

A Resource for Manipulating Gene Expression and Analyzing *cis*-Regulatory Modules in the *Drosophila* CNS

Laurina Manning,¹ Ellie S. Heckscher,¹ Maria D. Purice,¹ Jourdain Roberts,¹ Alysha L. Bennett,¹ Jason R. Kroll,¹ Jill L. Pollard,¹ Marie E. Strader,¹ Josh R. Lupton,¹ Anna V. Dyukareva,¹ Phuong Nam Doan,¹ David M. Bauer,¹ Allison N. Wilbur,¹ Stephanie Tanner,¹ Jimmy J. Kelly,¹ Sen-Lin Lai,¹ Khoa D. Tran,¹ Minoree Kohwi,¹ Todd R. Lavery,² Joseph C. Pearson,³ Stephen T. Crews,³ Gerald M. Rubin,² and Chris Q. Doe^{1,*}

¹Institutes of Neuroscience and Molecular Biology, Howard Hughes Medical Institute, University of Oregon, Eugene, OR 97403, USA

²Janelia Farm Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA

³Program in Molecular Biology and Biophysics, Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

*Correspondence: cdoe@uoregon.edu

<http://dx.doi.org/10.1016/j.celrep.2012.09.009>

SUMMARY

Here, we describe the embryonic central nervous system expression of 5,000 GAL4 lines made using molecularly defined *cis*-regulatory DNA inserted into a single *attP* genomic location. We document and annotate the patterns in early embryos when neurogenesis is at its peak, and in older embryos where there is maximal neuronal diversity and the first neural circuits are established. We note expression in other tissues, such as the lateral body wall (muscle, sensory neurons, and trachea) and viscera. Companion papers report on the adult brain and larval imaginal discs, and the integrated data sets are available online (<http://www.janelia.org/gal4-gen1>). This collection of embryonically expressed GAL4 lines will be valuable for determining neuronal morphology and function. The 1,862 lines expressed in small subsets of neurons (<20/segment) will be especially valuable for characterizing interneuronal diversity and function, because although interneurons comprise the majority of all central nervous system neurons, their gene expression profile and function remain virtually unexplored.

INTRODUCTION

All model organisms can benefit from tools that allow targeted gene expression. In mouse, collections of Cre lines allow floxed or flox-stop transgenes to be expressed or inactivated, respectively (reviewed in Branda and Dymecki, 2004). In *Drosophila* and zebrafish, the yeast GAL4/UAS system has become a standard method for generating cell-type-specific gene expression (reviewed in Baier and Scott, 2009; Scott, 2009; Scott et al., 2007; Venken and Bellen, 2012). *Drosophila* is a powerful model system for studying cell biology, development, and neuroscience, primarily because of its facile genetics. Over the past

two decades, genetic screens have identified key regulators of highly conserved cellular processes (e.g., cell-cycle progression, cell death, and cell migration), developmental events (e.g., Hox function, cellular memory, and cell-cell signaling pathways), and neurobiological processes (e.g., axon pathfinding, learning and memory, and behavior).

We are interested in characterizing the generation of cellular diversity within the embryonic central nervous system (CNS) and linking neuronal development with function and behavior. The CNS starts to develop at stage 8 with the formation of specialized ventral midline cells, and at stage 9 the bilateral neuroectoderm generates an array of ~30 neural progenitors, called neuroblasts (NBs) (Doe, 2008; Skeath and Thor, 2003). All ventral midline progenitors and NBs are stereotyped and individually identifiable (Broadus et al., 1995; Wheeler et al., 2006). Each NB divides asymmetrically to generate smaller ganglion mother cells (GMCs) and a self-renewed NB (reviewed in Skeath and Thor, 2003). Each GMC either divides once to generate a pair of sibling neurons (Pearson and Doe, 2003), divides more than once to generate a small number of glia (Akiyama-Oda et al., 1999), or directly differentiates into a neuron (Baumgardt et al., 2009). By stage 16, NB divisions have largely ceased and each segment of the CNS typically contains ~22 midline cells (Wheeler et al., 2006) and ~335 bilateral cells (~35 motoneurons, ~35 glia, and ~260 interneurons or unknown cell types; Beckervordersandforth et al., 2008; Landgraf et al., 1999; E.S.H. and C.Q.D., unpublished data).

Embryonic motoneurons are relatively well characterized. Most can be identified by the presence of nuclear phosphorylated Mothers against Dpp (pMad) protein or Zfh1 protein (Layden et al., 2006), and subsets can be identified that express the *even-skipped* (*eve*) gene, encoding a transcription factor that promotes axon targeting to dorsal muscle groups (Fujioka et al., 2003; Landgraf et al., 1999), or *hb9* (Flybase: *exex*), which encodes a transcription factor that promotes axon targeting to ventral muscles (Broihier and Skeath, 2002; Odden et al., 2002). GAL4 lines that are expressed in subsets of the *Eve*+ neurons (Fujioka et al., 1999) and the entire population of *Hb9*+ neurons (Broihier and Skeath, 2002) are available. Similarly, glial

cells can be identified based on position, molecular markers, and GAL4 line expression (Beckervordersandforth et al., 2008). Despite the work of many investigators over the past few decades, very little is known about the majority of interneurons within the CNS. There are few molecular markers for interneuronal subsets and even fewer GAL4 lines with well-characterized expression in subsets of interneurons. This is a barrier to understanding how interneurons develop and how they participate in neural circuits regulating larval behavior.

A powerful tool for characterizing neuronal morphology and function in *Drosophila* is the GAL4/UAS system (Brand and Perrimon, 1993). GAL4 transgenes with distinct patterns have been made by random insertion of constructs containing a basal promoter and the GAL4 coding sequence, a variation of the enhancer-trapping method (Bellen et al., 1989; Bier et al., 1989; O’Kane and Gehring, 1987) that has been widely used to identify cell-type-specific enhancers and their associated genes. The growing collection of *Drosophila* GAL4 transgenes is useful for marking neuronal cell types or expressing upstream activating sequence RNA interference (UAS-RNAi) constructs to assay gene function (see above), and in the Mosaic Analysis with Repressible Cell Marker (MARCM) system to generate positively-marked homozygous mutant neuronal clones (Lee and Luo, 2001). These are powerful methods, but they all have limitations. First, most GAL4 lines have been generated by random insertion of a basal GAL4 transgene randomly within the genome (enhancer traps), which makes it difficult to generate additional lines with the same pattern (e.g., LexA versions of the line) and impossible to perform bioinformatic comparisons of the DNA sequences that confer different GAL4 expression patterns. Second, relatively few GAL4 lines are expressed in small subsets of the CNS, which makes it impossible to characterize the morphology and function of the vast majority of neurons.

Here we describe the embryonic CNS expression of 5,000 GAL4 lines made by inserting molecularly defined *cis*-regulatory DNA into a single *attP* genomic location using the PhiC31 integrase system (Bischof et al., 2007; Groth et al., 2004). This will facilitate the production of non-GAL4 drivers that can be used in combination with GAL4 lines. It will also make it possible to use bioinformatics to look for *cis*-regulatory elements that are shared by GAL4 lines with coexpression in one or more neurons. We describe the embryonic CNS expression at stages 9–11 (when neurogenesis is maximal) and stage 16 (when neuronal diversity is maximal). We comprehensively annotate expression patterns in the CNS midline cells, NBs, and neurons at both stages. We note expression in other embryonic tissues, such as the lateral body wall (e.g., muscle, sensory neurons, and trachea) and viscera. Companion papers report on the expression of the same collection of GAL4 lines in the adult brain (see the accompanying paper by Jenett et al., 2012, in this issue of *Cell Reports*) and imaginal discs (see the accompanying paper by Jory et al., 2012, in this issue of *Cell Reports*), and all three data sets are publicly available online (<http://www.janelia.org/gal4-gen1>). The embryonic data derived from this collection of GAL4 lines will be valuable for determining the morphology and function of all neurons in the embryonic CNS, and especially for shedding light on the function of interneurons, which have

only begun to be defined although they comprise the majority of all neurons in the CNS.

RESULTS AND DISCUSSION

Overview of Embryonic Patterns and Database Features

We stained >5,500 enhancer-gal4 lines generated at Janelia Farm Research Campus (JFRC) to determine their pattern of expression during embryogenesis (Table S1), and acquired interpretable images for 5,000 lines representing 796 genes. The method used to generate the lines was described previously (Pfeiffer et al., 2008). Each line contains an average of 3 kb *cis*-regulatory DNA from intergenic and intronic regions of genes known or suspected to be expressed in the adult brain, and each transgene was integrated into the same *attP2* site on the second chromosome. We chose embryos at two stages for detailed analysis: germband elongated (GBE; embryonic stages 9–12) and germband shortened (stages 15–17; most embryos were at stage 16). We crossed each line to a line containing *UAS-GFP:NLS* (green fluorescent protein [GFP] with a nuclear localization signal) and/or *cd8:GFP* (membrane-tethered GFP), and stained for GFP to determine the expression pattern of each line. We costained each line for the Even-skipped (Eve) transcription factor, which detects a subset of neurons and muscle precursors and anal pad (Frasch et al., 1987); these cells can be used as well-characterized landmarks to help annotate the GAL4 patterns. Our goal was to comprehensively annotate the gene expression patterns in NBs, GMCs, neurons, and glia, including the specialized neurons and glia at the ventral midline. We noted expression outside the CNS but did not attempt to annotate nonneural tissue patterns.

