kaye, E.G., Booker, M., Kurland, J.V., Conicella, A.E., Fawzi, N.L., Bulyk, M.L., Tolstorukov, M.Y., and Larschan, E. (2018). Differential Occupancy of Two GA-Binding Proteins Promotes Targeting of the Drosophila Dosage Compensation Complex to the Male X Chromosome. Cell Rep. *22*, 3227–3239. https://doi.org/10.1016/j.celrep.2018.02.098.
42. Bag, I., Dale, R.K., Palmer, C., and Lei, E.P. (2019). The zinc-finger protein CLAMP promotes gypsy chromatin insulator function in Drosophila. J Cell Sci *132*.
43. Kuzu, G., Kaye, E.G., Chery, J., Siggers, T., Yang, L., Dobson, J.R., Boor, S., Bliss, J., Liu, W., Jogl, G., et al. (2016). Expansion of GA Dinucleotide Repeats Increases the Density of CLAMP Binding Sites on the X-Chromosome to Promote Drosophila Dosage Compensation. PLoS Genet *12*, e1006120.
44. Czesak, M., Lemonde, S., Peterson, E.A., Rogaeva, A., and Albert, P.R. (2006). Cell-specific repressor or enhancer activities of Deaf-1 at a serotonin 1A receptor gene polymorphism. J Neurosci *26*, 1864–1871.
45. Harkin, E.F., Nasrallah, G., Le Francois, B., and Albert, P.R. (2023). Transcriptional Regulation of the Human 5-HT1A Receptor Gene by Lithium: Role of Deaf1 and GSK3beta. Int J Mol Sci *24*.
46. Tsiarli, M.A., Conard, A.M., Xu, L., Nguyen, E., and Larschan, E.N. (2020). The transcription factor CLAMP is required for neurogenesis in Drosophila melanogaster. Preprint at bioRxiv, https://doi.org/10.1101/2020.10.09.333831 https://doi.org/10.1101/2020.10.09.333831.
47. Haussmann, I.U., White, K., and Soller, M. (2008). Erect wing regulates synaptic growth in Drosophila by integration of multiple signaling pathways. Genome Biol. *9*, R73. https://doi.org/10.1186/gb-2008-9-4-r73.
48. Stergachis, A.B., Neph, S., Reynolds, A., Humbert, R., Miller, B., Paige, S.L., Vernot, B., Cheng, J.B., Thurman, R.E., Sandstrom, R., et al. (2013). Developmental Fate and Cellular Maturity Encoded in Human Regulatory DNA Landscapes. Cell *154*, 888–903. https://doi.org/10.1016/j.cell.2013.07.020.
49. Nord, A.S., Blow, M.J., Attanasio, C., Akiyama, J.A., Holt, A., Hosseini, R., Phouanenavong, S., Plajzer-Frick, I., Shoukry, M., Afzal, V., et al. (2013). Rapid and Pervasive Changes in Genome-wide Enhancer Usage during Mammalian Development. Cell *155*, 1521–1531. https://doi.org/10.1016/j.cell.2013.11.033.
50. Philippe, T.J., Vahid-Ansari, F., Donaldson, Z.R., Le François, B., Zahrai, A.,

Turcotte-Cardin, V., Daigle, M., James, J., Hen, R., Merali, Z., et al. (2018). Loss of MeCP2 in adult 5-HT neurons induces 5-HT1A autoreceptors, with opposite sex-dependent anxiety and depression phenotypes. Sci. Rep. *8*, 5788. https://doi.org/10.1038/s41598-018-24167-8.
51. Ruiz, J.L., Yerbanga, R.S., Lefèvre, T., Ouedraogo, J.B., Corces, V.G., and Gómez-Díaz, E. (2019). Chromatin changes in Anopheles gambiae induced by Plasmodium falciparum infection. Epigenetics Chromatin *12*, 5. https://doi.org/10.1186/s13072-018-0250-9.
52. Kratsios, P., Pinan-Lucarre, B., Kerk, S.Y., Weinreb, A., Bessereau, J.L., and Hobert, O. (2015). Transcriptional coordination of synaptogenesis and neurotransmitter signaling. Curr Biol *25*, 1282–1295.
53. Konstantinides, N., Kapuralin, K., Fadil, C., Barboza, L., Satija, R., and Desplan, C. (2018). Phenotypic Convergence: Distinct Transcription Factors Regulate Common Terminal Features. Cell *174*, 622-635 e13.
54. Haussmann, I.U., and Soller, M. (2010). Differential activity of EWG transcription factor isoforms identifies a subset of differentially regulated genes important for synaptic growth regulation. Dev Biol *348*, 224–230.
55. Stefanakis, N., Carrera, I., and Hobert, O. (2015). Regulatory Logic of Pan-Neuronal Gene Expression in C. elegans. Neuron *87*, 733–750.
56. Noack, F., Vangelisti, S., Raffl, G., Carido, M., Diwakar, J., Chong, F., and Bonev, B. (2022). Multimodal profiling of the transcriptional regulatory landscape of the developing mouse cortex identifies Neurog2 as a key epigenome remodeler. Nat. Neurosci. *25*, 154–167. https://doi.org/10.1038/s41593-021-01002-4.
57. Adkins, N.L., Hagerman, T.A., and Georgel, P. (2006). GAGA protein: a multi-faceted transcription factorThis paper is one of a selection of papers published in this Special Issue, entitled 27th International West Coast Chromatin and Chromosome Conference, and has undergone the Journal's usual peer review process. Biochem. Cell Biol. *84*, 559–558. https://doi.org/10.1139/o06-062.
58. Bayarmagnai, B., Nicolay, B.N., Islam, A.B.M.M.K., Lopez-Bigas, N., and Frolov, M.V. (2012). Drosophila GAGA factor is required for full activation of the dE2f1-Yki/Sd transcriptional program. Cell Cycle *11*, 4191–4202. https://doi.org/10.4161/cc.22486.
59. Li, X., Tang, X., Bing, X., Catalano, C., Li, T., Dolsten, G., Wu, C., and Levine, M. (2023). GAGA-associated factor fosters loop formation in the *Drosophila* genome. Mol. Cell *83*, 1519-1526.e4. https://doi.org/10.1016/j.molcel.2023.03.011.
60. Leyva-Diaz, E., and Hobert, O. (2022). Robust regulatory architecture of pan-neuronal gene expression. Curr Biol.
61. Leyva-Díaz, E. (2023). CUT homeobox genes: transcriptional regulation of neuronal specification and beyond. Front. Cell. Neurosci. *17*. https://doi.org/10.3389/fncel.2023.1233830.
62. Vulto-van Silfhout, A.T., Rajamanickam, S., Jensik, P.J., Vergult, S., de Rocker, N., Newhall, K.J., Raghavan, R., Reardon, S.N., Jarrett, K., McIntyre, T., et al. (2014). Mutations affecting the SAND domain of DEAF1 cause intellectual disability with severe speech impairment and behavioral problems. Am J Hum Genet *94*, 649–661.
63. Chen, L., Jensik, P.J., Alaimo, J.T., Walkiewicz, M., Berger, S., Roeder, E., Faqeih, E.A., Bernstein, J.A., Smith, A.C.M., Mullegama, S.V., et al. (2017). Functional analysis of novel DEAF1 variants identified through clinical exome sequencing expands DEAF1-associated neurodevelopmental disorder (DAND) phenotype. Hum Mutat *38*, 1774–1785.
64. Bischof, J., Björklund, M., Furger, E., Schertel, C., Taipale, J., and Basler, K. (2013). A versatile platform for creating a comprehensive UAS-ORFeome library in Drosophila. Development *140*, 2434–2442. https://doi.org/10.1242/dev.088757.
65. Öztürk-Çolak, A., Marygold, S.J., Antonazzo, G., Attrill, H., Goutte-Gattat, D., Jenkins, V.K., Matthews, B.B., Millburn, G., Dos Santos, G., Tabone, C.J., et al. (2024). FlyBase: updates

to the Drosophila genes and genomes database. Genetics *227*.
66. R Core Team (2023). R: A Language and Environment for Statistical Computing.
67. Wickham, H., François, R., Henry, L., Müller, K., and Vaughan, D. (2023). dplyr: A Grammar of Data Manipulation.
68. Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., and Hornik, K. (2023). cluster: Cluster Analysis Basics and Extensions.
69. Kassambara, A., and Mundt, F. (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses.
70. Kolberg, L., Raudvere, U., Kuzmin, I., Adler, P., Vilo, J., and Peterson, H. (2023). g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Res *51*, W207–W212.
71. Sayols, S. (2023). rrvgo: a Bioconductor package to reduce and visualize Gene Ontology terms. MicroPublication Biol. https://doi.org/10.17912/micropub.biology.000811.
72. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573-3587 e29.
73. Marsh, S., Salmon, M., and Hoffman, P. (2024). samuel-marsh/scCustomize: Version 2.1.2. Version v2.1.2 (Zenodo). https://doi.org/10.5281/zenodo.10724532 https://doi.org/10.5281/zenodo.10724532.
74. Hu, Y., Comjean, A., Rodiger, J., Liu, Y., Gao, Y., Chung, V., Zirin, J., Perrimon, N., and Mohr, S.E. (2021). FlyRNAi.org-the database of the Drosophila RNAi screening center and transgenic RNAi project: 2021 update. Nucleic Acids Res *49*, D908–D915.
75. Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S., and Karolchik, D. (2010). BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics *26*, 2204–2207. https://doi.org/10.1093/bioinformatics/btq351.
76. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137. https://doi.org/10.1186/gb-2008-9-9-r137.
77. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.
78. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576–589.
79. Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME Suite. Nucleic Acids Res *43*, W39-49.
80. Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R. (2021). Single-cell chromatin state analysis with Signac. Nat. Methods *18*, 1333–1341. https://doi.org/10.1038/s41592-021-01282-5.
81. Schep, A. (2024). motifmatchr: Fast Motif Matching in R.
82. Shannon P, R.M. (2024). MotifDb: An Annotated Collection of Protein-DNA Binding Sequence Motifs.
83. Wang, B., Pourshafeie, A., Zitnik, M., Zhu, J., Bustamante, C.D., Batzoglou, S., and Leskovec, J. (2018). Network enhancement as a general method to denoise weighted biological networks. Nat Commun *9*, 3108.
84. Cheng, Y., Ma, X., Yuan, L., Sun, Z., and Wang, P. (2023). Evaluating imputation methods for single-cell RNA-seq data. BMC Bioinformatics *24*, 302.
85. Bravo Gonzalez-Blas, C., De Winter, S., Hulselmans, G., Hecker, N., Matetovici, I., Christiaens, V., Poovathingal, S., Wouters, J., Aibar, S., and Aerts, S. (2023). SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. Nat Methods *20*,

