Dynamic neurotransmitter specific transcription factor expression profiles during *Drosophila* development

Alicia Estacio-Gómez, Amira Hassan, Emma Walmsley, Lily Lee and Tony D. Southall*

Department of Life Sciences, Imperial College London, Sir Ernst Chain Building, London, UK

*Correspondence to t.southall@imperial.ac.uk

Abstract

The remarkable diversity of neurons in the nervous system is generated during development, when properties such as cell morphology, receptor profiles and neurotransmitter identities are specified. Neurotransmitters are essential for neuronal communication, and transcription factors are the key determinants of neurotransmitter fate. In order to gain a greater understanding of neurotransmitter specification we profiled the transcription state of cholinergic, GABAergic and glutamatergic neurons *in vivo* at multiple developmental time points. We identified 86 differentially expressed transcription factors that are uniquely enriched, or uniquely depleted, in a specific neurotransmitter subtype. Some transcription factors show a similar profile across development, others only show enrichment or depletion at specific developmental stages. Profiling of acj6 (cholinergic enriched) and Ets65A (cholinergic depleted) binding sites *in vivo* reveals that they both directly bind the *ChAT* locus, in addition to a wide spectrum of other key neuronal differentiation genes. Our data underlines that, similar to *C. elegans*, there are no simple transcription factor codes for neurotransmitter type specification, and provides a useful resource for studying the developmental basis of neurotransmitter specification in *Drosophila* and other model systems.

Introduction

The human brain is perhaps the most complex system known to mankind. It consists of approximately 85 billion neurons, which possess very diverse morphologies, neurotransmitter identities, electrical properties and preferences for synaptic partners. Understanding how this diversity is generated is one of the greatest challenges in biology and can only be achieved by identifying the underlying molecular mechanisms that determine these neuronal properties. Neurotransmitters allow neurons to communicate with each other, enabling organisms to sense, interpret and interact with their environment. Fast-acting neurotransmitters include acetylcholine and glutamate, which are, in general, excitatory, and GABA, which is inhibitory. The function of individual neurons depends on the specific types of neurotransmitters they produce, which in turn ensures proper information flow and can also influence the formation of neural circuits (Andreae and Burrone, 2018). Therefore the proper specification of neurotransmitter fate is fundamental for nervous system development.

Model organism studies in *C. elegans*, mice and *Drosophila* have provided a wealth of information about factors and mechanisms involved in neurotransmitter specification. Comprehensive neurotransmitter maps (Hobert, 2016) and the description of terminal selector genes in *C. elegans* (Hobert, 2008) have provided important contributions to the field. These terminal selectors are transcription factors (or a transcription factor complex) that regulate the expression of a battery of terminal differentiation genes in the last phase of neuronal differentiation, and maintain the expression of these genes during the lifetime of a neuron (Hobert, 2008). For example, the *C. elegans* transcription factors *ttx-3* and *unc-86* act as terminal selectors in distinct cholinergic and serotonergic neuron populations, respectively (Zhang et al., 2014).

Cellular context is important for the action of these specifying factors, as misexpression of terminal selectors in other neuronal subtypes is often not sufficient to reprogram their fate (Duggan et al., 1998; Wenick and Hobert, 2004), The presence of co-factors, and likely the chromatin state, can also influence this plasticity (Altun-Gultekin et al., 2001; Patel and Hobert, 2017). Related to this, there appears to be little evidence for master regulators of cholinergic, GABAergic or glutamatergic fate (Konstantinides et al., 2018; Lacin et al., 2019; Serrano-Saiz et al., 2013). Rather, individual lineages, or subpopulations, utilise different transcription factors (or combinations of transcription factors) to specify the fast-acting neurotransmitter that they will utilise. Developmental context also plays a role in the mechanisms governing neurotransmitter specification. In *Drosophila*, early born embryonic neurons in a given lineage can use different neurotransmitters (Landgraf et al., 1997; Schmid et al., 1999). However, strikingly, each post-embryonic lineage only uses one neurotransmitter (Lacin et al., 2019), implying that specification occurs at the stem cell level during larval stages.

Neurotransmitter specification studies across different organisms have highlighted conserved mechanisms. A prominent example is the binding of the transcription factors AST-1 (*C. elegans*) and Etv1 (vertebrates) to a phylogenetically conserved DNA motif to specify dopaminergic fate (Flames and Hobert, 2009). Furthermore, orthologues *acj6* (*Drosophila*), *unc-86* (*C. elegans*) and *Brn3A*/*POU4F1* (vertebrates) all have roles in cholinergic specification (Lee and Salvaterra, 2002; Serrano-Saiz et al., 2018; Zhang et al., 2014), while *PITX2* (vertebrates) and *unc-30* (*C. elegans*) both control GABAergic differentiation (Jin et al., 1994; Waite et al., 2011; Westmoreland et al., 2001).

In order to identify novel candidate genes, and investigate the dynamics of neurotransmitter specific transcription factors throughout development, we have performed cell specific profiling of RNA polymerase II occupancy, *in vivo*, in cholinergic, GABAergic and glutamatergic neurons of *Drosophila*. We identify 86 transcription factors that show differential expression between neurotransmitter subtypes, in at least one developmental time point. There are both uniquely enriched and uniquely depleted transcription factors, and we show that acj6 (cholinergic enriched) and Ets65A (cholinergic depleted) both directly bind the choline acetyltransferase gene (*ChAT*) required for cholinergic fate.

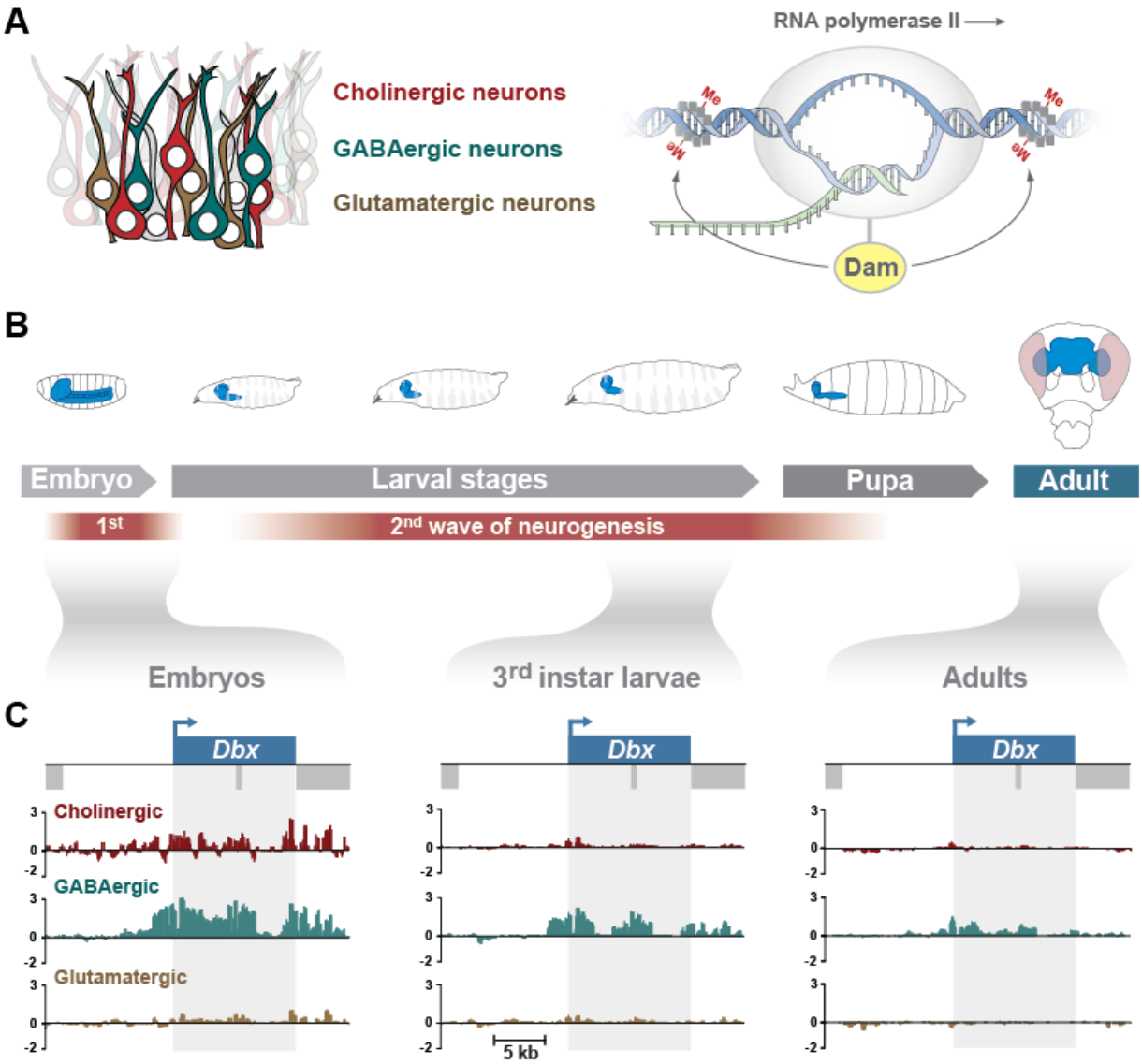

Figure 1. Cell specific profiling of RNA Pol II occupancy in different neuronal subtypes throughout Drosophila development. A) Profiling of RNA Pol II occupnacy in cholinergic, GABAergic and glutamatergic neurons using TaDa. B) Profling windows cover embryonic nervous system development (5 - 29 hr AEL), 3rd instar larval nervous system development (24 hr window before pupation) and the adult brain (heads from ~ 3-4 day old adults after a 24 hr expression window). Temporal restriction of Dam-Pol II expression was controlled using a temperature sensitive GAL80. C) An example of a transcription factor gene (Dbx) that is uniquely transcribed in GABergic neurons. Y-axis represent log2 ratios of Dam-Pol II over Dam-only.