Our data are included in a publicly available database hosted by JFRC (<http://www.janelia.org/gal4-gen1>; see Figures 1A–1C for screenshots describing how to search the database for embryonic patterns). Additional details regarding the website construction and database functions are described in an accompanying paper (Jenett et al., 2012). The database integrates expression data from embryos (this work) as well as from the adult brain (Jenett et al., 2012) and larval imaginal discs (Jory et al., 2012). Our embryonic data sets include a maximum intensity projection of the most relevant portion of the CNS staining pattern, as well as at least one movie that steps through a z-stack of confocal images to illustrate all aspects of the pattern. One can search the database using selected keywords (see Table 1 for GBE categories, Table 2 for stage 16 categories, and Table 3 for midline categories).

We found that nearly all of the 5,000 lines imaged had expression in both young GBE embryos (4,289; 86%) and older stage 16 embryos (4,672; 93%; Figure 1D). This is an underestimate of the GBE expression because we did not image this stage in 410 lines (see Experimental Procedures); thus, most of the lines contained active *cis*-regulatory modules (CRMs or enhancers). The most common pattern was head expression (GBE, 76%; stage 16, 74%; Figure 1D), possibly because the genes expressed in the adult brain were preferentially used to generate the GAL4 lines (Jenett et al., 2012). It should be noted, however, that head expression was observed in 59% of the 500 lacZ enhancer trap lines (Bellen et al., 1989), so the head may

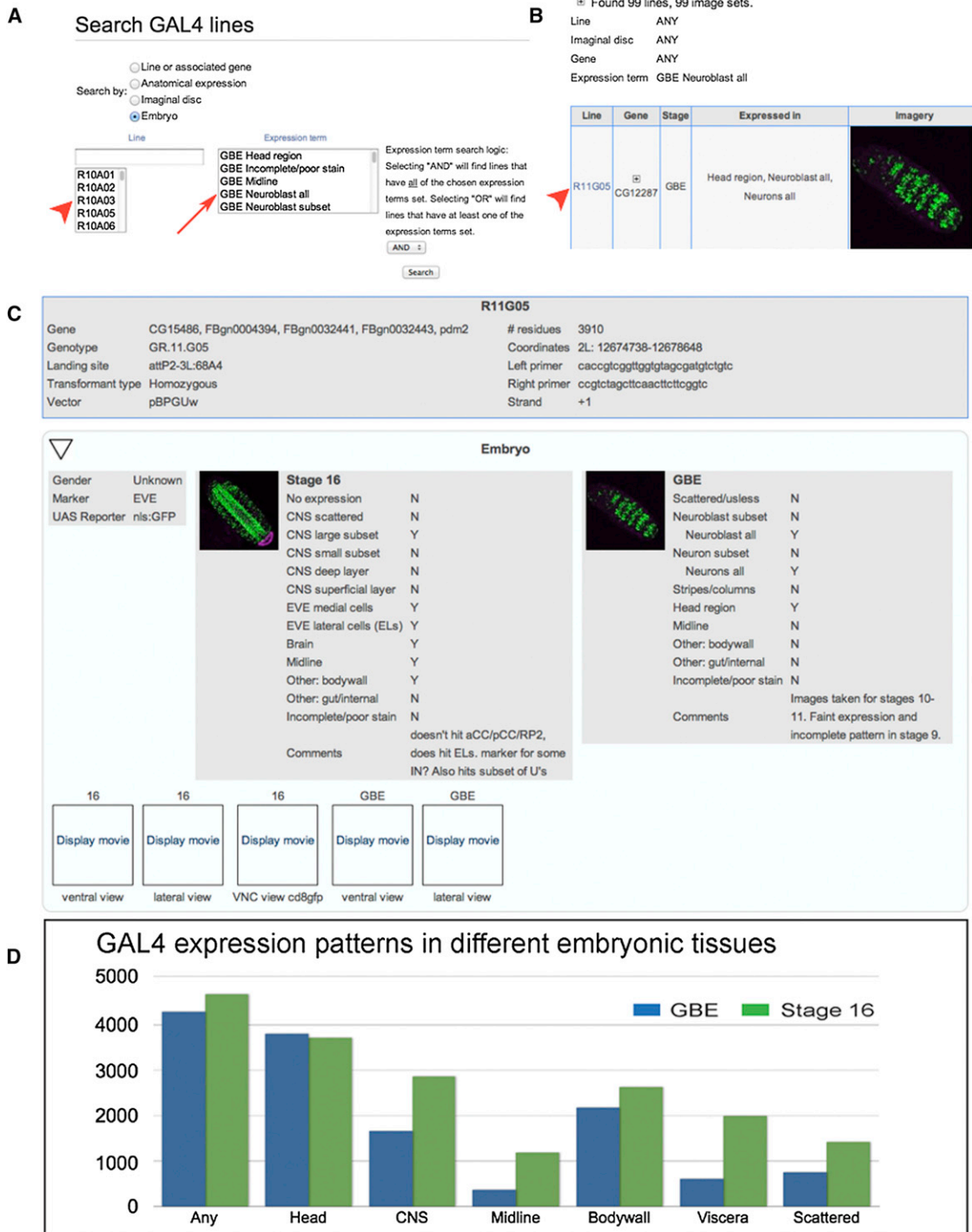


Figure 1. The GAL4 Database Interface

(A) Searching the GAL4 database. Searches can be done by GAL4 line name (red arrowhead) or by tissue expression (red arrow).

(B) Search results. Searches using tissue expression criteria can return multiple hits (99 in this example). Each hit can be viewed by selecting the GAL4 line name (red arrowhead).

(C) Information for each GAL4 line. This includes the associated genes for the line, genotype, genomic insertion site, size of the *cis*-regulatory DNA, and primers used to PCR amplify the *cis*-regulatory DNA (top box). It also includes gene expression data from the adult brain and ventral nerve cord (see Jenett et al., 2012, Jory et al., 2012) and the CNS in GBE (stages 9–12) and stage 16 embryos. For the embryonic CNS, the annotated expression patterns (Y, expression observed; N, no expression observed), maximum intensity projections, and QuickTime movies (click "Display movie" link to play movie) are given. Higher-resolution image data sets are available upon request.

Table 1. Lines Expressed in GBE Embryos: Stages 9–12

Expression Pattern	Number of Lines (%)	Number of Genes (%)
Total	5,000 (100%)	796 (100%)
Head	3,806 (76%)	736 (92%)
NBs, subset	662 (13%)	313 (39%)
NBs, most or all	99 (2%)	69 (9%)
GMCs/neurons, subset	1,411 (28%)	480 (60%)
GMCs/neurons, most or all	46 (1%)	36 (5%)
Stripe or column pattern	224 (4%)	140 (18%)
Other: body wall	2,176 (44%)	577 (72%)
Other: gut/internal	613 (12%)	336 (42%)
Scattered	756 (15%)	377 (47%)
Poor stain	385 (8%)	226 (28%)

utilize a large number of CRMs. The next most common patterns were CNS expression (GBE, 33%; stage 16, 57%) and body wall expression (GBE, 44%; stage 16, 53%; [Figure 1D](#)); this may be due to the complexity of cell types in these tissues. We saw relatively fewer lines expressed in the midline and viscera, and in a scattered pattern, but for each of these we observed nearly double the number of expressing lines observed in older embryos ([Figure 1D](#)), perhaps due to the increasing complexity of these tissues as development proceeds. We observed one expression pattern common to most of the 5,000 lines assayed: a pattern of four neuronal clusters in the stage 16 subesophageal CNS ([Figure S1](#)). This gene expression pattern is presumably derived from a DNA sequence shared by all GAL4 constructs (see [Discussion](#)).

GBE Patterns—Embryonic Stages 9–12

At this stage of neurogenesis, the neuroectoderm covers the ventral surface of the embryo, and just internal to these cells is the NB array. Progressively more internal are GMCs and newly born neurons (reviewed in [Skeath and Thor, 2003](#)). We used the Eve protein as a fiduciary marker ([Figure 2A](#)). Eve is detected in the first-born GMC from NB1-1 and NB4-2, the first five GMCs from NB7-1, and a segment-specific number of GMCs from NB3-3 (five in T1-T2, six in T3, and ten in A1-A7). Each of these GMCs divides to generate an Eve⁺ neuron and an Eve⁻ sibling ([Skeath and Doe, 1998](#)). In addition, Eve is detected in pericardial cells at the dorsal surface of the embryo and the anal pad ([Frasch et al., 1987; Figure 2A](#)).