1355–1367.
86. Mimno, D., Wallach, H., Talley, E., Leenders, M., and McCallum, A. (2011). Optimizing semantic coherence in topic models. In Proceedings of the 2011 conference on empirical methods in natural language processing, pp. 262–272.
87. Arun, R., Suresh, V., Veni Madhavan, C.E., and Narasimha Murthy, M.N. (2010). On Finding the Natural Number of Topics with Latent Dirichlet Allocation: Some Observations. In Advances in Knowledge Discovery and Data Mining, M. J. Zaki, J. X. Yu, B. Ravindran, and V. Pudi, eds. (Springer), pp. 391–402. https://doi.org/10.1007/978-3-642-13657-3_43.
88. Cao, J., Xia, T., Li, J., Zhang, Y., and Tang, S. (2009). A density-based method for adaptive LDA model selection. Neurocomputing *72*, 1775–1781. https://doi.org/10.1016/j.neucom.2008.06.011.
89. Veraksa, A., Kennison, J., and McGinnis, W. (2002). DEAF-1 function is essential for the early embryonic development of Drosophila. Genesis *33*, 67–76.
90. Amankwaa, B., Schoborg, T., and Labrador, M. (2022). Drosophila insulator proteins exhibit in vivo liquid–liquid phase separation properties. Life Sci. Alliance *5*. https://doi.org/10.26508/lsa.202201536.
91. Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol *37*, 907–915.
92. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. Gigascience *10*.
93. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.
94. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.
95. Stephens, M. (2017). False discovery rates: a new deal. Biostatistics *18*, 275–294.
96. Shen, L., and Sinai, I.S. of M. at M. (2023). GeneOverlap: Test and visualize gene overlaps https://doi.org/10.18129/B9.bioc.GeneOverlap.
97. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag New York).
98. Patil, I. (2021). Visualizations with statistical details: The “ggstatsplot” approach. J. Open Source Softw. *6*, 3167. https://doi.org/10.21105/joss.03167.

Figure Legends

Figure 1. Broad temporal coordination of synaptic gene transcription in *Drosophila* and mouse

(A) Average expression of 485 genes highly correlated with average synaptic expression pattern (blue) shows defined periods of up and down regulation during development, distinct from genes that regulate other neurodevelopmental processes including neurite development (orange), or a random set of nervous system expressed genes (grey). (B) GO analysis of genes highly correlated with average synaptic expression profile reveals enrichment for synaptic signaling, organization, and ion transport terms. Average expression scores of synaptic gene list across 3 temporal scRNA-seq datasets displayed as feature plots (C), and averaged per time point (D), shows a pattern closely matching the one from bulk data in (A). (E) Average profile of 742 mouse genes highly correlated with the average expression of mouse orthologs of our Drosophila synaptic 'footprint' genes shows initial upregulation around embryonic day 12. (F) GO analysis of the highly correlated mouse genes reveals high enrichment of synaptic signaling and ion transport terms matching Drosophila GO terms in (B). Shading around average lines represents the margin of the standard error. GO terms colored-coded based on shared hierarchical parent GO terms.

Figure S1. Temporally coordinated gene expression and clustering analysis

(A) Individual expression profiles of 20 well-characterized synaptic genes that were averaged for correlation analysis with all genes (Fig. 1A, blue). (B) Genes with expression values below a Spearman correlation cutoff of 0.80 show less enrichment for synaptic GO terms. (C) Individual expression profiles of 20 well-characterized genes involved with neurite development and axon guidance that were averaged for correlation analysis with all genes (Fig. 1A, orange). (D) GO analysis of 485 genes highly correlated with 'footprint' genes from (C) show enrichment for earlier neurodevelopmental terms. (E) Unbiased k-means clustering analysis identifies a pattern closely matching the synaptic pattern (cluster 8, light blue). (F) Distribution of the expression of synaptic genes overlaps mostly with cluster 8, and partially with cluster 1. (G) GO analysis of cluster 8 reveals enrichment for the same synaptic terms as highly correlated synaptic genes (Fig. 1B). GO terms colored-coded based on shared hierarchical parent GO terms.

Figure 2. DEAF1 motifs and binding are enriched at dynamically regulated accessible regions near synaptic genes

(A) Average profiles of DNase-seq data[25] surrounding TSSs of synaptic genes (aqua) reveals substantially increasing accessibility at 10-12 hours compared with random genes (grey). (B) The substantial increase in chromatin accessibility at synaptic gene TSS regions (aqua) precedes upregulation of synaptic transcripts (blue). (C) Motif analysis with Homer[78] of newly accessible regions within 3kb of synaptic genes at 10-12 hours reveals the DEAF1 binding motif as the most enriched. (D) Analysis of ChIP-seq data[33,34] identifies DEAF1 as the most enriched TF binding at newly accessible regions near synaptic genes (blue), compared with random genes (grey). Shading around average lines represents the margin of the standard error.

Figure S2. Synaptic genes have temporally coordinated chromatin accessibility and are bound by DEAF1.
(A) Synaptic gene locations are dispersed broadly across all chromosomes. (B) Average scores of 19 synaptic genes show DNase accessibility (aqua) preceding transcript (blue) and protein expression (purple). (C) Chromatin accessibility at synaptic genes TSSs (aqua) is temporally distinct from neurite development genes (orange) and is specific to neuronal tissue (top row) while little changed in mesoderm tissue (bottom row).

Figure 3. DEAF1 and CLAMP are predicted to co-repress synaptic gene expression
Input for NetREm[35] analysis: third instar larval brain scRNA-seq (A) and scATAC-seq (B) data[36] clustered and annotated to broad developmental stages. (C) TF coordination network centered on DEAF1, showing number of predicted synaptic target genes (ellipse size), predicted positive (red) to negative (dark blue) effect ratio on target genes, input experimental physical interaction scores (line width), calculated average weighted correlation score (line color intensity). (D) Embryonic temporal scRNA-seq data[21] as input for expression correlation analysis (E) of embryonic scRNA-seq data between DEAF1/CLAMP and all genes (top) or synaptic genes (bottom). (F) Embryonic scATAC-seq data[21] as input for co-accessibility analysis of (G) the number of DEAF1 and CLAMP binding motifs at synaptic genes. Co-accessible motifs for DEAF1 (gold) and CLAMP (red) at synaptic genes *brp* (H), and *Nlg4* (I).

Figure S3. DEAF1 is predicted to regulate most synaptic genes
(A) Additional motif analysis using MEME Suite[79] identified some additional binding motifs and all matching TFs were included as input for NetREm[35]. (B) Top 10 TFs predicted by NetREm to regulate the most synaptic genes, showing number of predicted synaptic target genes (ellipse size), predicted positive (red) to negative (dark blue) effect ratio on target genes, input experimental physical interaction scores (line width), calculated average weighted correlation score (line color intensity). Embryonic scRNA-seq data[21] UMAP prior to imputation (C) and following imputation (D). Expression of selected marker genes prior to (H) and following (I) imputations. Co-accessible motifs for DEAF1 (gold) and CLAMP (red) at synaptic genes futsch (J), and *mAChR-A* (K). Neural progenitors (N.P.), Immature neurons (I.N.), Mature neurons (M.N.)

Figure 4. DEAF1 and CLAMP colocalize at synaptic genes
(A) CLAMP and DEAF1 binding on polytene chromosomes displays widespread binding, and colocalization at a subset of sites. (B) Motif accessibility (top) and ChIP-seq binding peaks (bottom) for DEAF1 and CLAMP at synaptic genes (left) are enriched, and colocalized sites are more enriched than for either TF individually. No enrichment is found at random genes (right) (shaded by Jaccard index score). (C) ChIP-seq[34,43] peaks subset for 3kb around synaptic genes (blue). DEAF1 (gold) binds at 75% of synaptic genes and 85% of the synaptic genes that are also bound by CLAMP (red).

Figure S4. *deaf1* and *clamp* RNAi cause significant loss of DEAF1 and CLAMP protein on chromatin

Polytene chromosome spreads with control *gfp*RNAi (A), *deaf1*RNAi (B), and *clamp*RNAi (C). Close up images (right) show protein in a region that was traced for fluorescence intensity, which is quantified at each site along the trace, below each genotype. Total fluorescence along the trace for each sample in each genotype is quantified for DAPI (D), DEAF1 (E), and CLAMP (F). Each intensity score is normalized to the mean value in *gfp*RNAi controls.

Figure 5. Loss of DEAF1 or CLAMP derepresses synaptic gene expression
GO analysis, showing terms with the highest percentage of genes (recall) upregulated with *deaf1*-RNAi (A) or *clamp*-RNAi (C). (B) Overlap of synaptic genes (blue) and genes upregulated with *deaf1*-RNAi (gold) or *clamp*-RNAi (red). (D) Synaptic genes (left) are enriched for genes that are upregulated after either *deaf1*-RNAi or *clamp*-RNAi, and are more enriched within the set of common genes upregulated in both conditions. No enrichment is found for random neuronally-expressed genes (right) or for synaptic genes in downregulated gene sets (shaded by Jaccard index score). GO terms colored-coded based on shared hierarchical parent GO terms.

Figure S5. DEAF1 and CLAMP directly regulate repression of synaptic genes
(A) GO analysis, showing terms with the highest percentage of genes (recall) downregulated upon *deaf1*-RNAi (A) or *clamp*-RNAi (C). (B) Synaptic genes (blue) have little overlap with downregulated gene sets with *deaf1*-RNAi (gold) or *clamp*-RNAi (red), and no overlap with genes downregulated in both conditions. (D) ChIP-seq binding peaks for DEAF1 and CLAMP (columns) have significant overlap with genes that are upregulated after RNAi conditions but not with those downregulated after RNAi (shaded by Jaccard index score). GO terms colored-coded based on shared hierarchical parent GO terms.

Figure 6. DEAF1 and CLAMP regulate synapse formation.
(A-E) Representative images (A-C) of NMJ 4 in the indicated genotypes labeled with the neuronal membrane marker HRP (magenta) and active zone marker Brp (green) and quantification of synaptic bouton number (D) and active zone number (E). (F-H) Representative images (F, G) of NMJ 4 in the indicated genotypes labeled with the neuronal membrane marker HRP (magenta) and active zone marker Brp (green) and quantification of synaptic bouton number (H). Scale bar: 10 μm. Data points represent individual NMJs. Kruskal-Wallis test followed by Dunn multiple comparisons test and Benjamini-Hochberg p-value adjustment. Data is represented as mean ± SEM.