Results

Transcriptional profiling of neuronal subtypes across development

In order to investigate which genes participate in the specification of neuronal properties, namely, neurotransmitter choice, we applied the cell specific profiling technique Targeted DamID (TaDa). Targeted DamID is based on DamID (van Steensel and Henikoff, 2000) and allows the profiling of protein-DNA interactions without the need for cell isolation, specific antibodies or fixation (Aughey et al., 2019; Southall et al., 2013). Transcriptional profiling is also possible with TaDa using the core subunit of RNA polymerase II (Pol II) (Southall et al., 2013). We have mapped the occupancy of Pol II in cholinergic, GABAergic and glutamatergic neurons, using specific GAL4 drivers that trap the

97 expression of the genes *ChAT* (choline acetyltransferase), *Gad1* (Glutamic acid decarboxylase 1) and
98 *VGlut* (vesicular Glutamate transporter) (Diao et al., 2015). During *Drosophila* development, there are
99 two neurogenic periods, the first to produce the larval nervous system, and the second to produce the
100 adult nervous system. Therefore, to cover both developing stages and adult neurons, we profiled
101 embryonic neurons, larval postembryonic neurons and adult neurons (see Figure 1B). Windows of 20
102 hr (First instar larvae samples), and 24 hr (third instar larvae and adult samples) were used for TaDa
103 profiling and 3 replicates were performed for each experiment. The number of genes bound by Pol II
104 ranged from 1170 to 1612 (see Table S1). To investigate the global differences in Pol II occupancy
105 between neuronal subtypes and developmental stages, we generated a correlation matrix (Figure 2A).
106 We found that the greatest variability is between developmental stages, rather than between cell
107 types, with the adult brain data being more distinct from the embryonic and larval stages. When
108 focusing on transcription factor genes, a similar pattern is evident (Figure 2B).

109 For each developmental stage, we identified uniquely enriched genes (i.e. genes enriched in one
110 neurotransmitter compared to the other two neurotransmitter subtypes) (Table S2). Encouragingly, a
111 strong enrichment of Pol II occupancy is evident at *ChAT*, *Gad1* and *VGlut*, the genes encoding the key
112 enzymes involved in the biosynthesis of these neurotransmitters (Figure S1). Transcription factors
113 and non-coding RNAs make up a large proportion of all the enriched genes, at each developmental
114 stage (Figure 2C). In the adult, almost a quarter (23/97) of the enriched genes are transcription
115 factors. Other enriched genes include the immunoglobulin domain containing *beaten path* (*beat*) and
116 *Down syndrome cell adhesion molecule* (*Dscam*) genes, which play roles in axon guidance and dendrite
117 self-avoidance (Pipes et al., 2001; Soba et al., 2007). Glutamatergic genes include *twit* and *Dad*, both of
118 which are known to regulate synaptic homeostasis at the neuromuscular junction (Goold and Davis,
119 2007; Kim and Marques, 2012). Interestingly, there is an enriched expression of MAP kinase inhibitors
120 in cholinergic neurons (Figure 2C). Also, glutamatergic neurons express higher levels of the
121 monoamine neurotransmitter related genes *Vmat*, *DAT* and *Tdc2*, while GABAergic neurons are
122 enriched for serotonergic and dopaminergic receptors, relative to the other two fast-acting
123 neurotransmitter subtypes (Figure 2C). Very few genes show enrichment across all developmental
124 stages: five for cholinergic (*ChAT*, *ChT*, *acj6*, *Mef2* and *sosie*), five for GABAergic (*Gad1*, *Dbx*, *vg*,
125 *CG13739* and *CG14989*) and two for glutamatergic (*VGlut* and *oc*) (Table S2). There is consistent
126 enrichment of the GAL4-trapped genes (*ChAT*, *Gad1* and *VGlut*) (Figure S1) that provide subtype
127 specific expression for the TaDa experiments.

128 CATaDa profiling of chromatin accessibility (Aughey et al., 2018) reveals that, similar to RNA Pol II,
129 global chromatin accessibility does not vary greatly between cell types (Figure 2) but shows more
130 differences between developmental stages. Correspondingly, when examining regions of the genome
131 that display robust changes in chromatin accessibility (>10 RPM differences across 3 consecutive
132 GATC fragments) during embryo development, only 37 GATC fragments (13 individual peaks) are
133 identified, with 62% mapping to the loci of the three neurotransmitter synthesis genes (*ChAT*, *Gad1*

and *VGlut*) (Figure S2B and C). Differential accessibility is also present at sites outside of the gene and promoter for *Gad1* and *VGlut* (yellow arrows in Figure S2C). Weaker differences in accessibility are also observed at some of the differentially expressed transcription factor loci (Figure S2D).

Figure 2. Transcription factors and non-coding RNAs are enriched in specific neurotransmitter subtypes. A) Correlation matrix for RNA Pol II signal (log2 over Dam-only for all genes) and chromatin accessibility (CATaDa) for all gene loci (extended 5 kb upstream and 2 kb downstream). (B) Correlation matrix for RNA Pol II and chromatin accessibility at annotated transcription factor genes. (C) Proportion of transcription factors and non-coding RNAs at each developmental stage. Examples of enriched GO term categories for the remaining genes are also included.

Identification of transcription factors uniquely enriched, or uniquely depleted in neurotransmitter subtypes

Transcription factors play the major role in neurotransmitter specification and we have identified many with enriched Pol II occupancy in specific neurotransmitter subtypes (Figure 2A). Uniquely

enriched transcription factors are candidates for activators of neurotransmitter identity and conversely, if there is depletion (or absence) of a transcription factor from only one subtype, they are candidates for repressors of neurotransmitter identity. For example, a hypothetical transcription factor that represses GABAergic fate would be present in both cholinergic and glutamatergic neurons but absent from GABAergic neurons.

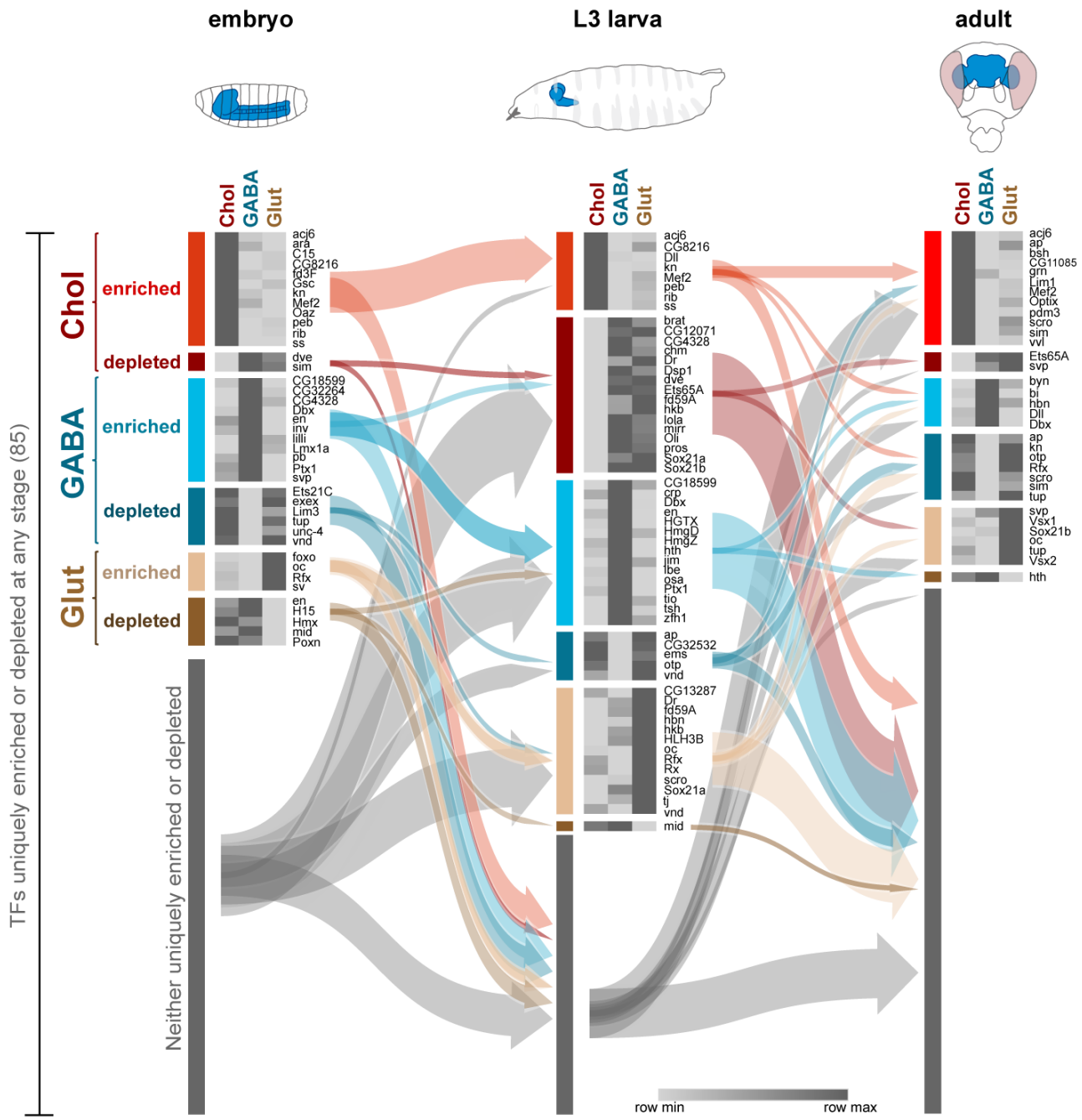

Figure 3. Transitions of uniquely enriched or depleted transcription factors during neural development. Transcription factors uniquely enriched or depleted in cholinergic, GABAergic and glutamatergic neurons. A total of 86 transcription factors are identified across all stages.

To investigate the expression pattern dynamics of both uniquely enriched and uniquely depleted transcription factors, we examined how their expression patterns transitioned across the stages of development (Figure 3). We observe a great deal of flux between transcription factor expression in cell types and developmental stages. While *acj6* (cholinergic), *Dbx* (GABAergic – see Figure 1) and *oc*

(glutamatergic) are enriched in their respective neurotransmitter type throughout all stages, many are only enriched in one or two of the developmental stages. In support of our data, Acj6 is known to promote cholinergic fate in the peripheral nervous system (Lee and Salvaterra, 2002) and Dbx is important for the proper differentiation of a subset of GABAergic interneurons (Lacin et al., 2009). When a gene is enriched in one particular subtype, it tends to either also be enriched in the next developmental stage, or is no longer enriched. *Dll* is an unusual case, as it is cholinergic enriched in larvae, however switches to being GABAergic enriched in adults (Figure 3 and Figure S3).