We found 4,320 lines expressed at GBE ([Table 1; Figure 1D](#)). We were interested in pan-NB lines, with the rationale that the associated genes may show NB-specific gene expression. We identified NBs based on their relatively large size and position just internal to the ventral neuroectoderm. Despite the lack of a definitive NB marker such as Deadpan or Miranda ([Doe, 2008](#)), we feel confident that size and position are sufficient criteria to identify NBs. Approximately 2% (99) of the GBE+ lines

Table 2. Lines Expressed in Stage 16 Embryos

Expression Pattern	Number of Lines (%)	Number of Genes (%)
Total	5,000 (100%)	796 (100%)
Head	3,712 (74%)	731 (92%)
Neurons, small subset	1,862 (37%)	553 (69%)
Neurons, most or all	822 (16%)	381 (48%)
Eve medial cells	661 (13%)	309 (39%)
Eve lateral cells	785 (16%)	366 (46%)
Other: body wall	2,631 (53%)	635 (80%)
Other: gut/internal	1,997 (40%)	619 (78%)
Scattered	1,423 (28%)	520 (65%)
Poor stain	764 (15%)	341 (43%)
No expression	28 (1%)	27 (3%)

were detected in most or all NBs ([Figure 2B](#)). We were also interested in identifying GAL4 lines that label subsets of NBs, particularly single NBs that could be used for lineage analysis. More than 15% (662) of the GBE+ lines were detected in a subset of NBs, with some being in just a few NBs per hemisegment ([Figures 2C and 2D](#)). Additional experiments will be necessary to map these GAL4+ NBs onto the NB map ([Broadus et al., 1995](#)), and this may be best achieved using new semiautomated methods for creating gene expression atlases ([Long et al., 2009; Peng et al., 2010; Qu et al., 2011](#)). The lines expressed in subsets of NBs should be useful as NB markers to study the specification of unique NB identity. They could also be used to drive yeast Flippase to excise an *FRT-stop-FRT* cassette ([del Valle Rodríguez et al., 2012](#)) to allow permanent tracing of the NB and its progeny beyond embryonic stages, which would help investigators align embryonic NB identity with larval NB identity ([Broadus et al., 1995; Truman and Bate, 1988](#)) and determine the contribution of specific NBs to the adult CNS.

We detected 46 lines expressed in most or all GMCs/neurons ([Figure 2E](#)) and 1,411 lines expressed in a subset of GMCs/neurons ([Figures 2F–2H](#)). We did not try to distinguish GMCs from neurons, because these populations are intermingled at this stage and have the same cell size. Some of the expression patterns were regionally restricted, e.g., in stripes at different anterior/posterior positions in a segment, in medial or lateral columns along the length of the CNS, or specific to a segment (thoracic only or abdominal only). These are annotated in the database and can be retrieved by checking the stripe/column box (quantified in [Table 1](#)). The GAL4 lines with regional expression (stripe or column) may be used to identify flanking genes that are involved in spatial patterning of the CNS, or as markers to study the effect of earlier patterning genes on the specification of regional neuronal identity. They may also be useful for designing split GAL4 transgenes ([Luan et al., 2006; Pfeiffer et al., 2010](#)) to restrict GAL4 expression to a very small subset of cells (e.g., using a columnar enhancer to drive the GAL4

(D) Histogram showing the number of GAL4 lines with the indicated expression pattern. Any, any expression in the entire embryo; head, the entire head region, including brain and other cell types; CNS, ventral nerve cord from T1-A8; midline, CNS midline cells; body wall, all cell types within the lateral body wall, including sensory neurons, muscle, trachea, histoblasts; viscera, internal organs excluding the CNS; scattered, unpatterned cells in the CNS or elsewhere, including hemocytes.

Table 3. Lines Expressed in the CNS Midline

Expression Pattern	Number of Lines (%)	Number of Genes (%)
Total	5,000 (100%)	796 (100%)
Midline expression	1,285 (26%)	471 (59%)
HS total	438 (9%)	253 (32%)
LS total	847 (17%)	387 (49%)
GBE midline expression	389 (8%)	244 (31%)
GBE HS	137 (3%)	104 (13%)
GBE HS all	24 (1%)	16 (2%)
GBE HS subset	113 (2%)	91 (11%)
GBE LS	252 (5%)	183 (23%)
St16 midline expression	1,261 (25%)	468 (59%)
St16 HS	435 (9%)	252 (32%)
St16 HS all	26 (1%)	18 (2%)
St16 HS midline glia	109 (2%)	83 (10%)
St16 HS midline glia (all)	83 (2%)	69 (9%)
St16 HS VUMs	89 (2%)	66 (8%)
St16 HS MNB progeny	26 (1%)	23 (3%)
St16 HS neuron subset	331 (7%)	205 (26%)
St16 HS neuron subset, VUM-MNB	223 (4%)	156 (20%)
St16 LS	826 (17%)	383 (48%)
Dorsal median cells	24 (1%)	20 (3%)
Channel glia	15 (<1%)	14 (2%)
MM-CBG	14 (<1%)	13 (2%)

DNA-binding domain and the stripe enhancer to drive the GAL4 VP16 transactivation domain). These experiments are straightforward because the enhancer and *attP* integration sites are defined, and thus new split GAL4 transgenes should have highly predictable expression patterns.

A large fraction of lines with GAL4 patterns showed expression in the head (3,806/4,320; 88%). We did not try to map the identity of these cells to NBs, neurons, glia, or nonneuronal cell types of the head, due to the complexity of the pattern (Figures 2I–2L). A bias toward head expression may reflect the fact that the most *cis*-regulatory DNA was selected based on its proximity to genes expressed in the adult brain.

We found many lines expressed in nonneural tissues (Figures 2M–2P), but did not comprehensively annotate these patterns. We note that 2,176 lines were expressed in the body wall, which could include trachea, muscles, sensory neurons, histoblasts, or other cell types (e.g., Figures 2M–2O). We also observed expression of 613 lines in the viscera, termed “internal/gut” in the database (Figure 2P). A total of 756 lines showed scattered expression (data not shown; retrieve from the database by checking the “scattered” box). This could reflect stochastic expression of the line within a reproducible pattern of cells, expression in hemocytes that have variable positions due to their migratory nature, or unknown cell types that have a variable pattern.

Stage 16 CNS Patterns

At stage 16, the ventral nerve cord is composed of ~335 bilateral cells, including ~35 motoneurons and ~35 glia (Beckervorder-

sandforth et al., 2008; Landgraf et al., 1999; E.S.H. and C.Q.D., unpublished data). Most of the remaining 260 cells are presumed to be interneurons, but in fact are mostly uncharacterized. A major goal of this study was to identify GAL4 lines that are expressed in subsets of these presumed interneurons to allow further morphological and functional characterization. We detected 4,672 lines expressed at stage 16 (Figure 1D). At stage 16, the CNS shows a segmentally reiterated group of medial *Eve*⁺ neurons consisting of the RP2 motoneuron, the aCC/pCC motoneuron/interneuron siblings, and the U1–U5 motoneurons (Skeath and Doe, 1998). In addition, there is a cluster of *Eve*-lateral (EL) interneurons (Figure 3A). We identified 1,862 lines (37%) that are expressed in small subsets of neurons per hemisegment (<30 neurons per hemisegment, or ~10% of total cells; see Figures 3B–3E for examples). The more-specific of these lines will be especially useful for characterizing neuronal morphology and function. In contrast, we identified just 822 lines (16%) expressed in most or all neurons in the CNS; some are specific to the CNS and others have additional expression outside the CNS (examples shown in Figures 3F and 3G). Interestingly, previous enhancer-trap experiments showed a similar or lower percentage of lines with all-CNS expression (7/49, 14%; O’Kane and Gehring, 1987) and 42/3768 (1%; Bier et al., 1989), even though each insertion presumably queries a broader *cis*-regulatory region. It is somewhat surprising how few lines are expressed in a pan-neuronal pattern. This suggests that there may be relatively few CRMs devoted to pan-neuronal expression, and that genes that are widely expressed may utilize multiple, dispersed CRMs to achieve pan-neuronal expression.

In the future, the GAL4 lines expressed in <10 neurons per hemisegment can be used for many experiments. We are currently characterizing about 100 of these lines in more detail, with the goal of mapping each of the GAL4⁺ neurons into a CNS atlas of gene expression that would allow us to register each GAL4⁺ neuron to known neurons or each other (E.S. Heckscher et al., unpublished data). The lines can be used to express axon or dendrite markers to determine their morphology, possibly using GFP reconstitution across synaptic partners (Feinberg et al., 2008) to identify candidate synaptic partners. The lines can be used to express Ca²⁺ sensors to identify neurons with rhythmic activity that matches the periodicity of larval body wall contractions during locomotion. They can be used to express neuronal silencers to screen for behavioral defects, and they can be used to express channelrhodopsin to allow the neurons to be light activated. The pan-neuronal lines, particularly those without additional expression, will also be quite useful. Surprisingly few GAL4 lines are known to be pan-neuronal without additional expression. The most commonly used pan-neuronal GAL4 line is *Elav-GAL4*, but this line is expressed in NBs that give rise to glia (Berger et al., 2007), and thus is not appropriate for driving neuron-specific expression. We have also observed lines expressed in subsets or all embryonic glia, such as enhancers near the glial cells missing gene (data not shown); some examples of glial patterns are shown in Figures 3H–3J.