Figure S6. DEAF1 and CLAMP mutations cause dysregulation of synapse formation
Representative images (A-D) of NMJ 4 in the indicated genotypes labeled with the neuronal membrane marker HRP (magenta) and active zone marker Brp (green) and quantification of synaptic bouton number (F). Scale bar: 10 μm. Data points represent individual NMJs. Welch's t-test for unequal variances. Data is represented as mean ± SEM.

Figure 1

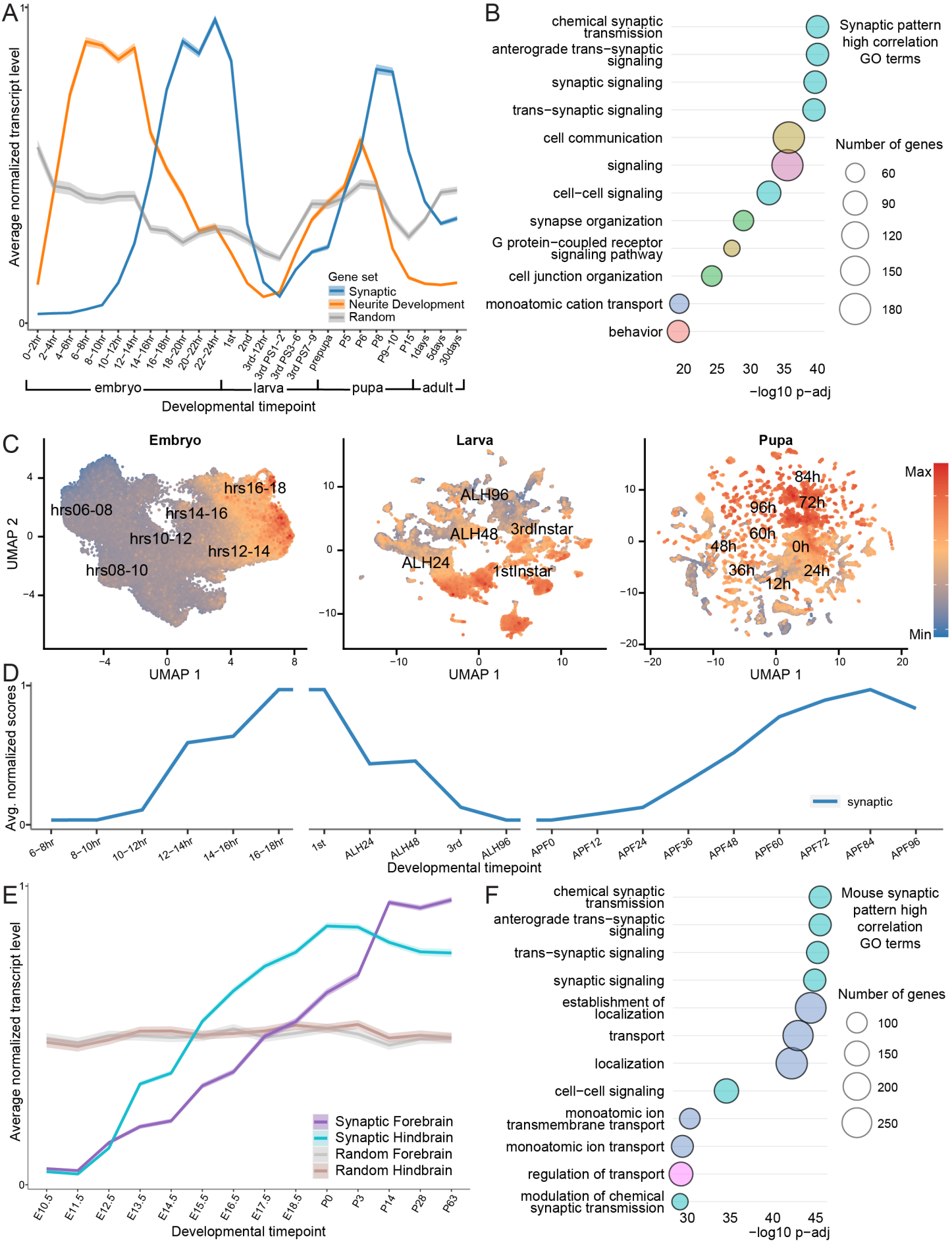

A
Average normalized transcript level
Gene set
Synaptic
Neurite Development
Random
0–2hr
2–4hr
4–6hr
6–8hr
8–10hr
10–12hr
12–14hr
14–16hr
16–18hr
18–20hr
20–22hr
22–24hr
1st
2nd
3rd-12hr
3rd PS1–2
3rd PS3–6
3rd PS7–9
prepupa
P5
P6
P8
P9–10
P15
1days
5days
30days
embryo
larva
pupa
adult
Developmental timepoint
B
chemical synaptic transmission
anterograde trans−synaptic signaling
synaptic signaling
trans−synaptic signaling
cell communication
signaling
cell−cell signaling
synapse organization
G protein−coupled receptor signaling pathway
cell junction organization
monoatomic cation transport
behavior
Synaptic pattern high correlation GO terms
Number of genes
60
90
120
150
180
20
25
30
35
40
−log10 p−adj
C
Embryo
Larva
Pupa
UMAP 2
UMAP 1
hrs06-08
hrs14-16
hrs16-18
hrs10-12
hrs12-14
hrs08-10
ALH96
ALH48
3rdInstar
ALH24
1stInstar
84h
96h
72h
60h
48h
0h
36h
24h
12h
Max
Min
D
Avg. normalized scores
synaptic
6–8hr
8–10hr
10–12hr
12–14hr
14–16hr
16–18hr
1st
ALH24
ALH48
3rd
ALH96
APF0
APF12
APF24
APF36
APF48
APF60
APF72
APF84
APF96
Developmental timepoint
E
Average normalized transcript level
Synaptic Forebrain
Synaptic Hindbrain
Random Forebrain
Random Hindbrain
E10.5
E11.5
E12.5
E13.5
E14.5
E15.5
E16.5
E17.5
E18.5
P0
P3
P14
P28
P63
Developmental timepoint
F
chemical synaptic transmission
anterograde trans−synaptic signaling
trans−synaptic signaling
synaptic signaling
establishment of localization
transport
localization
cell−cell signaling
monoatomic ion transmembrane transport
monoatomic ion transport
regulation of transport
modulation of chemical synaptic transmission
Mouse synaptic pattern high correlation GO terms
Number of genes
100
150
200
250
30
35
40
45
−log10 p−adj

Figure S1

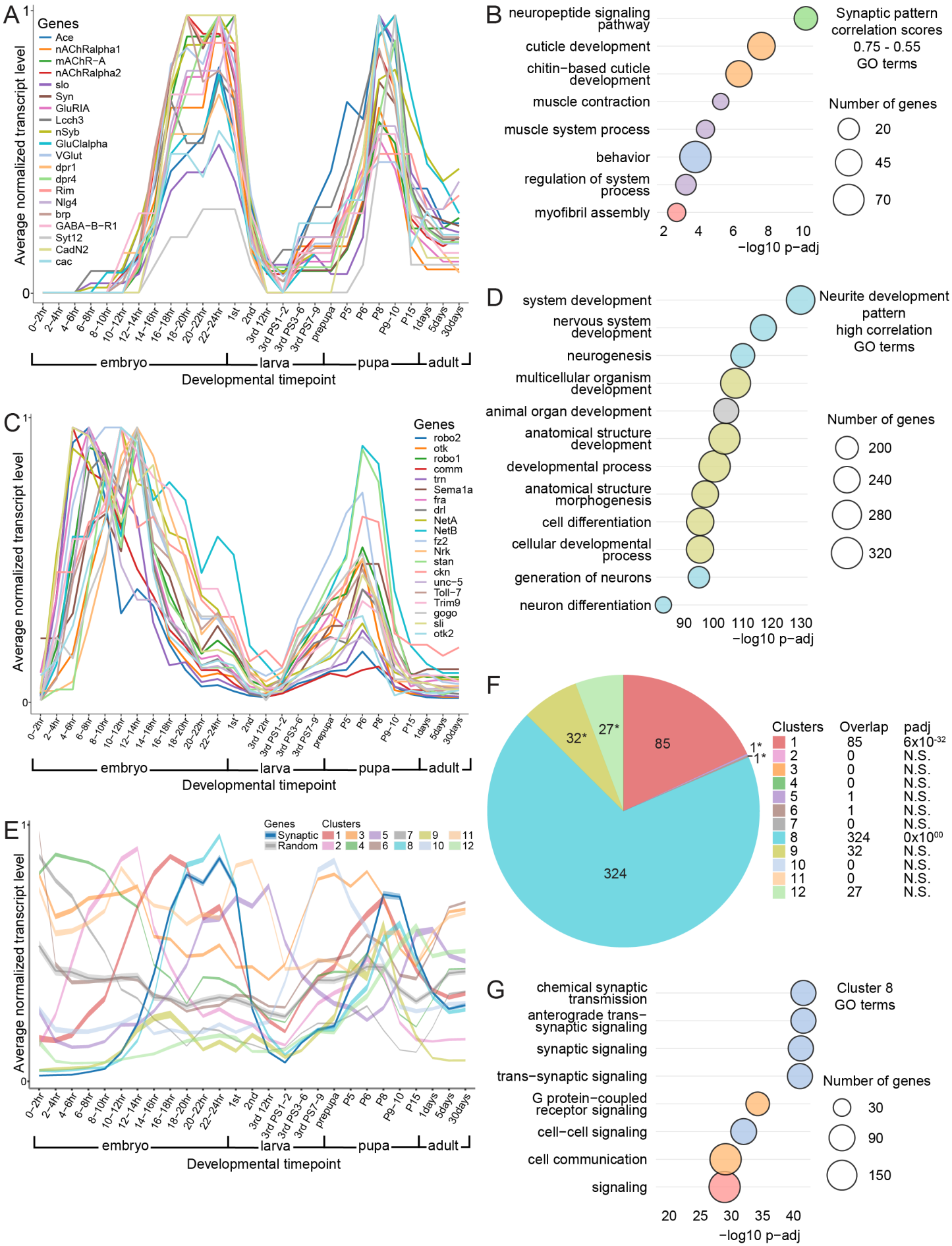
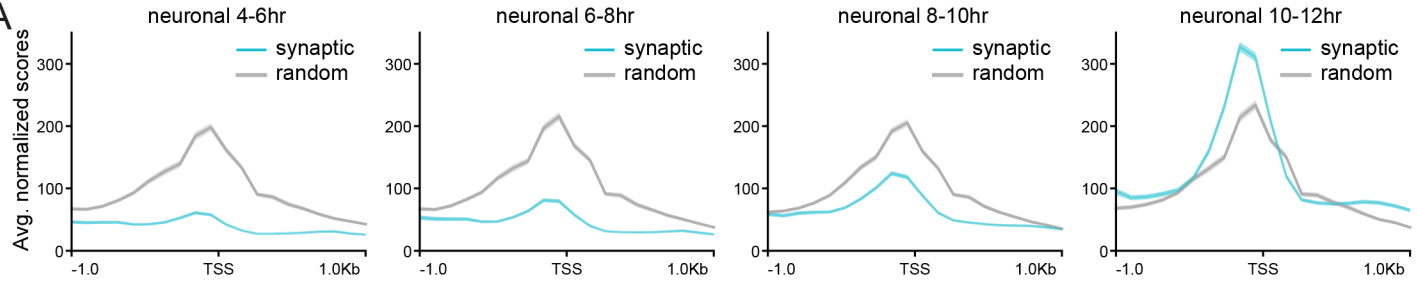
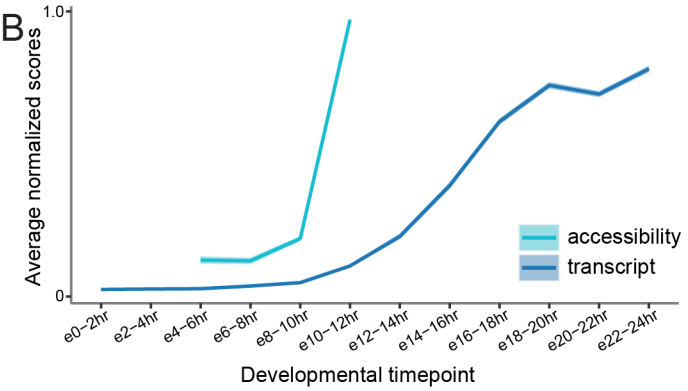