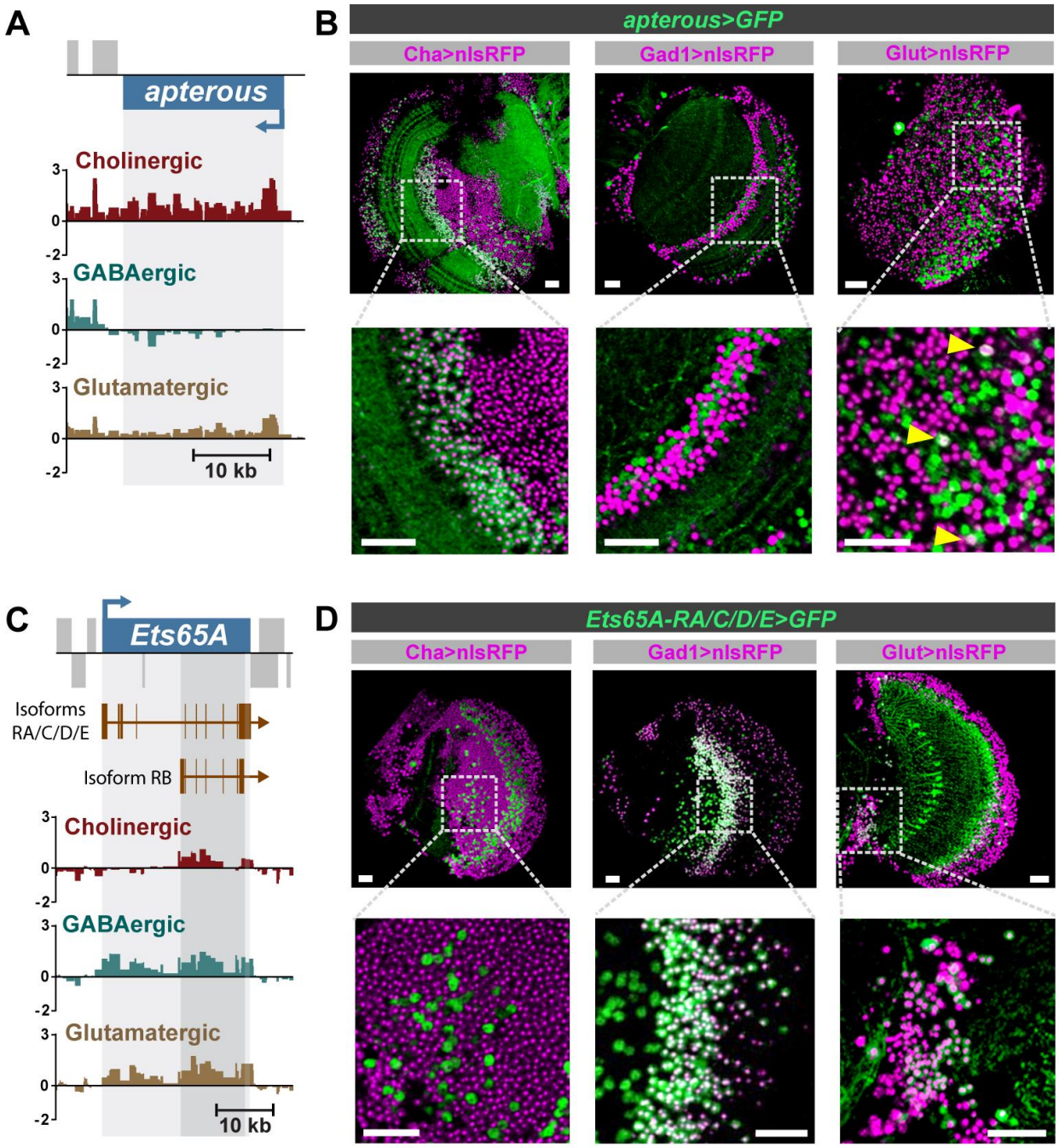

Figure 4. Absence of apterous and Ets65A-RA/C/D/E expression in cholinergic and GABAergic adult neurons, respectively. (A) Pol II occupancy at *apterous* in the adult brain. Y-axis represent log2 ratios of Dam-Pol II over Dam-only. (B) Expression pattern of apterous in the adult brain. (C) Pol II occupancy at Ets65A in the adult brain. Y-axis represent log2 ratios of Dam-Pol II over Dam-only. (D) Expression pattern of Ets65A-RA/C/D/E in the adult brain.

219 Candidate repressors of neurotransmitter fate (uniquely depleted transcription factors) also
220 demonstrate dynamic changes in expression pattern across development (Figure 3). Prominent
221 examples are the longer transcripts of *Ets65A* (absent in cholinergic) and *apterous* (absent in
222 GABAergic) (Figure 4A and 4C). We used genetic reporters to examine the expression pattern of
223 *Ets65A* and *apterous* in adult brains (Figures 4B and 4D). In agreement with our data, the cholinergic
224 reporter is absent in *Ets65A-RA/C/D/E* positive cells, and the GABAergic reporter absent in *apterous*
225 positive cells.

226 We have identified transcription factors with potentially novel roles in regulating neurotransmitter
227 identity. Therefore, we investigated candidate activators and candidate repressors for their potential
228 to elicit pan-neural reprogramming of neurotransmitter identity. Pan-neural expression and RNAi
229 knockdown of candidate activator transcription factors (*Dbx*, *en* and *CG4328*) and candidate repressor
230 transcription factors (*ap* and *otp*) during embryonic development did not result in any obvious
231 changes in neurotransmitter expression patterns (Figure S4A). To test whether Ets65A-PA can repress
232 cholinergic fate in any neuronal context, we generated overexpression clones in the adult. These
233 clones contained cells expressing *ChAT* (Figure S4B), therefore, Ets65A-PA does not have a universal
234 capacity to repress cholinergic fate (see Discussion).

235 Focusing on candidate transcription factors demonstrating binary differences (clear on and off), we
236 performed literature searches to examine whether they have been previously shown, or implicated in
237 regulating neurotransmitter identity (Figure 5). This included *C. elegans* and mouse orthologues, as
238 much of the work in this field has utilised these model organisms. For example, the orthologues of
239 cholinergic enriched *acj6* (*unc-86*), GABAergic enriched *Ptx1* (*PITX1* and *unc-30*) and glutamatergic
240 enriched *oc* (*OTX1/2* and *ttx-1*) have all shown to have a role in promoting cholinergic, GABAergic and
241 glutamatergic fate, respectively. However, there are many that have not been investigated in this
242 context (38%), or that are only supported by indirect evidence (38%). These include *Dll* (*DLX*, *ceh-43*),
243 *sox21a* (*SOX21, sox-3*), *hbn* (*ARX*, *alr1*, *unc-4*) and *otp* (*OTP*, *npax-1*). Given the strong conservation of
244 neurotransmitter specification mechanisms, many of these newly highlighted factors provide
245 promising research avenues for expanding our knowledge in this field.

246 While non-coding ribosomal RNAs and tRNAs and are transcribed by RNA polymerase I and III, micro
247 RNAs (miRNAs) and long non-coding RNAs (lncRNAs) are primarily transcribed by Pol II. Our Dam-Pol
248 II data identifies a set of differentially bound miRNAs and lncRNAs, between the neurotransmitter
249 subtypes (Figure S5A). These include non-characterised lncRNAs and GABAergic enriched *iab8*, which
250 is located in the Hox cluster and plays a role in the repression of *abd-A* (Gummalla et al., 2012). A small
251 number of miRNAs were also identified, most notably, *mir-87* (cholinergic), *mir-184* (GABAergic) and
252 *mir-190* (glutamatergic), which are enriched during the developing states but not in the adult.
253 Although annotated separately, *mir-184* is embedded in *CR44206* (Figure S5B).

254

		Gene	Mammalian orthologue	C. elegans orthologue	Expected role in neurotransmitter specification? (*Drosophila* and orthologue literature)			References
					Drosophila	Vertebrates	C. elegans	
Chol	enriched	***acj6***	*Brn3a/POU4F*	*unc-86*	✓	✓	✓	Lee and Salvaterra, 2002 Serrano-Saiz et al., 2018 Zhang et al., 2014
Chol	enriched	***Dll***	*DLX*	*ceh-43*	?	?	?	
Chol	enriched	***kn***	*EBF*	*unc-3*	?	?	✓	Kratsios et al., 2012
Chol	enriched	***peb***	*RREB1*	*sem-4*	?	?	?	
Chol	enriched	***rib***	*zbtb9*	-	?	?	?	
Chol	enriched	***ss***	*AHR*	*ahr-1*	?	(✓)	-	Hamzah and Abdullah, 2013
Chol	depleted	***CG4328***	*LMX1B*	*lim-6*	?	?	?	
Chol	depleted	***dve***	*SATB1*	*dve-1*	?	?	?	
Chol	depleted	***Ets65A***	*FLI1, ERG*	*ast-1*	?	(✓)	?	McKeon, et al., 1988
Chol	depleted	***fd59A***	*FOXD3*	*unc-130*	?	?	(✓)	Sarafi-Reinach and Sengupta, 2000
Chol	depleted	***hkb***	*EGR-1, KLF1*	*mnm-2, pat-9*	?	?	?	
Chol	depleted	***Sox21a***	*SOX21*	*sox-3*	?	?	?	
GABA	enriched	***CG18599***	*NOTO, VAX1*	*alr-1*	?	✓	(✓)	Taglialetela et al., 2004 Melkman and Sengupta, 2005
GABA	enriched	***Dbx***	*DBX1*	*ceh-51, egl-5*	(✓)	(✓)	?	
GABA	enriched	***en***	*EN1, EN2*	*ceh-16*	?	✓	?	
GABA	enriched	***HGTX***	*Nkx1 / 2*	*cog-1*	?	(✓)	?	Fogarty et al., 2007
GABA	enriched	***Ptx1***	*PITX1*	*unc-30*	?	✓	✓	Waite et al., 2011 Jin et al., 1994 Westmoreland et al., 2001
GABA	depleted	***ap***	*LHX2*	*ttx-3*	?	?	(✓)	Hobert, 2016
GABA	depleted	***CG32532***	*PROP1*	*unc-42*	?	?	(✓)	Serrano-Saiz et al., 2013 Hobert, 2016
GABA	depleted	***ems***	*EMX1 / 2*	*ceh-2*	?	(✓)	?	Gorski et al., 2002 Shinozaki et al., 2002
GABA	depleted	***otp***	*OTP*	*npax-1*	?	?	?	
GABA	depleted	***vnd***	*NKX2-2 / -8*	*ceh-22*	?	?	?	
Glut	enriched	***hbn***	*ARX*	*alr-1, unc-4*	?	(✓)	?	Beguin et al., 2013
Glut	enriched	***oc***	*OTX1 / 2*	*ttx-1*	?	✓	✓	Serrano-Saiz et al., 2013
Glut	enriched	***Rfx***	*RFX*	*daf-19*	?	(✓)	?	Ma et al., 2006
Glut	enriched	***Rx***	*RAX*	*ceh-8*	?	✓	?	Lu et al., 2013
Glut	enriched	***scro***	*NKX2-1 / -4*	*ceh-24*	?	?	?	
Glut	enriched	***vnd***	*NKX2-2 / -8*	*ceh-22*	?	?	?	
Glut	depleted	***mid***	*Tbx20*	*mab-9*	(✓)	?	?	Leal et al., 2009