Previously, other investigators and members of our laboratory characterized the genetic regulatory network that specifies the *Eve*⁺ RP2 motoneuron (McDonald and Doe, 1997; McDonald

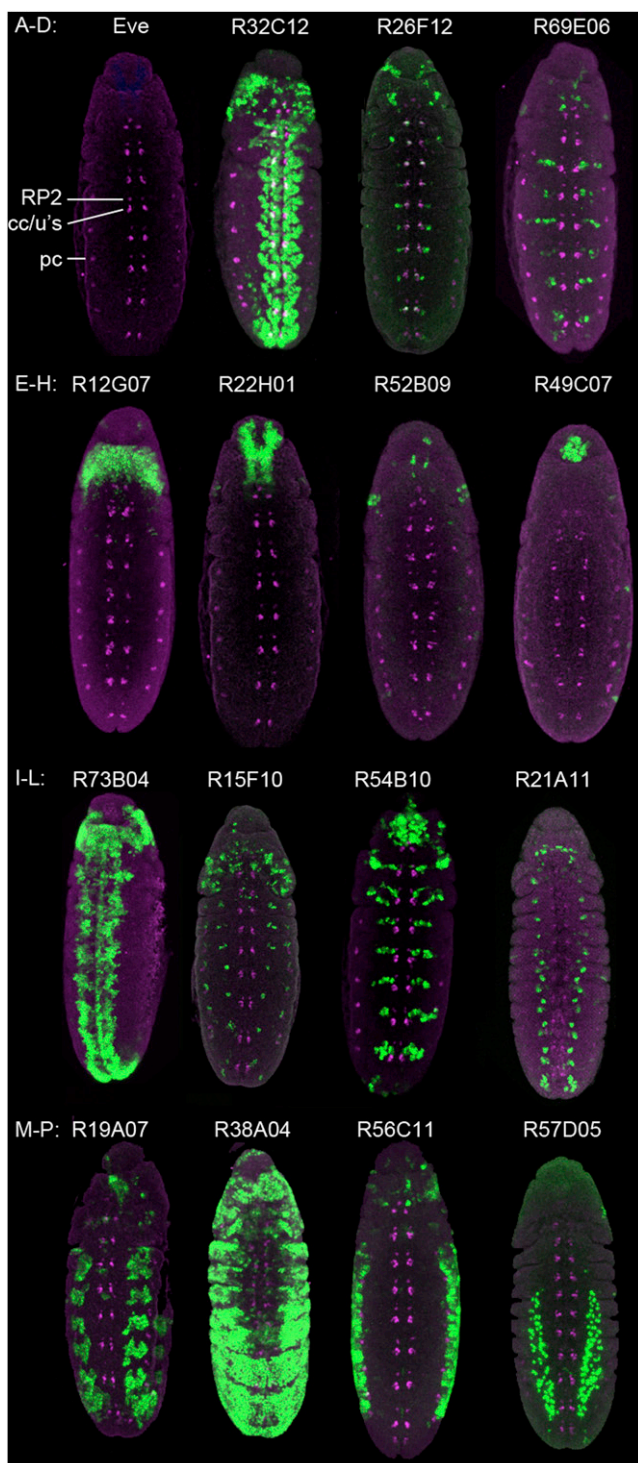


Figure 2. Representative GAL4 Expression Patterns in GBE Embryos

(A) Eve pattern at GBE stages. RP2: RP2 motoneuron; cc/u: aCC/pCC/U1/U2 (either GMCs or neurons); pc: pericardial precursors.

(B–D) Lines expressed in most NBs (B) or subsets of NBs (C–D). (B) Line R32C12, gene *stg*. (C) Line R26F12, gene *grn*. (D) Line R69E06, gene *gsb-n*.

(E–H) Lines expressed in most GMCs/neurons (E) or subsets of GMCs/neurons (F–H). (E) Line R73B04, gene *Sox21b*. (F) Line R15F10, gene *ss*. (G) Line R45B10, gene *pnt*. (H) Line R21A11, gene *eya*.

et al., 2003), and similar work has been done on the Eve+ U1–U5 neurons (Grosskortenhaus et al., 2006; Isshiki et al., 2001; Kohwi et al., 2011; Tran and Doe, 2008; Tran et al., 2010) and EL neurons (Tsuji et al., 2008). Identification of GAL4 lines expressed by Eve+ neurons would be useful for providing tools to manipulate gene expression in these neurons, and the flanking genes may be important for Eve+ neuron development or function. Thus, we searched for lines coexpressed by medial or lateral Eve+ neurons. We found 661 lines expressed in the medial Eve+ neurons, and 785 lines expressed in the lateral Eve+ neurons (ELs; Table 2; some examples of each are shown in Figures 3K–3O). In some cases many other cells express the line, which is not very useful, but in other cases the lines are fairly specifically expressed in the Eve+ neurons (Figures 3M and 3O). These lines should help define the genetic regulatory network that can be used to generate a specific neuron from an identified precursor, because the Eve+ neurons are among the best-characterized neurons in the CNS.

We also detected 3,712 lines expressed in the head (Figures 1D, 3P, and 3Q), although this number may be artificially high due to the “four clusters of neurons” general background pattern in nearly all lines (Figure S1). Many lines were also expressed in nonneural tissues such as the gut (e.g., Figures 1D and 3U–3W; quantified in Table 2) and body wall (which includes somatic body wall muscles, trachea, sensory neurons, and undefined cell types; Figures 1D, 3X, and 3Y), and in scattered cells that may include hemocytes (e.g., Figures 3R–3T).

CNS Midline Patterns

The ventral nerve cord contains a specialized set of neurons and glia that lie along the midline. By virtue of their appearance as a stripe down the midline of the embryo (Figure 4), it is relatively easy to identify expression of genes or transgenic lines that are midline expressed. Consequently, the midline cells are an attractive system for studying gene regulation, and this provides one rationale for our annotation of the JFRC GAL4 line collection for embryonic midline expression patterns.

There are two developmentally distinct groups of midline cells (Kearney et al., 2004). The cells that are generally considered midline cells are derived from *single-minded*+ mesectodermal cells and consist of the midline glia, which ensheath the axon commissures, and a diverse set of neurons. The second group is referred to as midline accessory cells. They reside at the midline but are either mesodermal cells that reside internal to the midline (dorsal-median cells) or are glia that arise from lateral neuroglia and migrate to the midline (channel glia and medial-most cell body glia [MM-CBG]). The ~22 mesectodermal-derived mature midline cells/segment at stage 16 are diverse and consist of (1) approximately three midline glia; (2)

(I–L) Lines expressed in the head. (I) Line R12G07, gene *kn*. (J) Line R22H01, gene *fru*. (K) Line R52B09, gene *gol*. (L) Line R49C07, gene *beat-IIIc*.

(M–P) Lines expressed in nonneural tissues. (M) Body wall muscle; line R19A07, gene *wl*. (N) Epithelial; line R38A04, gene *rst*. (O) Abdominal body wall; line R56C11, gene *lin-28*. (P) Internal cell types; line R57D05, gene *HLH54F*.

All panels show ventral views of the CNS of the indicated GAL4 UAS-nGFP lines stained for Eve (magenta) and nuclear GFP (green). Top: anterior.

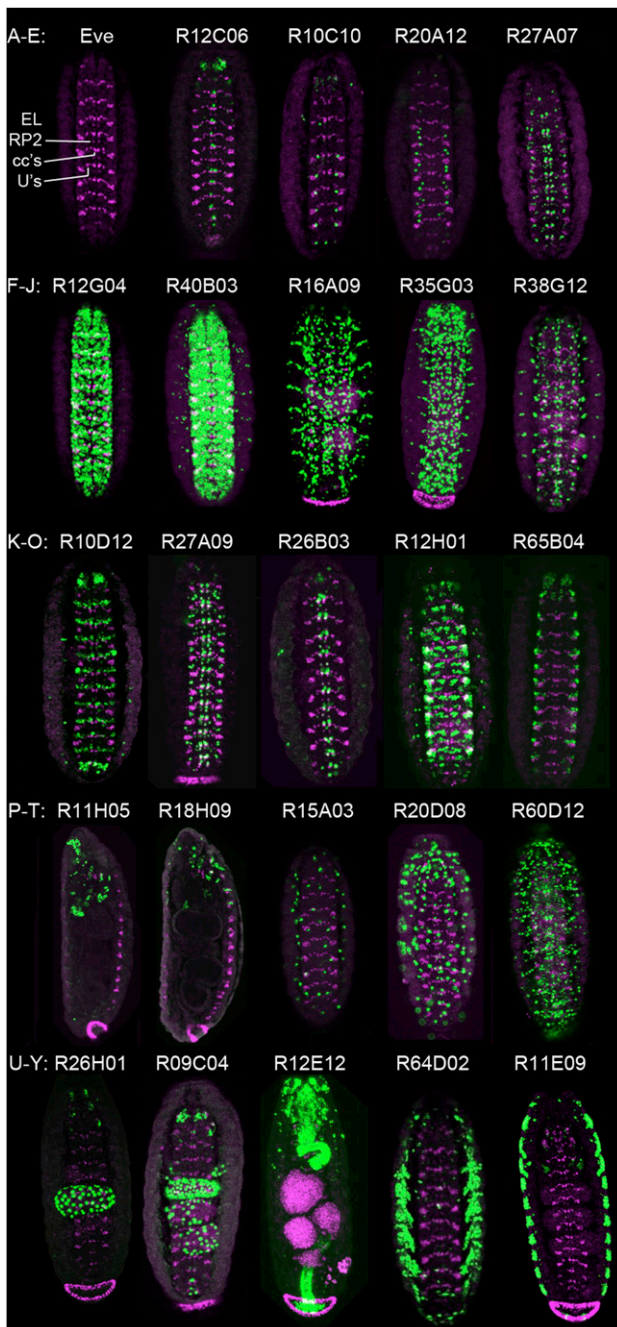


Figure 3. Representative GAL4 Expression Patterns in the Stage 16 CNS

(A) Eve pattern at stage 16. RP2: RP2 motoneuron; cc: aCC/pCC and U1/U2 neurons; U: U3–U5 neurons; EL: Eve lateral cluster.

(B–E) Lines expressed in a small subset of neurons. (A) Line R23C02, gene *Adar*. (B) Line R12C06, gene *Oli*. (C) Line R10C01, gene *bi*. (D) Line R20A12, gene *CG16805*. (E) Line R27A07, gene *Fur1*.