A
Genes
Ace
nAChRalpha1
mAChR-A
nAChRalpha2
slo
Syn
GluRIA
Lcch3
nSyb
GluClalpha
VGlut
dpr1
dpr4
Rim
Nlg4
brp
GABA-B-R1
Syt12
CadN2
cac
Average normalized transcript level
0–2hr
2–4hr
4–6hr
6–8hr
8–10hr
10–12hr
12–14hr
14–16hr
16–18hr
18–20hr
20–22hr
22–24hr
1st
2nd
3rd 12hr
3rd PS1–2
3rd PS3–6
3rd PS7–9
prepupa
P5
P6
P8
P9–10
P15
1days
5days
30days
embryo
larva
pupa
adult
Developmental timepoint
B
neuropeptide signaling pathway
cuticle development
chitin-based cuticle development
muscle contraction
muscle system process
behavior
regulation of system process
myofibril assembly
−log10 p−adj
Synaptic pattern correlation scores
0.75 - 0.55
GO terms
Number of genes
20
45
70
C
Genes
robo2
otk
robo1
comm
trn
Sema1a
fra
drl
NetA
NetB
fz2
Nrk
stan
ckn
unc-5
Toll-7
Trim9
gogo
sli
otk2
Average normalized transcript level
Developmental timepoint
D
system development
nervous system development
neurogenesis
multicellular organism development
animal organ development
anatomical structure development
developmental process
anatomical structure morphogenesis
cell differentiation
cellular developmental process
generation of neurons
neuron differentiation
−log10 p−adj
Neurite development pattern
high correlation
GO terms
Number of genes
200
240
280
320
E
Genes
Synaptic
Random
Clusters
Average normalized transcript level
Developmental timepoint
F
85
324
32*
27*
1*
1*
Clusters
Overlap
padj
1 85 6x10-32
2 0 N.S.
3 0 N.S.
4 0 N.S.
5 1 N.S.
6 1 N.S.
7 0 N.S.
8 324 0x100
9 32 N.S.
10 0 N.S.
11 0 N.S.
12 27 N.S.
G
chemical synaptic transmission
anterograde trans-synaptic signaling
synaptic signaling
trans-synaptic signaling
G protein-coupled receptor signaling
cell-cell signaling
cell communication
signaling
−log10 p−adj
Cluster 8
GO terms
Number of genes
30
90
150

Figure 2

A



B



C

Motif	P-value	% targets	% background	Best match
AGCTTTCG	$1x10^{-427}$	61.74 %	6.39 %	Deaf1
TTTCTTTT	$1x10^{-231}$	53.18 %	7.01 %	pan
ATTTTGCA	$1x10^{-222}$	41.51 %	6.51 %	byn
GCGCAGCC	$1x10^{-200}$	49.34 %	9.82 %	Med
CCTGCCCA	$1x10^{-179}$	36.48 %	2.78 %	pho
CCACCCCCCC	$1x10^{-178}$	36.34 %	5.19 %	run
AGAAATAA	$1x10^{-166}$	34.86 %	4.99 %	Mef2
ATTCGCTT	$1x10^{-161}$	34.27 %	6.49 %	Blimp-1
CACAGGAAAA	$1x10^{-154}$	42.98 %	9.11 %	pnt
CCCGTCAG	$1x10^{-144}$	77.84 %	30.69%	hth
GAGAGCCA	$1x10^{-140}$	40.92 %	7.19 %	GAF/Clamp
TTGTCAAC	$1x10^{-127}$	52.29 %	15.90 %	bin
TTTTCGGA	$1x10^{-121}$	37.96 %	8.09 %	kni
AGTCTGGTAT	$1x10^{-103}$	35.16 %	7.60 %	Hsf
CATGCATG	$1x10^{-94}$	60.56 %	23.95 %	vvl
TTTTAACATT	$1x10^{-84}$	38.40 %	12.67 %	CG11617
TACATACATA	$1x10^{-74}$	29.84 %	7.86 %	Top2
TGGATGCTGT	$1x10^{-64}$	49.34 %	22.01 %	odd
CAGCCAAC	$1x10^{-60}$	33.38 %	10.91 %	pho
ACAACAAA	$1x10^{-52}$	41.80 %	17.26 %	hb
GCTAATCGGT	$1x10^{-28}$	77.40 %	57.00 %	B-H1
CAGCAGCGGC	$1x10^{-18}$	81.98 %	68.18 %	Adf1

D

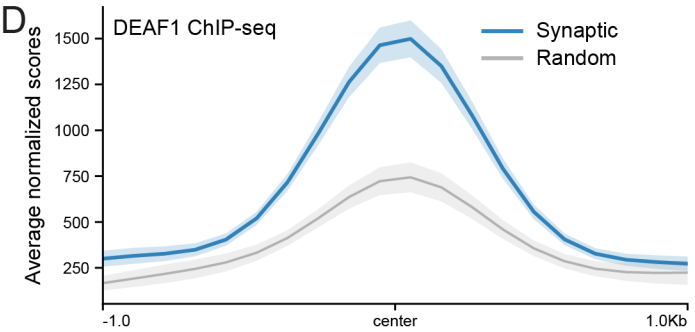


Figure S2

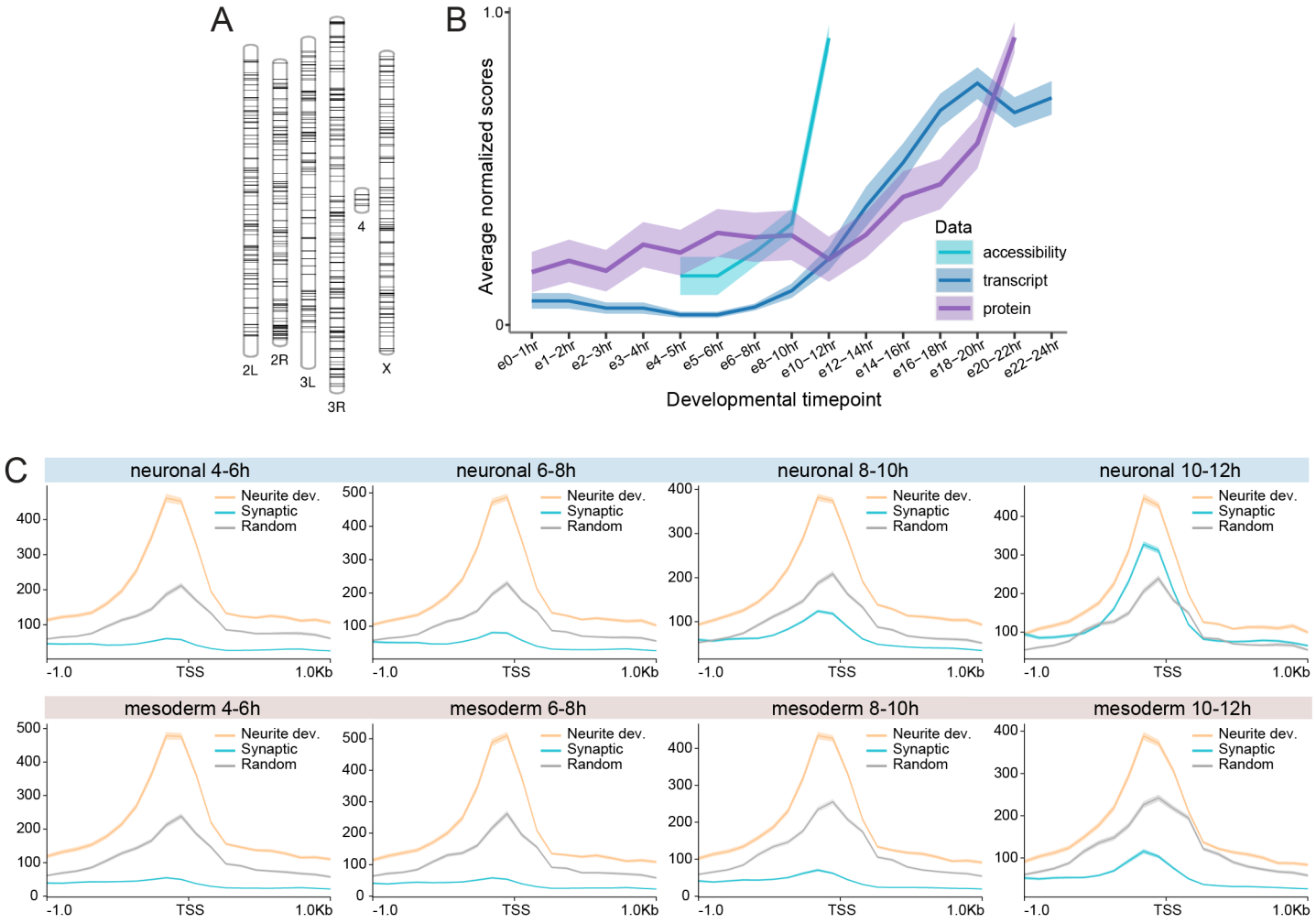

A
2L
2R
3L
3R
4
X
B
Average normalized scores
1.0
0
Data
accessibility
transcript
protein
e0–1hr
e1–2hr
e2–3hr
e3–4hr
e4–5hr
e5–6hr
e6–8hr
e8–10hr
e10–12hr
e12–14hr
e14–16hr
e16–18hr
e18–20hr
e20–22hr
e22–24hr
Developmental timepoint
C
neuronal 4-6h
neuronal 6-8h
neuronal 8-10h
neuronal 10-12h
mesoderm 4-6h
mesoderm 6-8h
mesoderm 8-10h
mesoderm 10-12h
Neurite dev.
Synaptic
Random
-1.0
TSS
1.0Kb