Figure 5. Evidence for predicted roles of identified transcription factors. Strongly enriched or depleted transcription factors identified in developing larval brains. Uniquely enriched factors are predicted to be candidates that promote the respective neurotransmitter fate, whilst uniquely depleted factors are predicted to repress the neurotransmitter fate. A full tick indicates direct evidence that the transcription factor directly promotes or represses the neurotransmitter fate, whilst indirect, supporting evidence is indicated by faded tick in brackets. A question mark signifies that nothing is currently known, regarding neurotransmitter specification.

282 **Acj6 and Ets65A-PA directly bind to *ChAT* and other key neuronal differentiation genes**

283 Acj6 is enriched in cholinergic neurons (Figure 3) and is known to promote cholinergic fate (Lee and
284 Salvaterra, 2002). Acj6 can bind to specific sites upstream of ChAT *in vitro* (Lee and Salvaterra, 2002),
285 however, the extent of Acj6 binding at the *ChAT* locus *in vivo*, and genome wide, is not known. In order
286 to only profile the cells that endogenously express *acj6*, and therefore gain an more accurate readout
287 of native Acj6 binding, we used an *acj6* GAL4 line (Lai et al., 2008) to drive the expression of the *Dam-*
288 *acj6* transgene. Furthermore, we generated an *Ets65A-RA/C/D/E* MiMIC GAL4 trap line to investigate
289 the *in vivo* binding of Ets65A-PA, with an interest to see whether, as a candidate cholinergic repressor,
290 it could directly bind the *ChAT* locus. In the adult brain, both factors directly bind the ChAT locus
291 (Figure 6A). Acj6 binds at the upstream region studied by (Lee and Salvaterra, 2002), as well as
292 strongly within intronic regions of *ChAT*. Ets65A-PA also binds at the same intronic region, however,
293 it's binding at the upstream region and transcriptional start site of *ChAT* is far more pronounced
294 (Figure 6A), which may reflect a different mode of regulation.

295 Acj6 and Ets65A-PA bind 2708 and 2277 genes, respectively, using a stringent false discovery rate
296 (FDR) (FDR < 0.0001) (Table S4). They co-bind 926 genes, which are highly enriched for nervous
297 system genes, including genes involved in axon development [GO:0061564] and chemical synaptic
298 transmission [GO:0007268] (Figure 6B). While both factors bind the cholinergic signalling regulator
299 gene *Acetylcholine esterase* (*Ace*) gene (Figure 6C), Acj6 uniquely binds *nicotinic Acetylcholine Receptor*
300 *α4* (*nAChRα4*) (Figure 6D) and Ets65A-RA binds multiple genes involved in MAP kinase signalling (e.g.
301 *hep*, *lic*, *Dsor* and *slpr*) (Figure 6E). Therefore, these factors have the potential to regulate not just a
302 single neuronal property, but also a multitude of other genes that govern a wide spectrum of neuronal
303 processes, such as their receptivity to extrinsic signals and synapse formation.

304

305 ## Discussion

306 Neurotransmitter identity is a key property of a neuron that needs to be tightly regulated in order to
307 generate a properly functioning nervous system. Here we have investigated the dynamics and extent
308 of transcription factor specificity in fast-acting neurotransmitter neuronal subtypes in *Drosophila*. We
309 profiled the transcription state of cholinergic, GABAergic and glutamatergic neurons in the developing
310 and adult brain of *Drosophila*. We observe enriched Pol II occupancy at the relevant neurotransmitter
311 synthesis genes (Figure S1) and other genes associated with the activity of the specific subtypes (Table
312 S2). The monoamine neurotransmitter related genes *Vmat*, *DAT* and *Tdc2* are enriched in
313 glutamatergic neurons (Figure 2C), which is not unprecedented, as monoamine populations can also
314 be glutamatergic (Aguilar et al., 2017; Trudeau and El Mestikawy, 2018). Cholinergic, GABAergic,
315 serotonergic and dopaminergic receptors are enriched in embryonic GABAergic neurons relative to
316 the other two fast-acting neurotransmitter subtypes (Figure 2C), which correlates with GABAergic

interneurons acting as integrative components of neural circuits. The enrichment of MAP kinase pathway regulators in cholinergic neurons is intriguing, suggesting that this signalling pathway may have a specific role in these neurons. This is supported by a recent study showing that MAP kinase signalling acts downstream of Gq-Rho signalling in *C. elegans* cholinergic neurons to control neuron activity and locomotion (Coleman et al., 2018).

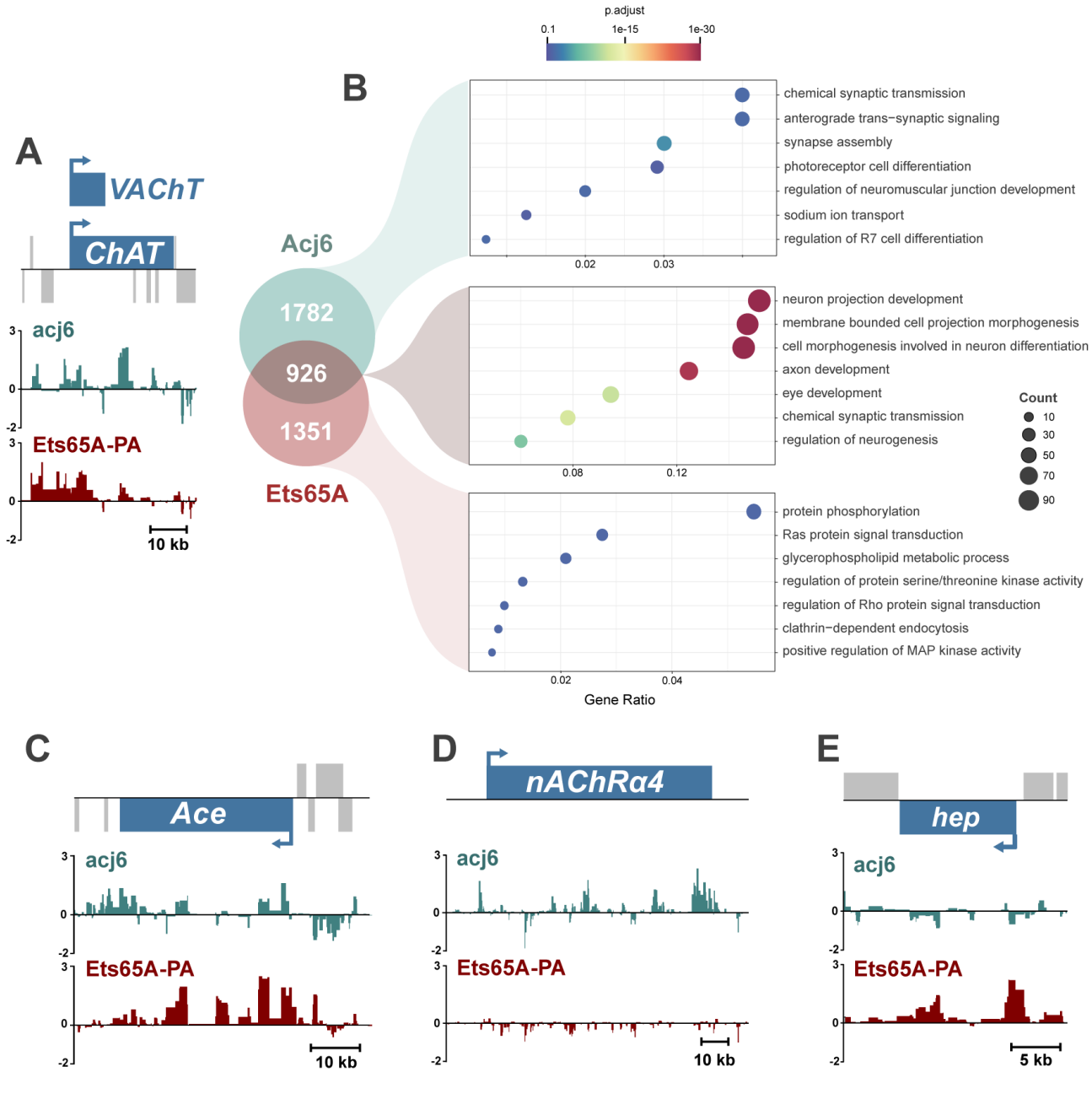

Figure 6. Acj6 and Ets65A-PA co-bind both *ChAT* and a whole suite of genes involved in neuronal differentiation. (A) Acj6 and Ets65-PA binding at *ChAT* (Y-axis represent log2 ratios of Dam-Pol II over Dam-only). (B) Enriched GO term categories for Acj6 and Ets65A-PA bound genes. (C)-(E) Binding at *Ace*, *nAChRα4* and *hep*.