(F and G) Lines expressed in large subsets of neurons. (F) Line R12G04, gene *Oli*. (G) Line R40B03, gene *CG2672*.

(H–J) Lines expressed in glia. (H) Line R16A09, gene *ey*. (I) Line R35G03, gene *Lag1*. (J) Line R38G12, gene *Nrt*.

(K–O) Lines expressed in Eve+ neurons. (K–M) Medial Eve neurons (aCC, pCC, RP2, or U1–U5). (N–O) Lateral Eve neurons. (K) Line R10D12, gene *Fas2*. (L)

two MP1 peptidergic neurosecretory cells; (3) H-cell, a dopaminergic interneuron; (4) H-cell sib, a glutamatergic interneuron; (5) three glutamatergic and octopaminergic motoneuron ventral unpaired median interneurons (mVUMs); (6) three GABAergic interneuron VUMs (iVUMs); and (7) the median NB (MNB) and its approximately eight interneuronal progeny (Wheeler et al., 2006). A previous large-scale in situ hybridization analysis documented midline expression of 286 genes (Kearney et al., 2004); these data are publically available from the *Drosophila* CNS Midline Gene Expression Database (MidExDB; http://midline.bio.unc.edu/MDB_Home.aspx; Wheeler et al., 2009). A confocal microscopic analysis of ~100 genes yielded detailed maps of midline cell expression at each embryonic stage, so that each midline cell type can be identified at each stage of development (Wheeler et al., 2006, 2008).

Analysis of horizontal views of GBE and stage 16 GAL4 transgenic embryos using anti-Eve staining was particularly useful for identifying lines with midline GFP expression, since medial Eve+ neurons, including aCC, pCC, and RP2 cells, bracket midline-expressing cells (Figures 4A–4J). At the GBE stage, midline cells are just acquiring their fates, and cannot easily be distinguished except by staining with cell-type-specific markers. Consequently, we designate GBE midline-expressing enhancers as expressed in either all midline cells or a subset of midline cells. In contrast, at stage 16 it is often possible to make judgments as to the identity of the midline cell type based on the position along the apical-basal axis and morphology (Wheeler et al., 2006, 2008). In this manner, the midline glia, VUM neurons, and progeny of the MNB can often be determined. In contrast, without appropriate markers, it is not possible to identify the other midline neurons. In annotating midline-expressed lines, we chose to liberally assign midline GFP+ cells to a specific midline cell type.

In our analysis, we analyzed 5,000 lines for midline expression corresponding to 796 genes (Table 3; Table S2). We identified 1,285 lines that had midline expression at either the GBE stage (389 lines) or at stage 16 (1,261 lines). Overall, 26% of all lines examined had midline expression, corresponding to 59% of the genes analyzed. We subdivided the 1,285 lines with midline expression into two classes: high-specificity (HS) lines and low-specificity (LS) lines. The LS lines (847 lines, 66% of total) had midline expression but usually also had expression in a large number of neurons in the lateral CNS. In contrast, the HS lines (438 lines, 34% of total) generally showed strong midline expression and were not generally expressed in large numbers of CNS cells. Nevertheless, many HS lines had expression in other

Line R27A09, gene *grn*. (M) Line R26B03, gene *grn*. (N) Line R12H01, gene *kn*. (O) Line R65B04, gene *Tk*.

(P–Q) Lines expressed in the brain. (P) Line R11H05, gene *norP*. (Q) Line R18H09, gene *wb*.

(R–T) Lines expressed in a scattered pattern. (R) Line R15A03, gene *Syt1*. (S) Line R20D08, gene *cib*. (T) Line R60D12, gene *arm*.

(U–Y) Lines expressed in the gut (U–W) or lateral body wall (X–Y). (U) Line R26H01, gene *lab*. (V) Line R9C04, gene *dac*. (W) Line R12E12, gene *h*. (X) Line R64D02, gene *14-3-3zeta*. (Y) Line R11E09, gene *bi*.

All panels show ventral views of the stage 16 CNS, with the exception of (S)–(U), which are lateral views. Magenta: Eve; green: nuclear GFP; top: anterior. See also Figure S3.

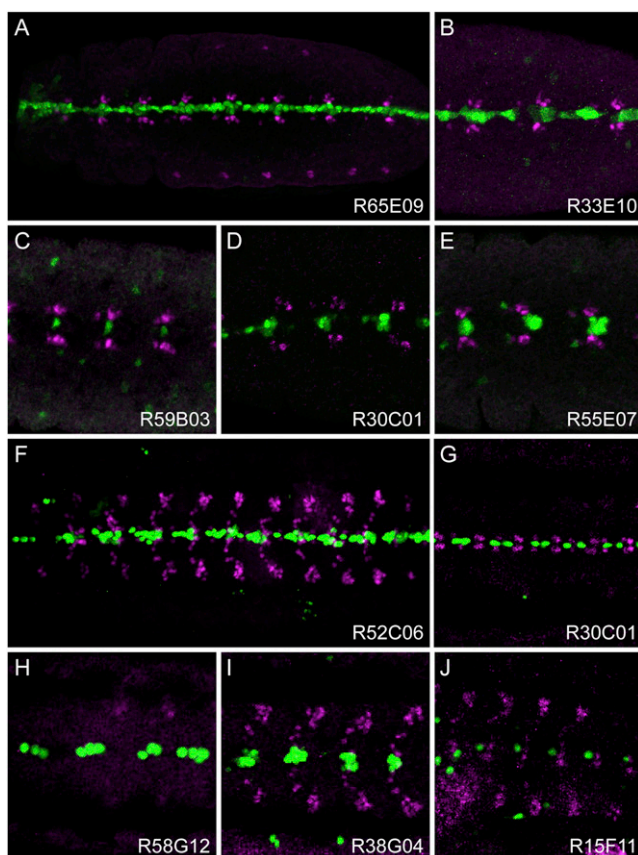


Figure 4. Representative GAL4 Expression Patterns in CNS Midline Cells

(A–E) Images of GAL4 lines at GBE stages showing midline primordium staining of (A) all midline cells (line R65E09, gene *Tk*), (B) a large subset (R33E10, gene *ct*), (C) a single cell (R59B03, gene *CG8867*), (D) a small subset of likely midline glia (line R30C01, gene *netB*), and (E) a small subset of neural precursors or neurons (line R55E07, gene *CG13253*).

(F–J) Images of stage 16 embryos showing (F) all midline cells (line R52C06, gene *cdl*), (G) midline glia (line R30C01, gene *netB*), (H) iVUMs (line R58G12, gene *Gad1*), (I) MNB progeny (line R38G04, gene *sca*), and (J) a single neuron (line R15F11, gene *Syt1*).

All panels show ventral views of the CNS of the indicated GAL4 UAS-*nGFP* lines stained for *Eve* (magenta) and nuclear GFP (green). Arrowhead: ventral midline; anterior is to the left.

embryonic and CNS cell types. The HS lines corresponded to 253 genes, and 111 of these genes had multiple HS lines with an average of 2.7 lines per gene. Generally, when a gene had multiple lines with the same midline pattern, they were overlapping (and thus likely contained the same CRM). In contrast, when a gene had multiple lines with distinct midline expression patterns, they were in nonoverlapping DNA segments and thus must represent different CRMs. Only rarely were lines with identical patterns of expression in the same gene found in nonoverlapping DNA segments (sometimes referred to as shadow enhancers; Barolo, 2012).

Many of the HS genes are known to be expressed in midline cells. For example, MidExDB currently contains expression data for 286 genes expressed in CNS midline cells. Of the 253

HS genes, 56 are listed in MidExDB. Another 20 HS genes are expressed in midline cells, based on published accounts. Consequently, although a significant fraction of the HS gene enhancers reside within known midline-expressed genes, new midline-expressed genes may be identified based on our discovery of new midline enhancers. However, there are examples of midline enhancers that reside within well-studied genes that do not have corresponding strong midline expression. For example, the *similar* (*Drosophila Hypoxia Inducible Factor* ortholog) bHLH-PAS gene, which is ubiquitously expressed (Nambu et al., 1996), has two intronic overlapping midline-expressed lines (R14D11 and R14E09). Consequently, the role of these enhancers in vivo remains unknown. Many of the midline enhancers reside in 5'-flanking regions, 3'-flanking regions, and introns, often within relatively close distance to a midline-expressed gene. However, there are examples of midline enhancers that may act as relatively long distances. For example, the overlapping lines R25F10 and R27A10 reside ~15 kb upstream of the *charlatan* gene, which is broadly expressed in the CNS but does not have prominent midline expression (Escudero et al., 2005; Yamasaki et al., 2011). However, the midline-expressed *hibris* gene is located 74 kb upstream of *charlatan* (Artero et al., 2001; Dworak et al., 2001), and the enhancer, which is ~58 kb downstream of *hibris*, may control *hibris* midline expression.

The HS midline-expressed lines include a variety of midline patterns that are listed in Table 3, Table S2, and Figure 4. At the GBE stage, 24 lines were expressed in all midline primordium cells (Figure 4A) and 113 were expressed in subsets (Figures 4B–4E). The subsets included large subsets (Figure 4B) and subsets with as few as a single cell per segment (Figure 4C). In some cases, the subsets of midline primordium cells can be tentatively assigned to midline glia or neural precursors based on the identity of the stage 16 midline cell types. For example, R30C01 (*netrin-B*) is expressed only in midline glia at stage 16 (Figure 4G) and is found in a subset of midline cells (likely midline glia) in GBE embryos (Figure 4D).