Figure 3

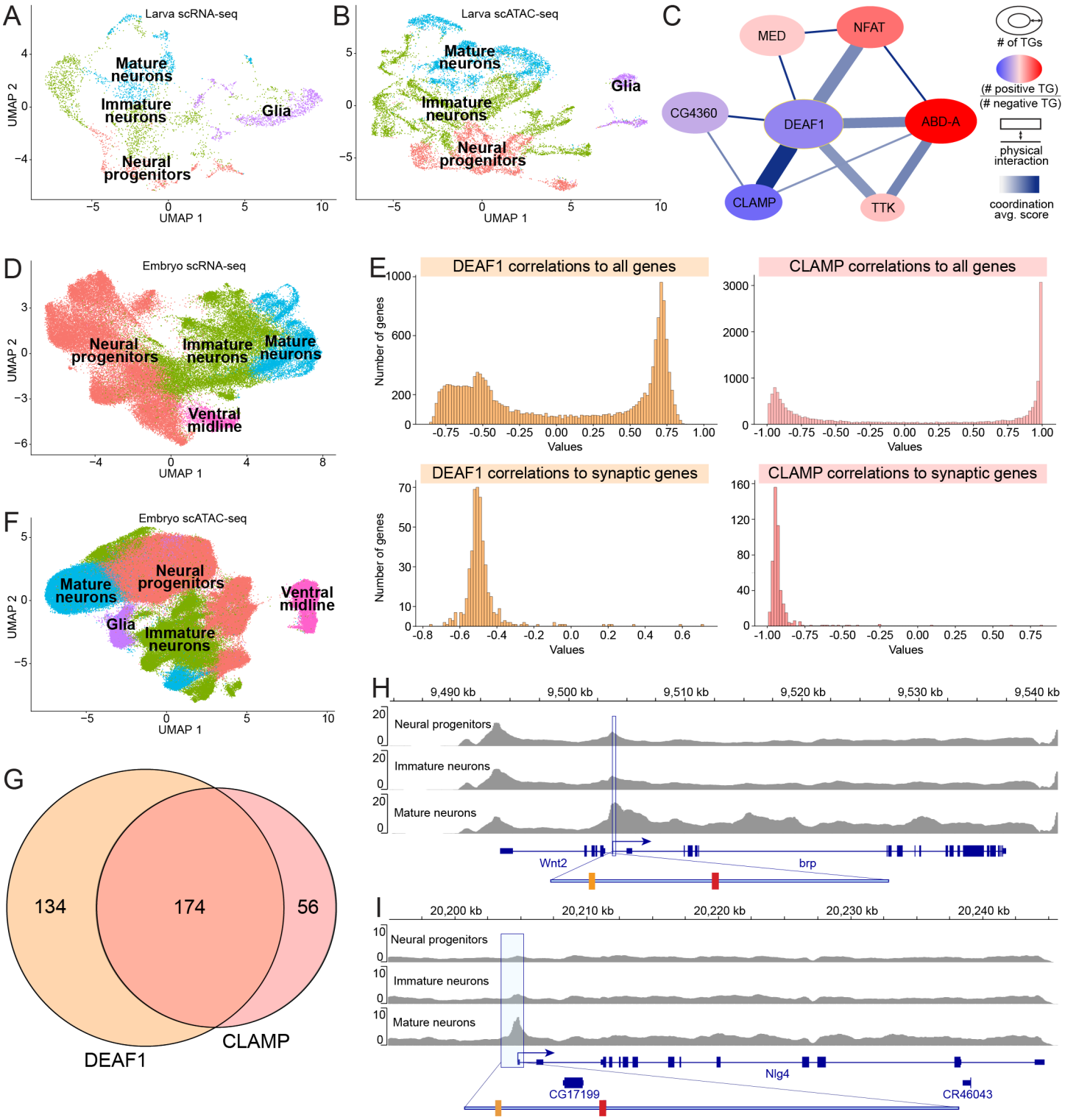
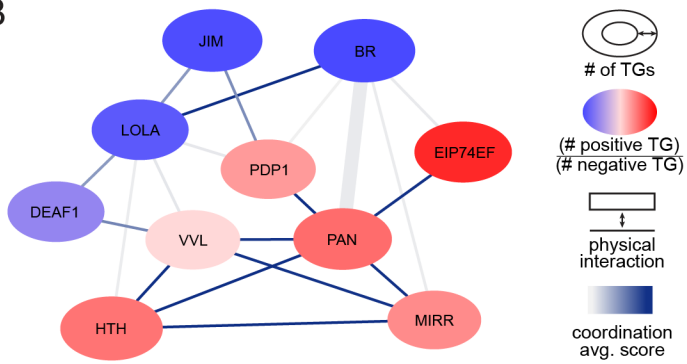

A
Larva scRNA-seq
Mature neurons
Immature neurons
Glia
Neural progenitors
UMAP 1
UMAP 2
B
Larva scATAC-seq
Mature neurons
Immature neurons
Glia
Neural progenitors
UMAP 1
C
MED
NFAT
CG4360
DEAF1
ABD-A
CLAMP
TTK
# of TGs
(# positive TG)
(# negative TG)
physical interaction
coordination avg. score
D
Embryo scRNA-seq
Neural progenitors
Immature neurons
Mature neurons
Ventral midline
UMAP 1
UMAP 2
E
DEAF1 correlations to all genes
CLAMP correlations to all genes
DEAF1 correlations to synaptic genes
CLAMP correlations to synaptic genes
Number of genes
Values
F
Embryo scATAC-seq
Mature neurons
Neural progenitors
Ventral midline
Glia
Immature neurons
UMAP 1
UMAP 2
G
134
174
56
DEAF1
CLAMP
H
9,490 kb
9,500 kb
9,510 kb
9,520 kb
9,530 kb
9,540 kb
Neural progenitors
Immature neurons
Mature neurons
Wnt2
brp
I
20,200 kb
20,210 kb
20,220 kb
20,230 kb
20,240 kb
Neural progenitors
Immature neurons
Mature neurons
Nlg4
CG17199
CR46043

Figure S3

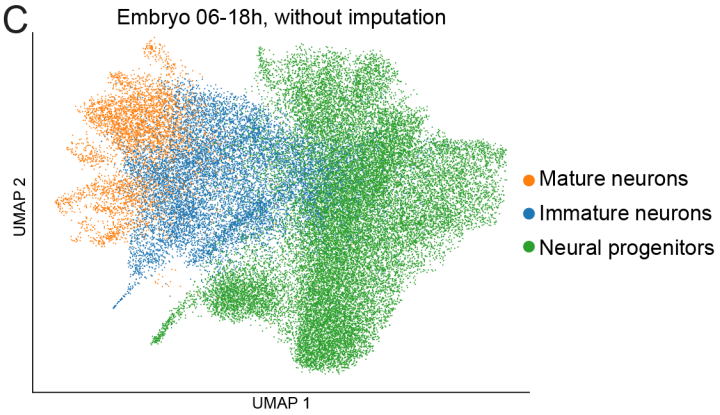
A

Motif	E-value	# of sites	Best match
	4.7×10^{-249}	334	klu
	3.4×10^{-206}	445	Trl
	3.7×10^{-111}	591	kni/jim
	1.0×10^{-34}	330	lola
	3.0×10^{-29}	178	Mad
	5.6×10^{-24}	70	l(3)neo38
	3.8×10^{-14}	110	Adf1
	4.7×10^{-5}	101	slp1