Importantly, we have uncovered and highlighted transcription factors and non-coding RNAs differentially expressed between these subtypes. Some of these are expected based on previous studies in *Drosophila*, including *acj6* (cholinergic) (Lee and Salvaterra, 2002) and *Dbx* (GABAergic) (Lacin et al., 2009). Also, studies in other model organisms fit with our findings, for example,

346 cholinergic enriched *knot*, whose orthologue, *UNC-3* (*C. elegans*), is a terminal selector for cholinergic
347 motor neuron differentiation (Kratsios et al., 2011). In addition, *RFX,* the vertebrate orthologue of *Rfx,*
348 which we identified as glutamatergic enriched, can increase the expression of the neuronal glutamate
349 transporter type 3 (Ma et al., 2006). However, we have identified many differentially expressed
350 transcription factors that have not had their role studied with respect to neurotransmitter
351 specification, or cases where there is supportive, but not direct, evidence for a role in neurotransmitter
352 specification. For instance, vertebrate neuronal precursors expressing *Nkx2.1 (HGTX* orthologue)
353 predominantly generate GABAergic interneurons (Fogarty et al., 2007), and a polyalanine expansion in
354 *ARX* (*hbn* orthologue) causes remodelling and increased activity of glutamatergic neurons in
355 vertebrates (Beguin et al., 2013). *Acj6* is expressed in a subset of cholinergic neurons (Lee and
356 Salvaterra, 2002) and *Dbx* in a subset of GABAergic neurons (Lacin et al., 2009). To the best of our
357 knowledge, none of the enriched transcription factors we identified are expressed in all of the neurons
358 of a particular neurotransmitter subtype. This highlights that, similar to *C. elegans*, there are no simple
359 transcription factor codes for neurotransmitter type specification in *Drosophila*.

360 Uniquely enriched factors are candidates for promoting a neurotransmitter fate, and we tested a
361 number of them for their ability to reprogram neurons on a global scale in embryos (Figure S4A). No
362 obvious changes were observed, however, this is not particularly surprising considering the
363 importance of cellular context for the reprogramming of neuronal properties (Duggan et al., 1998;
364 Wenick and Hobert, 2004). Successful reprograming may require intervention at a specific time point
365 (e.g. at the progenitor stage) and/or the co-expression of appropriate co-factors. Future work could
366 investigate these factors in specific and relevant lineages, to shed light on important contextual
367 information.

368 The majority of transcription factors identified as directly regulating neurotransmitter fate act in a
369 positive manner, whereas only a handful of studies describe the role of repressors. Incoherent
370 feedforward loops exist in *C. elegans*, where terminal selectors activate repressors, which feedback
371 onto effector genes (for review, see (Hobert, 2016)). In vertebrates, both *Neurogenin 2* and *Tlx3* are
372 required for the specification of certain glutamatergic populations but also act to repress GABAergic
373 fate (Cheng et al., 2004; Schuurmans et al., 2004). Whether this is direct repression of *Glutamic acid*
374 *decarboxylase* (*Gad*) genes (required for the synthesis of GABA), or indirectly, through another
375 transcription factor, is unclear. We have identified several transcription factors that are expressed in
376 two neurotransmitter subtypes, but absent from the other. These include *apterous* (*ap*), *Ets65A* (long
377 transcripts) and *orthopedia* (*otp*), which we hypothesise to be candidate repressors, given their
378 absence from cells with a specific neurotransmitter identity. Our profiling of Ets65A-PA binding *in*
379 *vivo*, reveals that it directly binds *ChAT* (Figure 6A), and therefore has the potential to directly regulate
380 cholinergic fate. Similar to the candidate activators, ectopic expression of these candidates did not
381 show any obvious repression of the respective neurotransmitter genes (Figure S4B), however, again,
382 this might be because they can only act as a repressor in specific contexts (e.g. when a co-repressor is

383 present), or that they regulate genes associated with specific subtypes but don't directly regulate
384 neurotransmitter identity.

385 The development of single cell RNA-seq (scRNA-seq) technology has led to the profiling of several
386 *Drosophila* tissues, including the whole adult brain (Davie et al., 2018), the central adult brain (Croset
387 et al., 2018) and the adult optic lobes (Konstantinides et al., 2018). Konstantinides and colleagues
388 investigated the association of transcription factors with neurotransmitter identity, and although it is
389 optic lobe, and not whole brain data, we observe similar trends. For example, they identified *apterous*
390 as cholinergic enriched and *Lim3* as GABAergic enriched, both of which show respective enrichment in
391 our data (1% FDR and 0.15 log2 thresholds). We also identified *apterous* as uniquely enriched in adult
392 brain cholinergic neurons, but also as GABAergic depleted in larval and adult stages (Figure 3), as
393 there is robust Pol II occupancy at *apterous* in glutamatergic neurons (Figure 4C) and co-expression in
394 the brain (Figure 4D), which fits well with the whole brain scRNA-seq (Davie et al., 2018) (Figure S6).
395 *ap*, *Dll* and *svp* were also identified in by Davis and colleagues when they sequenced different neural
396 populations of the adult optic lobe using TAPIN-seq (Davis et al., 2018).

397 We observed more differentially expressed transcription factors in the L3 larval stage (58) compared
398 to the embryo (40) or adults (33). This may reflect the existence of both the functioning larval nervous
399 system (built during embryogenesis) and the developing adult nervous system at this stage (Figure 3).
400 While both the embryo and larval data are similar on a global scale, Pol II occupancy and chromatin
401 accessibility in the adult brain is less correlated (Figure 2). It is currently unclear whether this is due
402 to adult VNCs being absent from the profiling experiments, or differences between immature and fully
403 mature neurons, such as overall lower transcriptional activity in adults. We have previously shown
404 that global chromatin accessibility distribution in adult neurons is distinct from larval neurons
405 (Aughey et al., 2018), which may account for some of these differences.

406 Apart from the neurotransmitter synthesis genes, the chromatin accessibility of the subtypes, at a
407 given stage, is surprisingly similar, as demonstrated in embryos (Figure 6). The enriched accessibility
408 is not just restricted to the gene bodies of the neurotransmitter genes, and peaks are present upstream
409 (*Gad1*) and downstream (*VGlut*) (Figure 6B), which are likely enhancers. Accessibility at the *ChAT* gene
410 is clearly higher in cholinergic neurons at the embryonic and adult stages, however, in L3 larvae, the
411 difference is less pronounced (Figure 6B). This could reflect increased plasticity at this stage, possibly
412 linked to the dramatic remodelling of larval neurons during metamorphosis (for a review, see (Yaniv
413 and Schuldiner, 2016)). While a subset of transcription factors display obvious contrasts in Pol II
414 occupancy, the same transcription factors have no observable, or minor, differences in accessibility
415 (Figure 6C). This could be due to transcription factors being expressed at relatively lower levels
416 and/or that they are only expressed in a subset of the cells, therefore the difference is less prominent.

417 Evidence is emerging for the roles of miRNAs in generating neuronal diversity, including the
418 differentiation of taste receptor neurons in worms (Chang et al., 2004; Johnston and Hobert, 2005) and

419 dopaminergic neurons in vertebrates (Kim et al., 2007). Here, we found the enriched expression of
420 *mir-184* in GABAergic cells (Figure S5B), which is intriguing, as *mir-184* has been shown to
421 downregulate *GABRA3* (GABA-A receptor) mRNA (possibly indirectly) in vertebrate cell lines (Luo et
422 al., 2017), and may be a mechanism to help prevent GABAergic neurons self-inhibiting. Furthermore,
423 *mir-87* has enriched RNA polymerase II occupancy in cholinergic neurons (Figure S5A), and when
424 mutated causes larval locomotion defects in *Drosophila* (Picao-Osorio et al., 2017).

425 *Acj6* is expressed in cholinergic neurons (Figure 3)(Lee and Salvaterra, 2002) , whilst *Ets65A-PA* is
426 expressed in non-cholinergic neurons (Figure 4C and D). However, despite this, they bind a large
427 number of common target genes (Figure 6). This includes 20% (101/493) of all genes annotated for a
428 role in *"neuron projection development"* (GO:0031175). This is quite striking, especially as this is in the
429 adult, where there is virtually no neurogenesis or axonogenesis. However, this may reflect dendritic
430 re-modelling processes, or a requirement of neurons to continuously express transcription factors,
431 even after development, to maintain their fate. The *acj6* orthologues, *unc-86* and *Brn3a* are both
432 required to maintain the fate of specific cholinergic populations (Serrano-Saiz et al., 2018), and
433 transcriptional networks that specific Tv1/Tv4 neurons in *Drosophila* are also required to maintain
434 them in the adult (Eade et al., 2012). Therefore, the binding of Acj6 and Ets65A-PA to developmental
435 genes and *ChAT* in adult neurons could be required for the continued activation (and repression) of
436 genes governing neuronal identity. MAP kinase signalling genes are enriched in cholinergic neurons
437 (Figure 2C) and Ets65A-PA specifically binds MAP kinase signalling genes (Figure 6), making it
438 tempting to speculate that Ets65A-PA acts to repress cholinergic specific genes such as *ChAT* and MAP
439 kinase genes. These Acj6 and Ets65A-PA data also emphasise the diverse set of neuronal
440 differentiation genes a single transcription factor could regulate.

441 The precise synthesis and utilisation of neurotransmitters ensures proper information flow and circuit
442 function in the nervous system. The mechanisms of specification are lineage specific, predominantly
443 through the action of transcription factors. Here we have provided further insights into the
444 complement of different transcription factors that regulate neurotransmitter identity throughout
445 development. Furthermore, we identified the genomic binding of a known activator, and a candidate
446 repressor, of cholinergic fate in the adult, emphasising the broad spectrum of neural identity genes
447 that they could be regulating outside of neurotransmitter use. Given the strong evidence for conserved
448 mechanisms controlling neurotransmitter specification, these data will be a useful resource for not
449 just researchers using *Drosophila* but other other model systems too. Continued work to elucidate the
450 mechanisms, co-factors and temporal windows in which these factors are acting will be fundamental
451 in gaining a comprehensive understanding of neurotransmitter specification.

452

453

454 # Materials and Methods

455 ## *Drosophila* lines

456 Lines used in this study are as follows:

457 *w; dvGlut-GAL4 [MI04979]/CyO act-GFP*, (Bloomington #60312)

458 *w;; ChAT-GAL4 [MI04508] / TM3 act GFP*, (Bloomington #60317)

459 *w;; Gad1-GAL4 [MI09277] / TM3 actin GFP* (Diao et al., 2015)

460 *w[*]; Mi{Trojan-lexA:QFAD.2}VGlut[MI04979-TlexA:QFAD.2]/CyO, P{Dfd-GMR-nvYFP}2*, (Bloomington
461 #60314)

462 *w[*]; Mi{Trojan-lexA:QFAD.0}ChAT[MI04508-TlexA:QFAD.0]/TM6B, Tb[1],* (Bloomington #60319)

463 *w[*]; Mi{Trojan-lexA:QFAD.2}Gad1[MI09277-TlexA:QFAD.2]/TM6B, Tb[1],* (Bloomington, #60324).