At stage 16, the midline cells can in many cases be distinguished as to midline cell type. In 26 HS lines, expression was present in all midline cells (Figure 4F); these were generally the same lines with GBE all-midline cell expression. Stage 16 expression in all midline cells may reflect perdurance from expression at the earlier midline primordium stages, because few genes, if any, are prominently expressed in all midline cells at stage 16, or this expression could be due to an incomplete CRM lacking a late repression element. Many HS lines were expressed in subsets of midline cells (Figures 4G–4J); in some cases expression was present in a single midline cell type, and in other cases expression was present in multiple midline cell types. We categorized HS lines expressed in specific cell types as midline glia, VUMs, MNB progeny, and small subsets of neurons. The midline glia can generally be identified based on their characteristic dorsal position, number, and elongated nuclear morphology. We annotated 83 HS lines with midline glia expression, excluding lines expressed in all midline cells. In 27 of these 83 lines, the only midline cell type is midline glia (Figure 4G). The VUM neurons can also be identified based on their characteristic ventral position and relatively large nuclei

(Figure 4H). In the small number of cases in which we assign a VUM subset to either iVUMs or mVUMs, this is based on the existing knowledge that the corresponding gene is expressed in either iVUMs or mVUMs. For example, *Gad1* is expressed in the GABAergic iVUMs (Wheeler et al., 2006), making it likely that line R58G12 whose fragment resides in an intron of *Gad1* is expressed in iVUMs (Figure 4H). Eighty-nine HS lines are expressed in the VUM neurons, many in VUM subsets, likely either iVUMs or mVUMs. The MNB progeny tend to form a cluster of small neuronal nuclei (Figure 4I), and were the predominant midline cell type in 26 HS lines. The neural subset category consists of 223 HS lines that express GFP in one or a small number of neuronal nuclei, but could not be assigned to a specific midline neuron without staining with cell-type-specific markers. In those rare cases in which a GFP⁺ neuron is assigned to a specific type (e.g., H-cell; R29A07), this is because the corresponding gene is known to be expressed in that cell type (*pale*; Wheeler et al., 2006). The last class of annotated midline lines (which includes HS and LS) corresponds to the midline accessory cells: the dorsal median cells (24 lines), channel glia (15 lines), and MM-CBG (14 lines). The large number of GAL4 lines that correspond to each midline cell type have the potential to be useful for functional studies of midline cell types, as well analysis of enhancer function and the regulation of midline gene expression.

Spatial and Temporal Changes in Gene Expression during Development

We identified >4,000 GAL4 lines with expression in early embryos (GBE) or in late embryos (stage 16), which allowed us to determine whether there are different CRMs for each tissue at each stage of development, or whether individual CRMs give persistent tissue-specific expression over time. We found that the vast majority of GAL4 lines have some embryonic expression in both early and late embryos (85%, with only 3% early only and 12% late only; Figure 5A). Within the CNS there is more stage-specific expression (6% early specific, 45% late specific, and 49% at both stages; Figure 5A). The increased number of lines specifically expressed in the stage 16 CNS is likely due to the greater complexity of cell types present in the CNS at this stage. Similar results were observed for the midline, body wall, and viscera patterns (Figure 5A), reflecting the increased cellular complexity in these tissues. Most tissues had a relatively small number of early-embryo-specific patterns (6% CNS, 2% midline, and 5% viscera; Figure 5A). This suggests that once a CRM is active in young embryos, it often may remain active as the embryo develops. Although we cannot rule out the possibility that the perdurance of GAL4 or the GFP reporter proteins contributes to the observed persistent expression from young to old embryos, we note that some tissues do in fact show a much higher percentage of early-embryo-specific expression (19% body wall and 22% scattered; Figure 5A), which shows that GAL4 and GFP can be turned over during early embryogenesis.

We were also interested in whether the GAL4 lines showed a bias toward any cell type or subset of the CNS. We chose to focus on the Eve⁺ neurons and the midline neurons, which were the most comprehensively annotated cell types. The

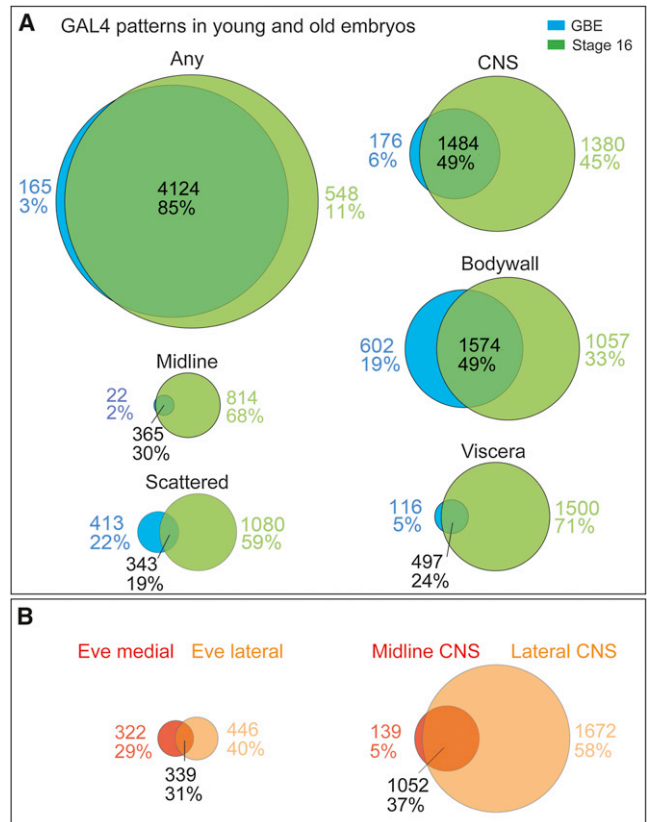


Figure 5. Spatial and Temporal Changes in Gene Expression during Development

(A) Venn diagrams showing the number of GAL4 lines expressed in different tissues (defined in Figure 1D legend) within early embryos (GBE, stages 9–12) and late embryos (stage 16). Only lines with data for both GBE and stage 16 embryos were included.

(B) Venn diagrams showing the number of GAL4 lines expressed in Eve medial and EL neurons (left) or midline and lateral CNS (right) within early embryos (GBE, stages 9–12) and late embryos (stage 16).

Eve⁺ neurons can be divided into a medial group that is primarily motoneurons (seven motoneurons and one interneuron) and a lateral group that is all interneurons. Interestingly, we found no significant bias in gene expression between these two groups: the lines were nearly equally divided among medial-specific, lateral-specific, and both (Figure 5B). This shows that the CRMs in this study appear equally likely to be expressed in the medial and lateral regions of the CNS, and within motoneurons and interneurons—with the caveat that Eve represents only a subset of motoneuron and interneurons in the CNS. Analysis of midline and lateral CNS expression gave a different result: 58% of the lines were expressed in the lateral CNS only, 37% were expressed in both the midline and lateral CNS, and only 5% were expressed in the midline only (Figure 5B). This suggests that the lateral CNS has more gene regulatory complexity than the midline, a result that is consistent with the greater number of unique cell types in the lateral CNS (Bossing et al., 1996; Schmid et al., 1999; Schmidt et al., 1997; Wheeler et al., 2009).

Conclusions and Future Directions

We have characterized the embryonic expression of 5,000 GAL4 lines, and provided the images and *cis*-regulatory DNA sequence information in a public website. All GAL4 lines were made with molecularly defined *cis*-regulatory DNA inserted into a single *attP* site in the genome, and thus the observed patterns can be replicated with the use of other gene products. For example, we have already used the R9D11 *cis*-regulatory DNA to make a direct fusion to red fluorescent protein, to be used as a marker in combination with other GAL4/UAS transgenes (O. Bayraktar and C.Q.D., unpublished data). Other applications could use *cis*-regulatory DNA to drive LexA or FLPase, which could then be used in combination with another GAL4/UAS transgene or to express split GAL4 components. The collection of GAL4 lines will allow investigators to target many different NB lineages or neuronal subsets using misexpression or UAS-RNAi screens. As mentioned above, perhaps the most powerful use of these lines will be in neural circuit analysis: lines with GAL4 expression in fewer than ten neurons per hemisegment may be further refined by using smaller fragments of *cis*-regulatory DNA or a split GAL4 methodology, allowing single neurons to be targeted. Combining GAL4 lines with highly restricted expression patterns with permanent-labeling transgenes will be a powerful method for linking embryonic neurons to their adult morphology and function. Lastly, it should be possible to perform bioinformatic analysis on the *cis*-regulatory DNA from lines with overlapping expression patterns to identify motifs that may drive cell-specific gene expression, as has been done for larval imaginal disc CRMs (Jory et al., 2012).

EXPERIMENTAL PROCEDURES

Generation of Fragment Enhancer-GAL4 Lines

Details regarding the design and manufacture of the enhancer-GAL4 constructs, transformant fly stocks, and all related protocols were previously described (Pfeiffer et al., 2008).

Drosophila Genetics

Males from each GAL4-expressing line were crossed to virgin females of *y,w; UAS-GFP::lacZ.nls* (Bloomington Indiana stock #6452), which produces nuclear localized GFP. In a few cases, the male GAL4 line was crossed to virgin females of *y,w; UAS-mCD8::GFP, w+* (Bloomington Indiana stock #5137) that produces membrane-bound GFP.