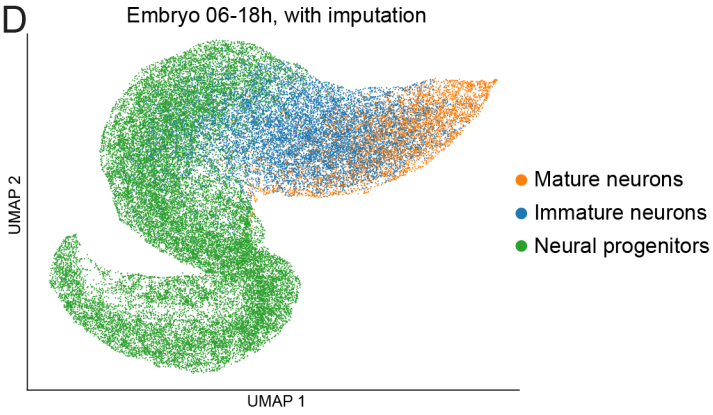
B



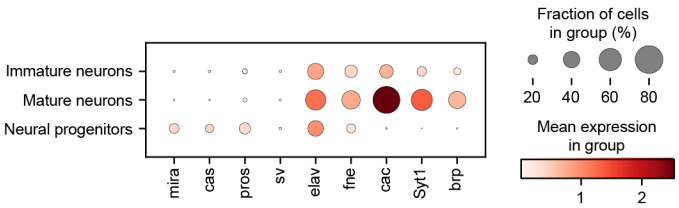
C



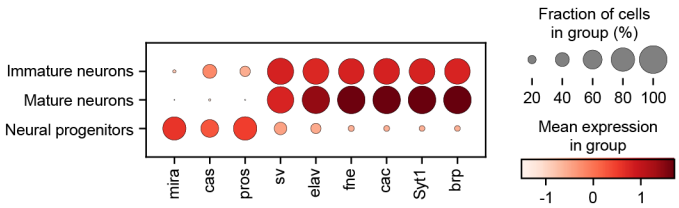
D



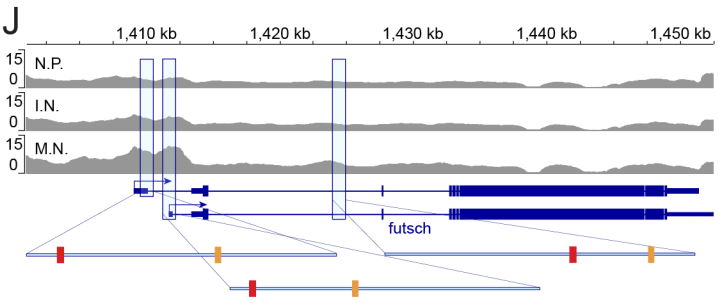
H



I



J



K

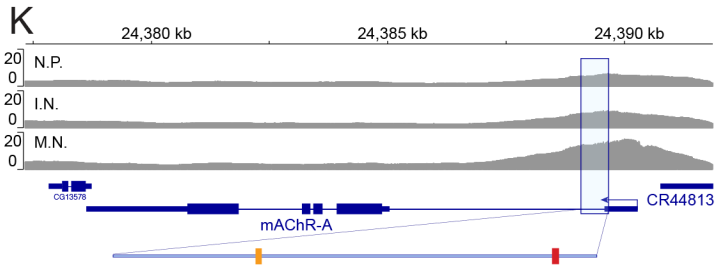
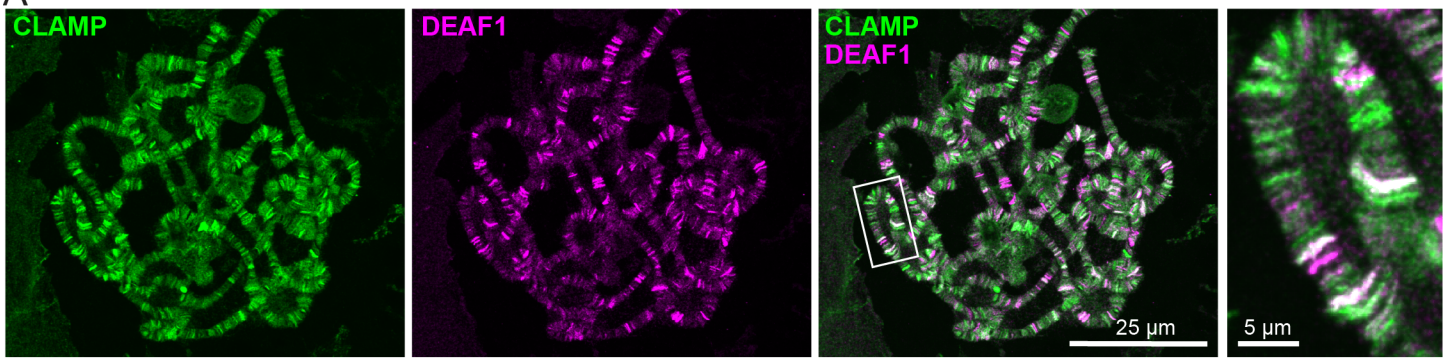
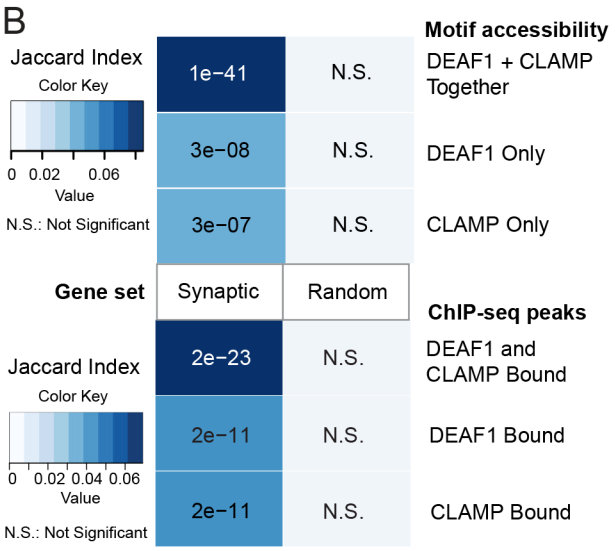


Figure 4

A


CLAMP
DEAF1
CLAMP
DEAF1
25 µm
5 µm

B


Jaccard Index
Color Key
0 0.02 0.06
Value
N.S.: Not Significant
1e−41
N.S.
DEAF1 + CLAMP Together
3e−08
N.S.
DEAF1 Only
3e−07
N.S.
CLAMP Only
Motif accessibility
Gene set
Synaptic
Random
ChIP-seq peaks
Jaccard Index
Color Key
0 0.02 0.04 0.06
Value
N.S.: Not Significant
2e−23
N.S.
DEAF1 and CLAMP Bound
2e−11
N.S.
DEAF1 Bound
2e−11
N.S.
CLAMP Bound

C

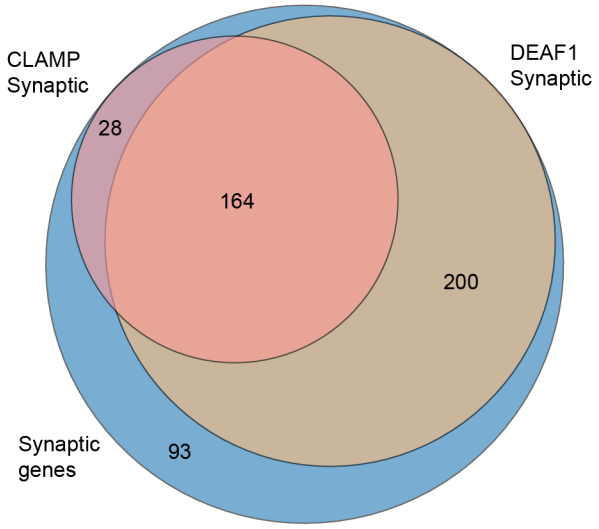

CLAMP Synaptic
DEAF1 Synaptic
28
164
200
Synaptic genes
93

Figure S4

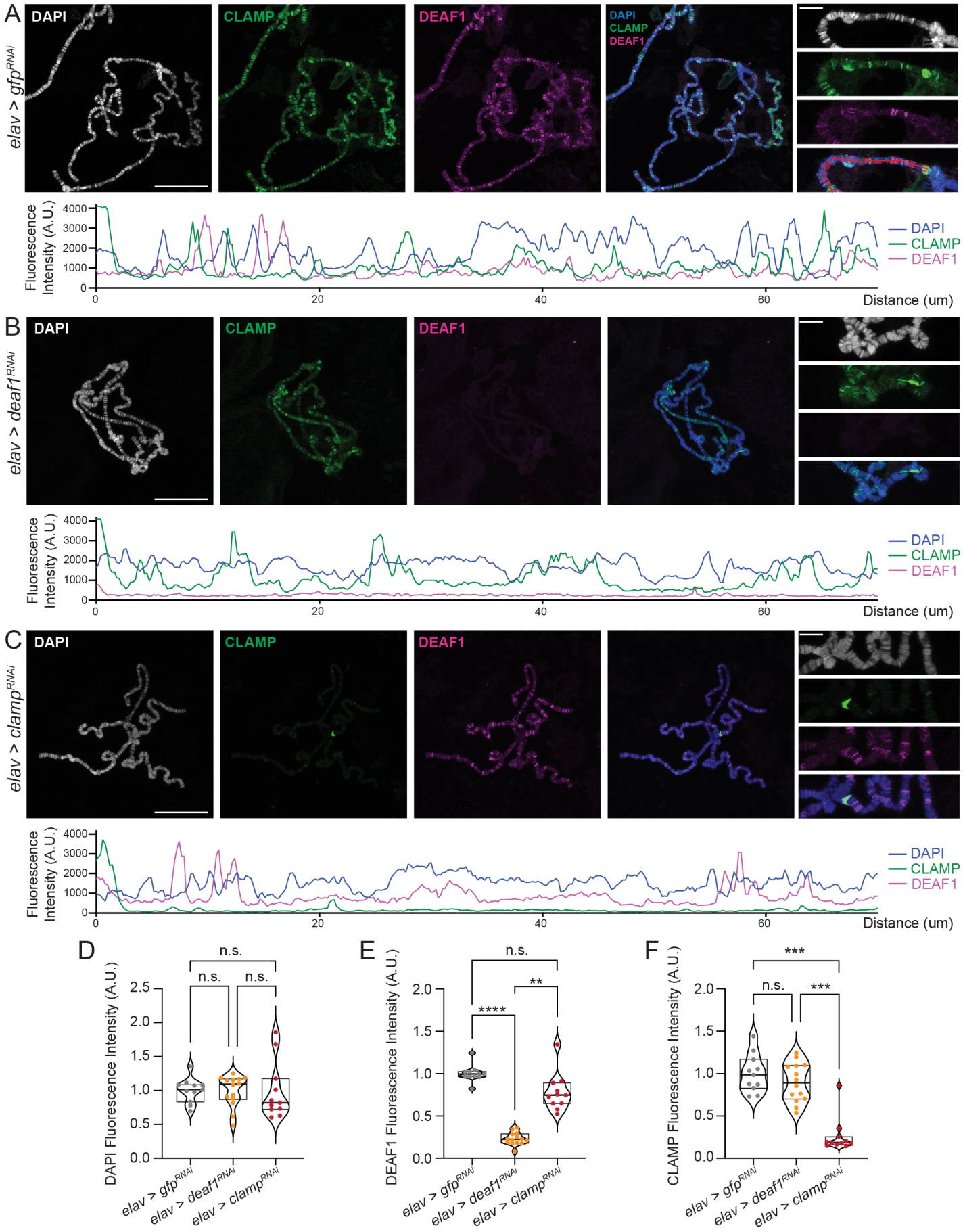

A
DAPI
CLAMP
DEAF1
DAPI
CLAMP
DEAF1
elav > gfp^RNAi
Fluorescence Intensity (A.U.)
4000
3000
2000
1000
0
0
20
40
60
Distance (um)
DAPI
CLAMP
DEAF1
B
DAPI
CLAMP
DEAF1
elav > deaf1^RNAi
Fluorescence Intensity (A.U.)
Distance (um)
DAPI
CLAMP
DEAF1
C
DAPI
CLAMP
DEAF1
elav > clamp^RNAi
Fluorescence Intensity (A.U.)
Distance (um)
DAPI
CLAMP
DEAF1
D
DAPI Fluorescence Intensity (A.U.)
n.s.
n.s.
n.s.
2.5
2.0
1.5
1.0
0.5
0.0
elav > gfp^RNAi
elav > deaf1^RNAi
elav > clamp^RNAi
E
DEAF1 Fluorescence Intensity (A.U.)
n.s.
****
**
2.0
1.5
1.0
0.5
0.0
elav > gfp^RNAi
elav > deaf1^RNAi
elav > clamp^RNAi
F
CLAMP Fluorescence Intensity (A.U.)
***
n.s.
***
2.0
1.5
1.0
0.5
0.0
elav > gfp^RNAi
elav > deaf1^RNAi
elav > clamp^RNAi