464 (All obtained from M. Landgraf)

465 *UAS-LT3-NDam, tub-GAL80^{ts}; UAS-LT3-NDam-RNA Pol II* (from Andrea Brand)

466 *Ets65A-RA/C/D/E-GAL4 [MI07721]* (this study)

467 *apterous-GAL4; UAS-GFP* (from F Jiménez Díaz-Benjumea)

468 *acj6-GAL4-UAS-mCD8-GFP/FM7c; Pin/CyO* (from DJ Luginbuhl) (Lai et al., 2008)

469 *elavG4;; Mi{PT-GFSTF.2}Gad [MI09277]/TM3 actin-GFP* (from Bloomington, #59304)

470 *UAS-Dbx* (Bloomington, #56826)

471 *UAS-apterous* (from F Jiménez Díaz-Benjumea),

472 *UAS-engrailed [E9]* (from Andrea Brand)

473 *UAS-otp* (Fly ORF #F000016)

474 *UAS-CG4328* (FlyORF, #F0019111)

475 *UAS-Dbx sh RNAi attP40* (VDRC #330536)

476 *UAS-ap sh RNAi attP40* (VDRC #330463)

477 *UAS-Ets65A-RA attP2* (this study)

478 *yw, hs-Flp 1; +; Dr/TM6B*

479 *yw, hs-Flp 1; +; UAS-Ets65A-RA*

480 *AyGal4, UAS-mCD8-GFP/(CyO); Cha lexAQF, mCherry /TM6B*

481 ## Generation of *Ets65A* and *acj6* Targeted DamID lines

482 Details and sequences of all primers used for generating constructs are shown in Supplemental
483 Experimental Procedures. *pUAST-LT3-NDam-acj6-RF* and *pUAST-LT3-NDam-Ets65A-RA* were
484 generated by PCR amplifying *acj6-RF* and *Ets65A-RA* from an embryonic cDNA library. The resulting

485 PCR products were cloned into *pUAST-LT3-Dam* plasmid (Southall et al., 2013) with NotI and XhoI
486 sites, using Gibson assembly.

487 acj6-RF FW:
488 CATCTCTGAAGAGGATCTGGCCGGCGCAGATCTGCGGCCGCTCATGACAATGTCGATGTATTCGACGACGG

489 acj6-RF RV:
490 GTCACACCACAGAAGTAAGGTTCCTTCACAAAGATCCTCTAGATCAGTATCCAAATCCCGCCGAACCG
491 Ets65A-RA FW:

492 CTGAAGAGGATCTGGCCGGCGCAGATCTGCGGCCGCTCATGTACGAGAACTCCTGTTCGTATCAGACG

493 Ets65A-RA RV:

494 ACAGAAGTAAGGTTCCTTCACAAAGATCCTCTAGATCATGCGTAGTGGGGATAGCTGCTC

495 Generation of *Ets65A-RA-GAL4* line

496 *Ets65A-RA/C/D/E-GAL4* was generated by inserting a GAL4 trap cassette into the MI07721 MiMIC line
497 (Bloomington #43913) using the triplet donor *in vivo* system described in (Diao et al., 2015).

498 Targeted DamID (TaDa) for RNA-Pol II mapping

499 Crosses producing larvae with the following genotypes were allowed to lay eggs over a minimum of
500 two days at 25˚C before timed collections were performed:

501 *tub-GAL80^{ts}/+; UAS-LT3-NDam/ ChAT-GAL4^{MI04508}*

502 *tub-GAL80^{ts}/+; UAS-LT3-NDam-RNA Pol II/ ChAT-GAL4^{MI04508}*

503 *tub-GAL80^{ts}/+; UAS-LT3-NDam/ Gad1-GAL4 ^{MI09277}*

504 *tub-GAL80^{ts}/+; UAS-LT3-NDam-RNA Pol II/ Gad1-GAL4 ^{MI09277}*

505 *tub-GAL80^{ts}/ dvGlut-GAL4 ^{MI04979}; UAS-LT3-NDam/ +*

506 *tub-GAL80^{ts}/ dvGlut-GAL4 ^{MI04979}; UAS-LT3-NDam-RNA Pol II/ +*

507 **First instar larvae samples:** Crosses of the right genotype were allowed to lay old eggs for 2 hours at
508 25˚C in fly cages. Wet yeast and two drops of 10% acetic acid were added to apple juice plates to
509 promote egg laying. Then, egg laying was done for 5 hours at 25˚C, apple juice plates containing those
510 embryos were transferred to 29˚C (permissive temperature) for 20 hours. After this time, first instar
511 larvae were collected and stored in 1x PBS. Samples were flash-frozen in dry ice, and stored at -80˚C
512 till the appropriate amount of tissue was enough to start the experiment. No selection for the right
513 genotype was done, and 20 µl worth of volume of tissue was used as a proxy to determine the
514 appropriate amount of material for each replicate. 3 replicates were done for each experiment. With
515 this husbandry protocol, the collected first instar larvae were around 12 hours ALH (after larvae
516 hatching), just before the first larval neurons are being generated, then providing the transcriptome of
517 embryonic neurogenesis.

518 **Third instar larvae samples:** Crosses of the right genotypes were allowed to lay eggs for 6 hours at
519 25˚C in fly food vials. These vials were then transferred to 18˚C (restrictive temperature) for 7 days.
520 They were then moved to 29˚C (permissive temperature) for 24 hours. Wandering stage larvae,
521 around 96h ALH, were selected with a GFP scope for the right genotype. Larvae were dissected in 1x
522 PBS, leaving the anterior half of the larvae partly dissected, containing the CNS, but removing the gut
523 and all the fat tissue. Samples were flash-frozen in dry ice, and stored at -80˚C till the appropriate
524 amount of tissue was enough to start the experiment. 100 partly dissected CNS were used for each
525 replicate. 3 replicates were done for each experiment.

526 **Adult samples:** Crosses of the right genotypes were allowed to lay eggs for 2 days at 18˚C in fly food
527 vials. Vials containing those eggs were kept at 18˚C (restrictive temperature) till adult flies eclosed.
528 They were then kept at 18˚C for 5-10 days. After that, they were selected according for the right
529 genotype, and transferred to 29˚C (permissive temperature) for 24 hours. Then, adult flies were flash-
530 frozen in dry ice, and stored at -80˚C. Around 50 fly heads were used for each replicate. 3 replicates
531 were done for each experiment.

532 When preparing the tissue to be used, larvae nor flies were sex sorted. It has been recently reported
533 that transcriptomes from males and females, obtained with a cell specific driver combination
534 expressed in neurons in the adult optic lobe, do not present major differences in their transcriptomes.
535 Only a small number of genes, known sex-specific genes showed differences between sexes. (Davis et
536 al., 2018).

537 Our DamID protocol was based on (Southall et al., 2013), and (Marshall et al., 2016). Briefly, DNA was
538 extracted using Qiagen DNeasy kit, and a minimum of 3 µg of DNA was precipitated for first instar
539 larvae, 6ug for third instar larvae, and 2.5 µg for adult samples. DNA was digested with DpnI overnight
540 at 37˚C. The next morning, 0.5 µl of DpnI was added for 1 hour extra incubation, followed by DpnI heat
541 inactivation (20 mins 80˚C). Either Advantage cDNA polymerase, or Advantage 2 cDNA polymerase
542 mix, 50x, Clontech, were used in PCR amplification. Enzymes Sau3AI or AlwI were used to remove
543 DamID adaptors, from sonicated DNA.

544 Libraries were sequenced using Illumina HiSeq single-end 50 bp sequencing. Three replicates were
545 performed for each experiment. A minimum of 25 million reads were obtained from the first instar
546 larvae samples, 30 million reads from the third instar larvae, and 9 million reads from the adults'
547 samples.

548 **Targeted DamID (TaDa) for identification of acj6 and Ets65A-RA binding sites**

549 Crosses producing larvae with the following genotypes were allowed to lay eggs over a minimum of
550 two days at 25˚C:

551 *acj6-GAL4-UAS-GFP; tub-GAL80^{ts}/+; UAS-LT3-NDam/+*

552 *acj6-GAL4-UAS-GFP; tub-GAL80^{ts}/+; UAS-LT3-NDam-acj6-RF/+*

553 *tub-GAL80^{ts}/+; UAS-LT3-NDam/ Ets65A-RA-GAL4^{MI07721}*

554 *tub-GAL80^{ts}/+; UAS-LT3-NDam-Ets65A-RA/ Ets65A-RA-GAL4^{MI07721}*

555 Crosses of the right genotypes were allowed to lay eggs for 2 days at 18˚C in fly food vials. Vials
556 containing those eggs were kept at 18˚C (restrictive temperature) till adult flies eclosed. They were
557 then kept at 18˚C for around 10 days. Then, they were transferred to 29˚C (permissive temperature)
558 for 24 hours, selected according for the right genotype, flash-frozen in dry ice, and stored at -80˚C. A
559 minimum of 150 fly heads were used for each replicate. 2 replicates were done for each experiment.

560 The DamID protocol used for these samples is the same as described above, with minor changes, 6 µg
561 of DNA were precipitated, Bioline Polymerase was used in the PCR amplification, and only AlwI was
562 used to remove adaptors from sonicated DNA. Libraries were sequenced using Illumina HiSeq single-
563 end 50 bp sequencing. Two replicates were acquired for each experiment. A minimum of 10 million
564 reads were obtained these samples.