Antibody Staining and Image Analysis

We performed 24 hr embryo collections at room temperature (22.5°C) on apple agar caps with yeast paste. Standard methods were used to fix the embryos (Odden et al., 2002), which were stored in 100% ethanol at -20°C until antibody staining was performed. Embryos were incubated in primary antibody for 1.5 hr at room temperature (22.5°C), rinsed in phosphate-buffered saline (PBS) with 0.1% Triton X-100, 3% bovine serum albumin, and 10 mM glycine (PTBG) for 60 min. They were then incubated in secondary antibody for 1.5 hr at room temperature (22.5°C), rinsed twice for 10 min in PTBG, followed by a single 15 min rinse in 10% glycerol in PBS on a rocker. Embryos were dehydrated by passing through a 25%, 50%, 90% glycerol series and stored overnight in mounting media (90% glycerol with 2% N-propyl gallate). We used the following primary and secondary antibodies: chicken anti-GFP (1:1000; Aves, Tigard OR), mouse anti-Eve 2B10 concentrate with a working concentration of 4 µg/ml, preincubated on fixed wild-type embryos prior to use (Developmental Studies Hybridoma Bank developed under the auspices of the National Institute of Child Health and Human Development [NICHD] and maintained by the Department of Biology, University of Iowa [Iowa City, IA]), donkey anti-chicken

DyLight 488, and donkey anti-mouse DyLight 549 (1:400; Jackson Immunochemicals, West Grove PA). Reduction of the rinsing time dramatically increased the signal of Eve antibody. Embryos were imaged on a Bio-Rad Radiance 2100MP or a Zeiss 700 confocal microscope. We did not determine the gender of the imaged embryos. ImageJ was used to produce QuickTime movies and TIFF projections.

Data Collection

We collected z-stack images of the stage 16 ventral nerve cord in the T2-A3 segment region. For each line, our goal was to customize the region imaged to the extent of the staining. We also imaged GBE embryos (stages 9–12), typically imaging the entire embryo. We did not image GBE stages for 410 lines due to the lack of appropriately staged embryos or poor staining quality. Most of the image stacks were collected at 20× + 1.5 zoom at 1.0 µm z-steps, with the number of z-steps depending on the extent of the staining. Eve protein was a marker for the location and general dimensions of the CNS, and allowed accurate orientation and staging of the embryos.

SUPPLEMENTAL INFORMATION

Supplemental Information includes one figure and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2012.09.009>.

LICENSING INFORMATION

This is an open-access article distributed under the terms of the Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 Unported License (CC-BY-NC-ND; <http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode>).

ACKNOWLEDGMENTS

We thank R. Mann for comments on the manuscript, Tallon Lamoreaux for assistance with fly stocks, and Janet Hanawalt for administrative support. This work was supported by an NRSA postdoctoral award to J.C.P. (NICHD), NIH grants R01 NS64264 (NINDS) and R37 RD25251 (NICHD) to S.T.C., and NIH grant HD27056 and support from the Howard Hughes Medical Institute to C.Q.D.

Received: August 15, 2012

Revised: September 11, 2012

Accepted: September 17, 2012

Published online: October 11, 2012

WEB RESOURCES

The URLs for data presented herein are as follows:

Drosophila CNS Midline Gene Expression Database (MidExDB), http://midline.bio.unc.edu/MDB_Home.aspx

FlyLight Image Database, <http://www.janelia.org/gal4-gen1>

REFERENCES

- Akiyama-Oda, Y., Hosoya, T., and Hotta, Y. (1999). Asymmetric cell division of thoracic neuroblast 6-4 to bifurcate glial and neuronal lineage in *Drosophila*. *Development* 126, 1967–1974.
- Artero, R.D., Castanon, I., and Baylies, M.K. (2001). The immunoglobulin-like protein Hibris functions as a dose-dependent regulator of myoblast fusion and is differentially controlled by Ras and Notch signaling. *Development* 128, 4251–4264.
- Baier, H., and Scott, E.K. (2009). Genetic and optical targeting of neural circuits and behavior—zebrafish in the spotlight. *Curr. Opin. Neurobiol.* 19, 553–560.
- Barolo, S. (2012). Shadow enhancers: frequently asked questions about distributed *cis*-regulatory information and enhancer redundancy. *Bioessays* 34, 135–141.

- Baumgardt, M., Karlsson, D., Terriente, J., Díaz-Benjumea, F.J., and Thor, S. (2009). Neuronal subtype specification within a lineage by opposing temporal feed-forward loops. *Cell* 139, 969–982.
- Beckervordersandforth, R.M., Rickert, C., Altenhein, B., and Technau, G.M. (2008). Subtypes of glial cells in the *Drosophila* embryonic ventral nerve cord as related to lineage and gene expression. *Mech. Dev.* 125, 542–557.
- Bellen, H.J., O’Kane, C.J., Wilson, C., Grossniklaus, U., Pearson, R.K., and Gehring, W.J. (1989). P-element-mediated enhancer detection: a versatile method to study development in *Drosophila*. *Genes Dev.* 3, 1288–1300.
- Berger, C., Renner, S., Lüer, K., and Technau, G.M. (2007). The commonly used marker ELAV is transiently expressed in neuroblasts and glial cells in the *Drosophila* embryonic CNS. *Dev. Dyn.* 236, 3562–3568.
- Bier, E., Vaessin, H., Shepherd, S., Lee, K., McCall, K., Barbel, S., Ackerman, L., Carretto, R., Uemura, T., Grell, E., et al. (1989). Searching for pattern and mutation in the *Drosophila* genome with a P-lacZ vector. *Genes Dev.* 3, 1273–1287.
- Bischof, J., Maeda, R.K., Hediger, M., Karch, F., and Basler, K. (2007). An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. *Proc. Natl. Acad. Sci. USA* 104, 3312–3317.
- Bossing, T., Udolph, G., Doe, C.Q., and Technau, G.M. (1996). The embryonic central nervous system lineages of *Drosophila melanogaster*. I. Neuroblast lineages derived from the ventral half of the neuroectoderm. *Dev. Biol.* 179, 41–64.
- Brand, A.H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118, 401–415.
- Branda, C.S., and Dymecki, S.M. (2004). Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice. *Dev. Cell* 6, 7–28.
- Broadus, J., Skeath, J.B., Spana, E.P., Bossing, T., Technau, G., and Doe, C.Q. (1995). New neuroblast markers and the origin of the aCC/pCC neurons in the *Drosophila* central nervous system. *Mech. Dev.* 53, 393–402.
- Broihier, H.T., and Skeath, J.B. (2002). *Drosophila* homeodomain protein dHB9 directs neuronal fate via crossrepressive and cell-nonautonomous mechanisms. *Neuron* 35, 39–50.
- del Valle Rodríguez, A., Didiano, D., and Desplan, C. (2012). Power tools for gene expression and clonal analysis in *Drosophila*. *Nat. Methods* 9, 47–55.
- Doe, C.Q. (2008). Neural stem cells: balancing self-renewal with differentiation. *Development* 135, 1575–1587.
- Dworak, H.A., Charles, M.A., Pellerano, L.B., and Sink, H. (2001). Characterization of *Drosophila hibris*, a gene related to human nephrin. *Development* 128, 4265–4276.
- Escudero, L.M., Caminero, E., Schulze, K.L., Bellen, H.J., and Modolell, J. (2005). Charlatan, a Zn-finger transcription factor, establishes a novel level of regulation of the proneural *achaete/scute* genes of *Drosophila*. *Development* 132, 1211–1222.
- Feinberg, E.H., Vanhoven, M.K., Bendesky, A., Wang, G., Fetter, R.D., Shen, K., and Bargmann, C.I. (2008). GFP reconstitution across synaptic partners (GRASP) defines cell contacts and synapses in living nervous systems. *Neuron* 57, 353–363.
- Frasch, M., Hoey, T., Rushlow, C., Doyle, H., and Levine, M. (1987). Characterization and localization of the even-skipped protein of *Drosophila*. *EMBO J.* 6, 749–759.
- Fujioka, M., Emi-Sarker, Y., Yusibova, G.L., Goto, T., and Jaynes, J.B. (1999). Analysis of an even-skipped rescue transgene reveals both composite and discrete neuronal and early blastoderm enhancers, and multi-stripe positioning by gap gene repressor gradients. *Development* 126, 2527–2538.
- Fujioka, M., Lear, B.C., Landgraf, M., Yusibova, G.L., Zhou, J., Riley, K.M., Patel, N.H., and Jaynes, J.B. (2003). Even-skipped, acting as a repressor, regulates axonal projections in *Drosophila*. *Development* 130, 5385–5400.
- Grosskortenhaus, R., Robinson, K.J., and Doe, C.Q. (2006). Pdm and Castor specify late-born motor neuron identity in the NB7-1 lineage. *Genes Dev.* 20, 2618–2627.
- Groth, A.C., Fish, M., Nusse, R., and Calos, M.P. (2004). Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics* 166, 1775–1782.
- Isshiki, T., Pearson, B., Holbrook, S., and Doe, C.Q. (2001). *Drosophila* neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. *Cell* 106, 511–521.
- Jenett, A., Rubin, G., Ngo, T., Shepard, D., Murphy, C., Dionne, H., Pfeiffer, B., Hibbard, K., Cavallaro, A., Hall, D., et al. (2012). A GAL4-driver line resource for *Drosophila* neurobiology. *Cell Rep.* Published online October 11, 2012. <http://dx.doi.org/10.1016/j.celrep.2012.09.011>.
- Jory, A., Estella, C., Giorgianni, M., Slattery, M., Lavery, T., Rubin, G., and Mann, R. (2012). A survey of 6300 genomic fragments for cis-regulatory activity in the imaginal discs of *Drosophila melanogaster*. *Cell Rep.* Published online October 11, 2012. <http://dx.doi.org/10.1016/j.celrep.2012.09.010>.
- Kearney, J.B., Wheeler, S.R., Estes, P., Parente, B., and Crews, S.T. (2004). Gene expression profiling of the developing *Drosophila* CNS midline cells. *Dev. Biol.* 275, 473–492.
- Kohwi, M., Hiebert, L.S., and Doe, C.Q. (2011). The pipsqueak-domain proteins Distal antenna and Distal antenna-related restrict Hunchback neuroblast expression and early-born neuronal identity. *Development* 138, 1727–1735.
- Landgraf, M., Roy, S., Prokop, A., VijayRaghavan, K., and Bate, M. (1999). Even-skipped determines the dorsal growth of motor axons in *Drosophila*. *Neuron* 22, 43–52.
- Layden, M.J., Odden, J.P., Schmid, A., Garces, A., Thor, S., and Doe, C.Q. (2006). Zfh1, a somatic motor neuron transcription factor, regulates axon exit from the CNS. *Dev. Biol.* 291, 253–263.
- Lee, T., and Luo, L. (2001). Mosaic analysis with a repressible cell marker (MARCM) for *Drosophila* neural development. *Trends Neurosci.* 24, 251–254.
- Long, F., Peng, H., Liu, X., Kim, S.K., and Myers, E. (2009). A 3D digital atlas of *C. elegans* and its application to single-cell analyses. *Nat. Methods* 6, 667–672.
- Luan, H., Lemon, W.C., Peabody, N.C., Pohl, J.B., Zelensky, P.K., Wang, D., Nitabach, M.N., Holmes, T.C., and White, B.H. (2006). Functional dissection of a neuronal network required for cuticle tanning and wing expansion in *Drosophila*. *J. Neurosci.* 26, 573–584.
- McDonald, J.A., and Doe, C.Q. (1997). Establishing neuroblast-specific gene expression in the *Drosophila* CNS: huckebein is activated by Wingless and Hedgehog and repressed by Engrailed and Gooseberry. *Development* 124, 1079–1087.
- McDonald, J.A., Fujioka, M., Odden, J.P., Jaynes, J.B., and Doe, C.Q. (2003). Specification of motoneuron fate in *Drosophila*: integration of positive and negative transcription factor inputs by a minimal eve enhancer. *J. Neurobiol.* 57, 193–203.
- Nambu, J.R., Chen, W., Hu, S., and Crews, S.T. (1996). The *Drosophila melanogaster* similar bHLH-PAS gene encodes a protein related to human hypoxia-inducible factor 1 alpha and *Drosophila* single-minded. *Gene* 172, 249–254.
- O’Kane, C.J., and Gehring, W.J. (1987). Detection in situ of genomic regulatory elements in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 84, 9123–9127.
- Odden, J.P., Holbrook, S., and Doe, C.Q. (2002). *Drosophila* HB9 is expressed in a subset of motoneurons and interneurons, where it regulates gene expression and axon pathfinding. *J. Neurosci.* 22, 9143–9149.
- Pearson, B.J., and Doe, C.Q. (2003). Regulation of neuroblast competence in *Drosophila*. *Nature* 425, 624–628.
- Peng, H., Ruan, Z., Long, F., Simpson, J.H., and Myers, E.W. (2010). V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets. *Nat. Biotechnol.* 28, 348–353.
- Pfeiffer, B.D., Jenett, A., Hammonds, A.S., Ngo, T.T., Misra, S., Murphy, C., Scully, A., Carlson, J.W., Wan, K.H., Lavery, T.R., et al. (2008). Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 105, 9715–9720.