Figure 5

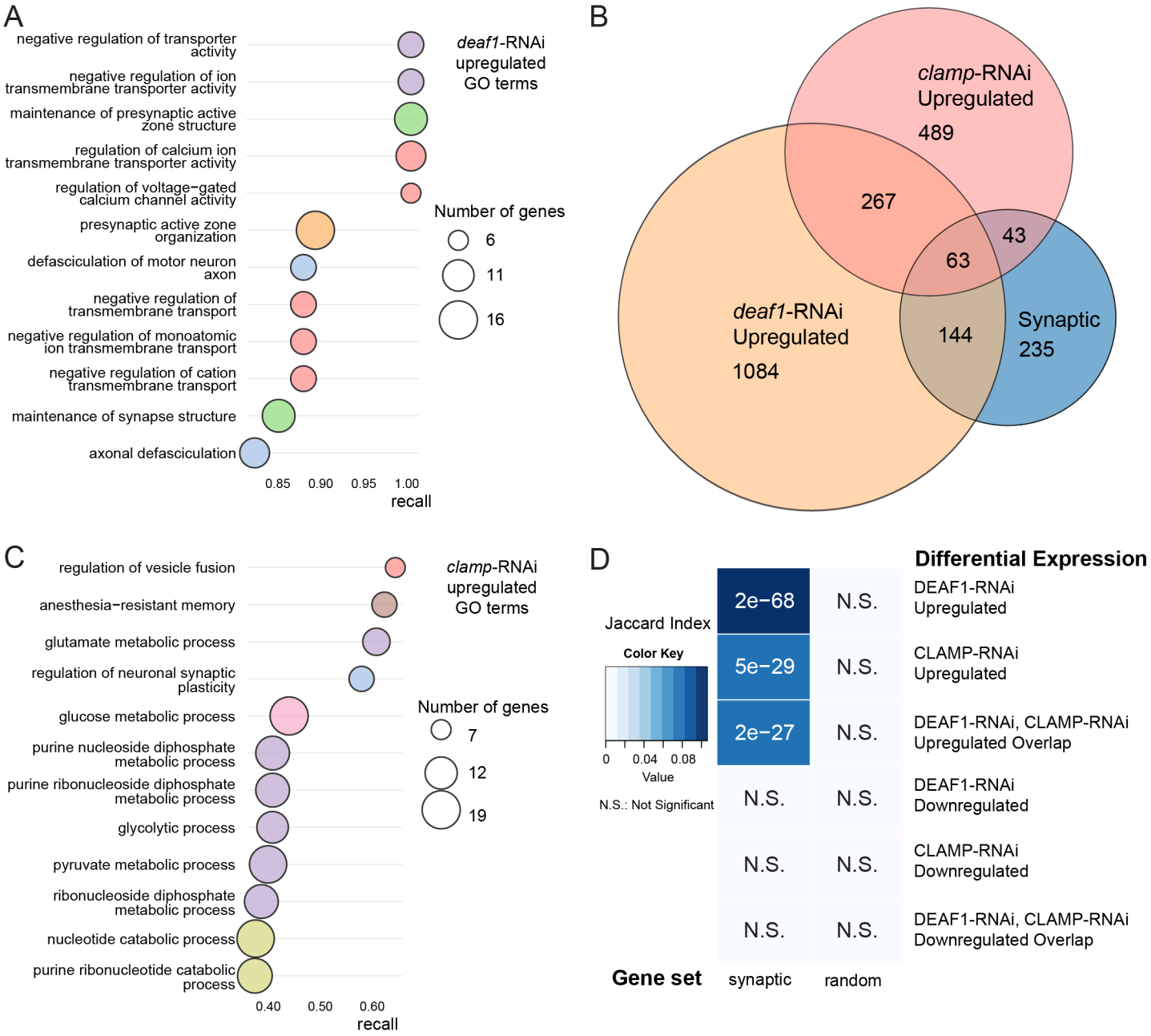

A
negative regulation of transporter activity
negative regulation of ion transmembrane transporter activity
maintenance of presynaptic active zone structure
regulation of calcium ion transmembrane transporter activity
regulation of voltage−gated calcium channel activity
presynaptic active zone organization
defasciculation of motor neuron axon
negative regulation of transmembrane transport
negative regulation of monoatomic ion transmembrane transport
negative regulation of cation transmembrane transport
maintenance of synapse structure
axonal defasciculation
0.85
0.90
0.95
1.00
recall
deaf1-RNAi upregulated GO terms
Number of genes
6
11
16
B
clamp-RNAi Upregulated
489
267
43
63
deaf1-RNAi Upregulated
1084
144
Synaptic
235
C
regulation of vesicle fusion
anesthesia−resistant memory
glutamate metabolic process
regulation of neuronal synaptic plasticity
glucose metabolic process
purine nucleoside diphosphate metabolic process
purine ribonucleoside diphosphate metabolic process
glycolytic process
pyruvate metabolic process
ribonucleoside diphosphate metabolic process
nucleotide catabolic process
purine ribonucleotide catabolic process
0.40
0.50
0.60
recall
clamp-RNAi upregulated GO terms
Number of genes
7
12
19
D
Jaccard Index
Color Key
0
0.04
0.08
Value
N.S.: Not Significant
Differential Expression
2e−68
N.S.
DEAF1-RNAi Upregulated
5e−29
N.S.
CLAMP-RNAi Upregulated
2e−27
N.S.
DEAF1-RNAi, CLAMP-RNAi Upregulated Overlap
N.S.
N.S.
DEAF1-RNAi Downregulated
N.S.
N.S.
CLAMP-RNAi Downregulated
N.S.
N.S.
DEAF1-RNAi, CLAMP-RNAi Downregulated Overlap
Gene set
synaptic
random