565 **Targeted DamID data analysis**

566 Sequencing data for TaDa and CATaDa were mapped back to release 6.03 of the *Drosophila* genome
567 using a previously described pipelines (Marshall and Brand, 2015; Aughey et al., 2018). Transcribed
568 genes (defined by Pol II occupancy) were identified using a Perl script described in (Mundorf et al.,
569 2019) based on one developed by (Southall et al. 2013) (available at
570 https://github.com/tonysouthall/Dam-RNA_POLII_analysis). *Drosophila* genome annotation release
571 6.11 was used, with 1% FDR and 0.2 log2 ratio thresholds. To compare data sets, log2 ratios were
572 subtracted, in this case, producing 3 replicate comparison files (as 3 biological replicates were
573 performed). These data were then analysed as described above to identify genes with significantly
574 different Pol II occupancy. Due to the presence of negative log2 ratios in DamID experiments, these
575 genes were filtered to check that any significantly enriched genes were also bound by Pol II in the
576 experiment of interest (numerator data set). A gene list was generated from the transcript data using
577 the values from the associated transcript with the most significant FDR. Correlation values (Figure 2)
578 were visualised using Morpheus (https://software.broadinstitute.org/morpheus/). The transition plot
579 (Figure 3) was generated in R using the transitionPlot function from the Gmisc R package
580 (http://gforge.se/). Enrichment GO analysis was performed using the R package clusterProfiler (Yu et
581 al., 2012)

582 For acj6 and Ets65A-RA TaDa, peaks were called and mapped to genes using a custom Perl program
583 (available at https://github.com/tonysouthall/Peak_calling_DamID) In brief, a false discovery rate
584 (FDR) was calculated for peaks (formed of two or more consecutive GATC fragments) for the
585 individual replicates. Then each potential peak in the data was assigned a FDR. Any peaks with less
586 than a 0.01% FDR were classified as significant. Significant peaks present in all replicates were used to

form a final peak file. Any gene (genome release 6.11) within 5 kb of a peak (with no other genes in between) was identified as a potentially regulated gene.

Immunostaining and imaging

Third instar larval CNS or adult brains were dissected in 1x PBS. They were fixed in 4% formaldehyde (methanol free) 0.1% Triton X-100 PBS (PBST), for 30 minutes at room temperature. Samples were then rinsed twice with 0.1% PBST, and washed four times for 1 hour with 0.1% PBST. 5% Normal Goat Serum in 0.1% PBST was used as a blocking agent for 1 hour at room temperature. Brains were then incubated overnight at 4˚C with primary antibodies in 5% Normal Goat Serum in 0.1% PBST. The primary antibodies used were: anti-Chicken-GFP (Abcam #13970, 1:2000), and anti-Rabbit-DsRed (Clontech #632496, 1:500). Brains were rinsed twice with 0.1% PBST, and washed four times with 0.1% PBST for 1 hour. Secondary antibodies were diluted in 5% Normal Goat Serum in 0.1% PBST and incubated with the brains for 1 hour at room temperature. The secondary antibodies used were: anti-Chicken-Alexa 488 (Thermo Scientific #A11039, 1:500), and anti-Rabbit-Alexa 546 (Thermo Scientific #A11010, 1:500). Samples were then rinsed twice with 0.1% PBST, and washed four times for 1 hour. Brains were mounted on glass cover slides in Vectashield (Vector laboratories). All incubations and washes were performed in a rotator. After dissection of first instar larvae CNS, they were placed in a polylisine coated microscope slide, where we performed all the incubations. Both experimental CNS and wild-type CNS were placed on the same slide. Images were acquired using a Zeiss LSM 510 confocal microscope and edited using Fiji/Image J. For the Ets65A-RA overexpression clones, *AyGal4, UAS-mCD8-GFP/(CyO); Cha lexAQF, mCherry /TM6B* was crossed with *UAS-Ets65A-RA*, and adults were heat-shocked for 1 hour at 37^{0}C, at both 24 hours ALH (after larvae hatching) and 48 hours ALH. Brains were dissected from 5-8 days old adults.

Accession numbers

All raw sequence files and processed files have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (data uploaded 4/11/19 – should be accessible by 11/11/19).

Acknowledgements

We would like to thank Matthias Landgraf, Eva Higginbotham, and Gabriel Aughey for feedback and advice on this project. We would like to thank Matthias Landgraf, Holly Ironfield, Benjamin White, Andrea Brand, Fernando J. Diaz-Benjumea, David John Luginbuhl, for providing fly stocks. For other fly stocks, we also thank Bloomington *Drosophila* Stock Center (NIH P40OD018537), the Vienna *Drosophila* Resource Center (VDRC, www.vdrc.at), and the FlyORF, Zurich ORFeome Project (https://flyorf.ch/). This work was funded by Wellcome Trust Investigator grant 104567 to T.D.S.

621 # Figure Legends

622 **Figure 1. Cell specific profiling of RNA Pol II occupancy in different neuronal subtypes**
623 **throughout Drosophila development.** Profiling of RNA Pol II occupancy in cholinergic, GABAergic
624 and glutamatergic neurons using TaDa. Profling windows cover embryonic nervous system
625 development (5 - 25 hr AEL), 3rd instar larval nervous system development (24 hr window before
626 pupation) and the adult brain (heads from ~ 3-4 day old adults after a 24 hr expression window).
627 Temporal restriction of Dam-Pol II expression was controlled using a temperature sensitive GAL80.
628 Bottom panels show an example of a transcription factor (*Dbx*) that is uniquely transcribed in
629 GABergic neurons. Y-axis represent log2 ratios of Dam-Pol II over Dam-only.

630 **Figure 2. Correlation of RNA Pol II occupancy and chromatin accessibility for neurotransmitter**
631 **subtypes.** (A) Correlation matrix for RNA Pol II signal (log2 over Dam-only for all genes) and
632 chromatin accessibility (CATaDa) for all gene loci (extended 5 kb upstream and 2 kb downstream). (B)
633 Correlation matrix for RNA Pol II and chromatin accessibility at annotated transcription factor genes.
634 (C) Characterisation of subtype-enriched genes at each developmental stage. Examples of enriched GO
635 term categories for the remaining genes are also included.

636 **Figure 3. Transitions of uniquely enriched or depleted transcription factors during neural**
637 **development.** Transcription factors uniquely enriched or depleted in cholinergic, GABAergic and
638 glutamatergic neurons. A total of 86 transcription factors are identified across all stages.

639 **Figure 4. Absence of *Ets65A-RA/C/D/E* and *apterous* expression in cholinergic and GABAergic**
640 **adult neurons, respectively.** (A) Pol II occupancy at *apterous* in the adult brain. Y-axis represent log2
641 ratios of Dam-Pol II over Dam-only. (B) Expression pattern of *apterous* in the adult brain. (C) Pol II
642 occupancy at Ets65A in the adult brain. Y-axis represent log2 ratios of Dam-Pol II over Dam-only. (D)
643 Expression pattern of *Ets65A-RA/C/D/E* in the adult brain.

644 **Figure 5. Evidence for predicted roles of identified transcription factors and their orthologues.**
645 Strongly enriched or depleted transcription factors identified in developing larval brains. Uniquely
646 enriched factors are predicted to be candidates that promote the respective neurotransmitter fate,
647 whilst uniquely depleted factors are predicted to repress the neurotransmitter fate. A full tick
648 indicates direct evidence that the transcription factor directly promotes or represses the
649 neurotransmitter fate, whilst indirect, supporting evidence is indicated by faded tick in brackets. A
650 question mark signifies that nothing is currently known, regarding neurotransmitter specification.

651 **Figure 6. Acj6 and Ets65A-PA co-bind both *ChAT* and a whole suite of genes involved in**
652 **neuronal differentiation.** (A) Acj6 and Ets65-PA binding at ChAT (Y-axis represent log2 ratios of
653 Dam-Pol II over Dam-only). (B) Enriched GO term categories for Acj6 and Ets65A-PA bound genes. (C)-
654 (E) Binding at *Ace*, *nAChRα4* and *hep*.

References

Aguilar, J.I., Dunn, M., Mingote, S., Karam, C.S., Farino, Z.J., Sonders, M.S., Choi, S.J., Grygoruk, A., Zhang, Y., Cela, C., Choi, B.J., Flores, J., Freyberg, R.J., McCabe, B.D., Mosharov, E.V., Krantz, D.E., Javitch, J.A., Sulzer, D., Sames, D., Rayport, S., Freyberg, Z., (2017). Neuronal Depolarization Drives Increased Dopamine Synaptic Vesicle Loading via VGLUT. *Neuron* **95**, 1074-1088 e1077.

Altun-Gultekin, Z., Andachi, Y., Tsalik, E.L., Pilgrim, D., Kohara, Y., Hobert, O., (2001). A regulatory cascade of three homeobox genes, ceh-10, ttx-3 and ceh-23, controls cell fate specification of a defined interneuron class in C. elegans. *Development* **128**, 1951-1969.

Andreae, L.C., Burrone, J., (2018). The role of spontaneous neurotransmission in synapse and circuit development. *J Neurosci Res* **96**, 354-359.

Aughey, G.N., Cheetham, S.W., Southall, T.D., (2019). DamID as a versatile tool for understanding gene regulation. *Development* **146**.

Aughey, G.N., Estacio Gomez, A., Thomson, J., Yin, H., Southall, T.D., (2018). CATaDa reveals global remodelling of chromatin accessibility during stem cell differentiation in vivo. *Elife* **7**.

Beguin, S., Crepel, V., Aniksztejn, L., Becq, H., Pelosi, B., Pallesi-Pocachard, E., Bouamrane, L., Pasqualetti, M., Kitamura, K., Cardoso, C., Represa, A., (2013). An epilepsy-related ARX polyalanine expansion modifies glutamatergic neurons excitability and morphology without affecting GABAergic neurons development. *Cereb Cortex* **23**, 1484-1494.

Chang, S., Johnston, R.J., Jr., Frokjaer-Jensen, C., Lockery, S., Hobert, O., (2004). MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode. *Nature* **430**, 785-789.

Cheng, L., Arata, A., Mizuguchi, R., Qian, Y., Karunaratne, A., Gray, P.A., Arata, S., Shirasawa, S., Bouchard, M., Luo, P., Chen, C.L., Busslinger, M., Goulding, M., Onimaru, H., Ma, Q., (2004). Tlx3 and Tlx1 are post-mitotic selector genes determining glutamatergic over GABAergic cell fates. *Nat Neurosci* **7**, 510-517.

Coleman, B., Topalidou, I., Ailion, M., (2018). Modulation of Gq-Rho Signaling by the ERK MAPK Pathway Controls Locomotion in Caenorhabditis elegans. *Genetics* **209**, 523-535.

Croset, V., Treiber, C.D., Waddell, S., (2018). Cellular diversity in the Drosophila midbrain revealed by single-cell transcriptomics. *Elife* **7**.