- Pfeiffer, B.D., Ngo, T.T., Hibbard, K.L., Murphy, C., Jenett, A., Truman, J.W., and Rubin, G.M. (2010). Refinement of tools for targeted gene expression in *Drosophila*. *Genetics* *186*, 735–755.
- Qu, L., Long, F., Liu, X., Kim, S., Myers, E., and Peng, H. (2011). Simultaneous recognition and segmentation of cells: application in *C. elegans*. *Bioinformatics* *27*, 2895–2902.
- Schmid, A., Chiba, A., and Doe, C.Q. (1999). Clonal analysis of *Drosophila* embryonic neuroblasts: neural cell types, axon projections and muscle targets. *Development* *126*, 4653–4689.
- Schmidt, H., Rickert, C., Bossing, T., Vef, O., Urban, J., and Technau, G.M. (1997). The embryonic central nervous system lineages of *Drosophila melanogaster*. II. Neuroblast lineages derived from the dorsal part of the neuroectoderm. *Dev. Biol.* *189*, 186–204.
- Scott, E.K. (2009). The Gal4/UAS toolbox in zebrafish: new approaches for defining behavioral circuits. *J. Neurochem.* *110*, 441–456.
- Scott, E.K., Mason, L., Arrenberg, A.B., Ziv, L., Gosse, N.J., Xiao, T., Chi, N.C., Asakawa, K., Kawakami, K., and Baier, H. (2007). Targeting neural circuitry in zebrafish using GAL4 enhancer trapping. *Nat. Methods* *4*, 323–326.
- Skeath, J.B., and Doe, C.Q. (1998). Sanpodo and Notch act in opposition to Numb to distinguish sibling neuron fates in the *Drosophila* CNS. *Development* *125*, 1857–1865.
- Skeath, J.B., and Thor, S. (2003). Genetic control of *Drosophila* nerve cord development. *Curr. Opin. Neurobiol.* *13*, 8–15.
- Tran, K.D., and Doe, C.Q. (2008). Pdm and Castor close successive temporal identity windows in the NB3-1 lineage. *Development* *135*, 3491–3499.
- Tran, K.D., Miller, M.R., and Doe, C.Q. (2010). Recombineering Hunchback identifies two conserved domains required to maintain neuroblast competence and specify early-born neuronal identity. *Development* *137*, 1421–1430.
- Truman, J.W., and Bate, M. (1988). Spatial and temporal patterns of neurogenesis in the central nervous system of *Drosophila melanogaster*. *Dev. Biol.* *125*, 145–157.
- Tsuji, T., Hasegawa, E., and Isshiki, T. (2008). Neuroblast entry into quiescence is regulated intrinsically by the combined action of spatial Hox proteins and temporal identity factors. *Development* *135*, 3859–3869.
- Venken, K.J., and Bellen, H.J. (2012). Genome-wide manipulations of *Drosophila melanogaster* with transposons, Flp recombinase, and Φ C31 integrase. *Methods Mol. Biol.* *859*, 203–228.
- Wheeler, S.R., Kearney, J.B., Guardiola, A.R., and Crews, S.T. (2006). Single-cell mapping of neural and glial gene expression in the developing *Drosophila* CNS midline cells. *Dev. Biol.* *294*, 509–524.
- Wheeler, S.R., Stagg, S.B., and Crews, S.T. (2008). Multiple *Notch* signaling events control *Drosophila* CNS midline neurogenesis, gliogenesis and neuronal identity. *Development* *135*, 3071–3079.
- Wheeler, S.R., Stagg, S.B., and Crews, S.T. (2009). MidExDB: a database of *Drosophila* CNS midline cell gene expression. *BMC Dev. Biol.* *9*, 56.
- Yamasaki, Y., Lim, Y.M., Niwa, N., Hayashi, S., and Tsuda, L. (2011). Robust specification of sensory neurons by dual functions of *charlatan*, a *Drosophila* NRSF/REST-like repressor of *extramacrochaetae* and *hairy*. *Genes Cells* *16*, 896–909.

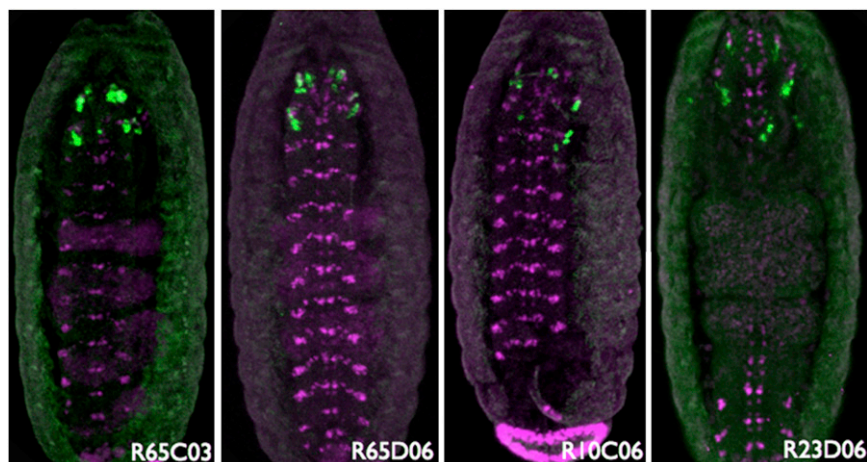


Figure S1. Background Quad Pattern of Four Neuronal Clusters, Related to Figure 3

Common expression pattern in the stage 16 subesophageal CNS observed in nearly all of the GAL4 lines. All panels show ventral views of the CNS of the indicated *GAL4 UAS-nGFP* lines stained for Eve (magenta) and nuclear GFP (green). Anterior is at the top.