Figure S5

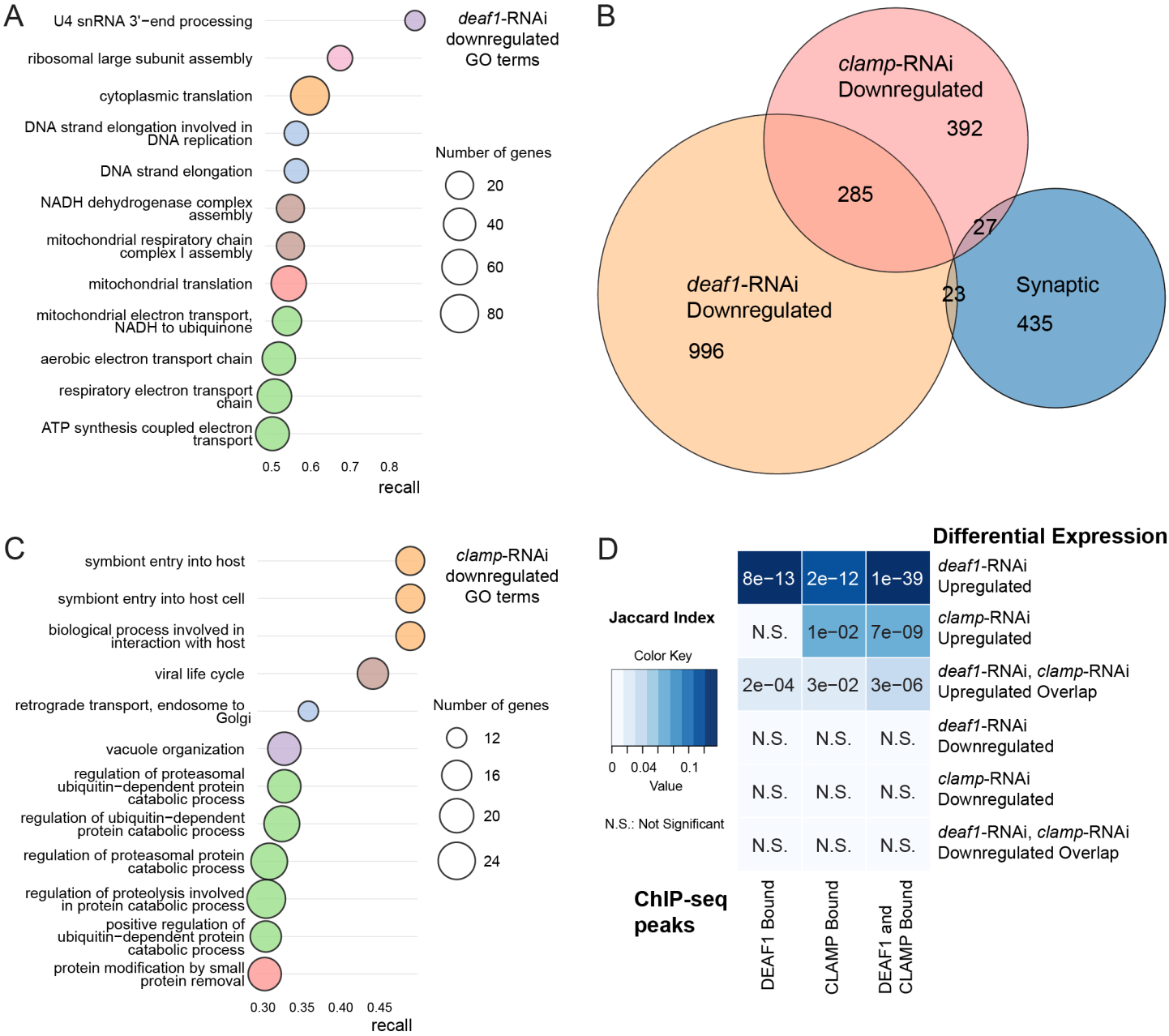

A
U4 snRNA 3'−end processing
ribosomal large subunit assembly
cytoplasmic translation
DNA strand elongation involved in DNA replication
DNA strand elongation
NADH dehydrogenase complex assembly
mitochondrial respiratory chain complex I assembly
mitochondrial translation
mitochondrial electron transport, NADH to ubiquinone
aerobic electron transport chain
respiratory electron transport chain
ATP synthesis coupled electron transport
0.5 0.6 0.7 0.8
recall
deaf1-RNAi downregulated GO terms
Number of genes
20
40
60
80
B
clamp-RNAi Downregulated
392
285
27
23
deaf1-RNAi Downregulated
996
Synaptic
435
C
symbiont entry into host
symbiont entry into host cell
biological process involved in interaction with host
viral life cycle
retrograde transport, endosome to Golgi
vacuole organization
regulation of proteasomal ubiquitin−dependent protein catabolic process
regulation of ubiquitin−dependent protein catabolic process
regulation of proteasomal protein catabolic process
regulation of proteolysis involved in protein catabolic process
positive regulation of ubiquitin−dependent protein catabolic process
protein modification by small protein removal
0.30 0.35 0.40 0.45
recall
clamp-RNAi downregulated GO terms
Number of genes
12
16
20
24
D
Jaccard Index
Color Key
0 0.04 0.1
Value
N.S.: Not Significant
Differential Expression
8e−13 2e−12 1e−39 deaf1-RNAi Upregulated
N.S. 1e−02 7e−09 clamp-RNAi Upregulated
2e−04 3e−02 3e−06 deaf1-RNAi, clamp-RNAi Upregulated Overlap
N.S. N.S. N.S. deaf1-RNAi Downregulated
N.S. N.S. N.S. clamp-RNAi Downregulated
N.S. N.S. N.S. deaf1-RNAi, clamp-RNAi Downregulated Overlap
ChIP-seq peaks
DEAF1 Bound
CLAMP Bound
DEAF1 and CLAMP Bound

Figure 6

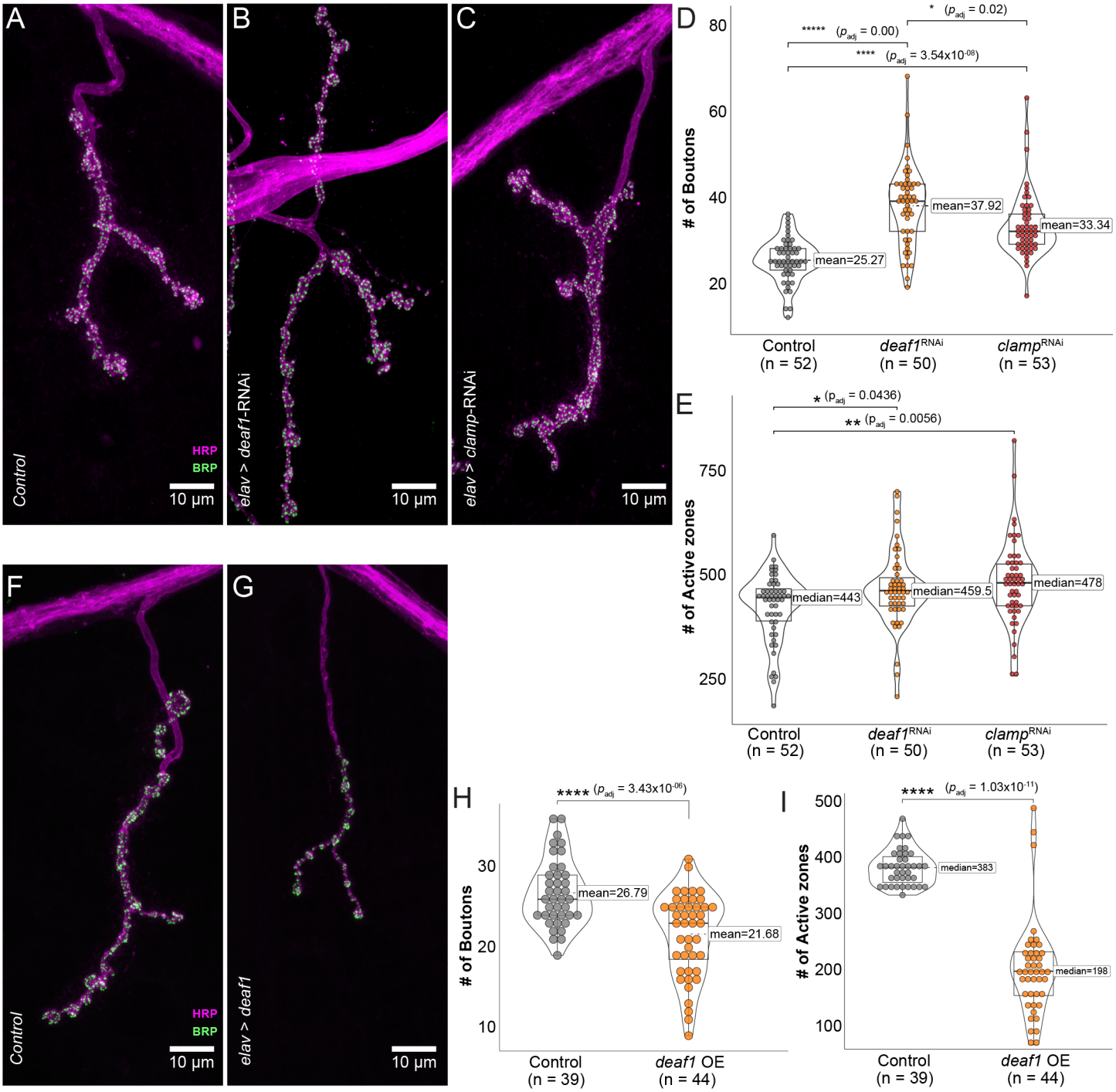
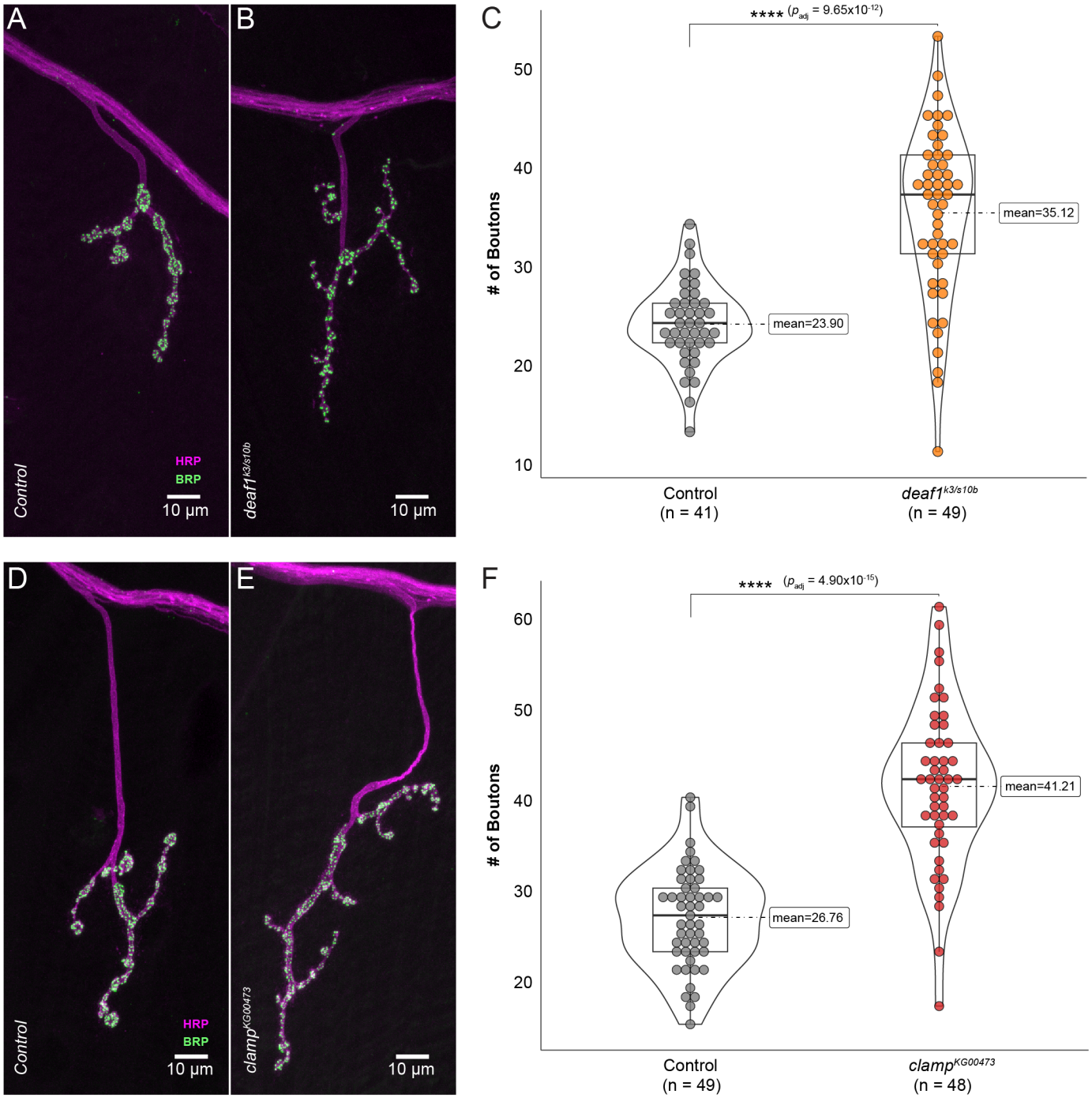

A
B
C
D
E
F
G
H
I
Control
elav > deaf1-RNAi
elav > clamp-RNAi
HRP
BRP
10 µm
elav > deaf1
# of Boutons
# of Active zones
mean=25.27
mean=37.92
mean=33.34
median=443
median=459.5
median=478
mean=26.79
mean=21.68
median=383
median=198
Control (n = 52)
deaf1RNAi (n = 50)
clampRNAi (n = 53)
Control (n = 39)
deaf1 OE (n = 44)
***** (padj = 0.00)
**** (padj = 3.54x10-08)
* (padj = 0.02)
* (padj = 0.0436)
** (padj = 0.0056)
**** (padj = 3.43x10-06)
**** (padj = 1.03x10-11)

Figure S6


A
B
C
D
E
F
Control
HRP
BRP
10 µm
deaf1k3/s10b
clampKG00473
**** (padj = 9.65x10-12)
mean=35.12
mean=23.90
# of Boutons
Control
(n = 41)
deaf1k3/s10b
(n = 49)
**** (padj = 4.90x10-15)
mean=41.21
mean=26.76
Control
(n = 49)
clampKG00473
(n = 48)