Davie, K., Janssens, J., Koldere, D., De Waegeneer, M., Pech, U., Kreft, L., Aibar, S., Makhzami, S., Christiaens, V., Bravo Gonzalez-Blas, C., Poovathingal, S., Hulselmans, G., Spanier, K.I., Moerman, T., Vanspauwen, B., Geurs, S., Voet, T., Lammertyn, J., Thienpont, B., Liu, S., Konstantinides, N.,

688 **Fiers, M., Verstreken, P., Aerts, S.**, (2018). A Single-Cell Transcriptome Atlas of the Aging Drosophila
689 Brain. *Cell* **174**, 982-998 e920.

690 **Davis, F.P., Nern, A., Picard, S., Reiser, M.B., Rubin, G.M., Eddy, S.R., Henry, G.L.**, (2018). A genetic,
691 genomic, and computational resource for exploring neural circuit function. *bioRxiv*, 385476.

692 **Diao, F., Ironfield, H., Luan, H., Diao, F., Shropshire, W.C., Ewer, J., Marr, E., Potter, C.J., Landgraf,**
693 **M., White, B.H.**, (2015). Plug-and-play genetic access to drosophila cell types using exchangeable exon
694 cassettes. *Cell Rep* **10**, 1410-1421.

695 **Duggan, A., Ma, C., Chalfie, M.**, (1998). Regulation of touch receptor differentiation by the
696 Caenorhabditis elegans mec-3 and unc-86 genes. *Development* **125**, 4107-4119.

697 **Eade, K.T., Fancher, H.A., Ridyard, M.S., Allan, D.W.**, (2012). Developmental transcriptional
698 networks are required to maintain neuronal subtype identity in the mature nervous system. *PLoS*
699 *Genet* **8**, e1002501.

700 **Flames, N., Hobert, O.**, (2009). Gene regulatory logic of dopamine neuron differentiation. *Nature* **458**,
701 885-889.

702 **Fogarty, M., Grist, M., Gelman, D., Marin, O., Pachnis, V., Kessaris, N.**, (2007). Spatial genetic
703 patterning of the embryonic neuroepithelium generates GABAergic interneuron diversity in the adult
704 cortex. *J Neurosci* **27**, 10935-10946.

705 **Goold, C.P., Davis, G.W.**, (2007). The BMP ligand Gbb gates the expression of synaptic homeostasis
706 independent of synaptic growth control. *Neuron* **56**, 109-123.

707 **Gummalla, M., Maeda, R.K., Castro Alvarez, J.J., Gyurkovics, H., Singari, S., Edwards, K.A., Karch,**
708 **F., Bender, W.**, (2012). abd-A regulation by the iab-8 noncoding RNA. *PLoS Genet* **8**, e1002720.

709 **Hobert, O.**, (2008). Regulatory logic of neuronal diversity: terminal selector genes and selector motifs.
710 *Proc Natl Acad Sci U S A* **105**, 20067-20071.

711 **Hobert, O.**, (2016). A map of terminal regulators of neuronal identity in Caenorhabditis elegans. *Wiley*
712 *Interdiscip Rev Dev Biol* **5**, 474-498.

713 **Jin, Y., Hoskins, R., Horvitz, H.R.**, (1994). Control of type-D GABAergic neuron differentiation by C.
714 elegans UNC-30 homeodomain protein. *Nature* **372**, 780-783.

715 **Johnston, R.J., Jr., Hobert, O.**, (2005). A novel C. elegans zinc finger transcription factor, lsy-2,
716 required for the cell type-specific expression of the lsy-6 microRNA. *Development* **132**, 5451-5460.

717 **Kim, J., Inoue, K., Ishii, J., Vanti, W.B., Voronov, S.V., Murchison, E., Hannon, G., Abeliovich, A.**,
718 (2007). A MicroRNA feedback circuit in midbrain dopamine neurons. *Science* **317**, 1220-1224.

Kim, N.C., Marques, G., (2012). The Ly6 neurotoxin-like molecule target of wit regulates spontaneous neurotransmitter release at the developing neuromuscular junction in Drosophila. *Dev Neurobiol* **72**, 1541-1558.

Konstantinides, N., Kapuralin, K., Fadil, C., Barboza, L., Satija, R., Desplan, C., (2018). Phenotypic Convergence: Distinct Transcription Factors Regulate Common Terminal Features. *Cell* **174**, 622-635 e613.

Kratsios, P., Stolfi, A., Levine, M., Hobert, O., (2011). Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene. *Nat Neurosci* **15**, 205-214.

Lacin, H., Chen, H.M., Long, X., Singer, R.H., Lee, T., Truman, J.W., (2019). Neurotransmitter identity is acquired in a lineage-restricted manner in the Drosophila CNS. *Elife* **8**.

Lacin, H., Zhu, Y., Wilson, B.A., Skeath, J.B., (2009). dbx mediates neuronal specification and differentiation through cross-repressive, lineage-specific interactions with eve and hb9. *Development* **136**, 3257-3266.

Lai, S.L., Awasaki, T., Ito, K., Lee, T., (2008). Clonal analysis of Drosophila antennal lobe neurons: diverse neuronal architectures in the lateral neuroblast lineage. *Development* **135**, 2883-2893.

Landgraf, M., Bossing, T., Technau, G.M., Bate, M., (1997). The origin, location, and projections of the embryonic abdominal motorneurons of Drosophila. *J Neurosci* **17**, 9642-9655.

Lee, M.H., Salvaterra, P.M., (2002). Abnormal chemosensory jump 6 is a positive transcriptional regulator of the cholinergic gene locus in Drosophila olfactory neurons. *J Neurosci* **22**, 5291-5299.

Luo, Y., Liu, S., Yao, K., (2017). Transcriptome-wide Investigation of mRNA/circRNA in miR-184 and Its r.57c > u Mutant Type Treatment of Human Lens Epithelial Cells. *Mol Ther Nucleic Acids* **7**, 71-80.

Ma, K., Zheng, S., Zuo, Z., (2006). The transcription factor regulatory factor X1 increases the expression of neuronal glutamate transporter type 3. *J Biol Chem* **281**, 21250-21255.

Marshall, O.J., Southall, T.D., Cheetham, S.W., Brand, A.H., (2016). Cell-type-specific profiling of protein-DNA interactions without cell isolation using targeted DamID with next-generation sequencing. *Nat Protoc* **11**, 1586-1598.

Patel, T., Hobert, O., (2017). Coordinated control of terminal differentiation and restriction of cellular plasticity. *Elife* **6**.

Picao-Osorio, J., Lago-Baldaia, I., Patraquim, P., Alonso, C.R., (2017). Pervasive Behavioral Effects of MicroRNA Regulation in Drosophila. *Genetics* **206**, 1535-1548.

Pipes, G.C., Lin, Q., Riley, S.E., Goodman, C.S., (2001). The Beat generation: a multigene family encoding IgSF proteins related to the Beat axon guidance molecule in Drosophila. *Development* **128**, 4545-4552.

Schmid, A., Chiba, A., Doe, C.Q., (1999). Clonal analysis of Drosophila embryonic neuroblasts: neural cell types, axon projections and muscle targets. *Development* **126**, 4653-4689.

Schuurmans, C., Armant, O., Nieto, M., Stenman, J.M., Britz, O., Klenin, N., Brown, C., Langevin, L.M., Seibt, J., Tang, H., Cunningham, J.M., Dyck, R., Walsh, C., Campbell, K., Polleux, F., Guillemot, F., (2004). Sequential phases of cortical specification involve Neurogenin-dependent and -independent pathways. *EMBO J* **23**, 2892-2902.

Serrano-Saiz, E., Leyva-Diaz, E., De La Cruz, E., Hobert, O., (2018). BRN3-type POU Homeobox Genes Maintain the Identity of Mature Postmitotic Neurons in Nematodes and Mice. *Curr Biol* **28**, 2813-2823 e2812.

Serrano-Saiz, E., Poole, R.J., Felton, T., Zhang, F., De La Cruz, E.D., Hobert, O., (2013). Modular control of glutamatergic neuronal identity in C. elegans by distinct homeodomain proteins. *Cell* **155**, 659-673.

Soba, P., Zhu, S., Emoto, K., Younger, S., Yang, S.J., Yu, H.H., Lee, T., Jan, L.Y., Jan, Y.N., (2007). Drosophila sensory neurons require Dscam for dendritic self-avoidance and proper dendritic field organization. *Neuron* **54**, 403-416.

Southall, T.D., Gold, K.S., Egger, B., Davidson, C.M., Caygill, E.E., Marshall, O.J., Brand, A.H., (2013). Cell-type-specific profiling of gene expression and chromatin binding without cell isolation: assaying RNA Pol II occupancy in neural stem cells. *Dev Cell* **26**, 101-112.

Trudeau, L.E., El Mestikawy, S., (2018). Glutamate Cotransmission in Cholinergic, GABAergic and Monoamine Systems: Contrasts and Commonalities. *Front Neural Circuits* **12**, 113.

van Steensel, B., Henikoff, S., (2000). Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat Biotechnol* **18**, 424-428.

Waite, M.R., Skidmore, J.M., Billi, A.C., Martin, J.F., Martin, D.M., (2011). GABAergic and glutamatergic identities of developing midbrain Pitx2 neurons. *Dev Dyn* **240**, 333-346.

Wenick, A.S., Hobert, O., (2004). Genomic cis-regulatory architecture and trans-acting regulators of a single interneuron-specific gene battery in C. elegans. *Dev Cell* **6**, 757-770.

Westmoreland, J.J., McEwen, J., Moore, B.A., Jin, Y., Condie, B.G., (2001). Conserved function of Caenorhabditis elegans UNC-30 and mouse Pitx2 in controlling GABAergic neuron differentiation. *J Neurosci* **21**, 6810-6819.

Yaniv, S.P., Schuldiner, O., (2016). A fly's view of neuronal remodeling. *Wiley Interdiscip Rev Dev Biol* **5**, 618-635.

Yu, G., Wang, L.G., Han, Y., He, Q.Y., (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287.

Zhang, F., Bhattacharya, A., Nelson, J.C., Abe, N., Gordon, P., Lloret-Fernandez, C., Maicas, M., Flames, N., Mann, R.S., Colon-Ramos, D.A., Hobert, O., (2014). The LIM and POU homeobox genes ttx-3 and unc-86 act as terminal selectors in distinct cholinergic and serotonergic neuron types. *Development* **141**, 422-435.