RESEARCH ARTICLE

TECHNIQUES AND RESOURCES

Atlas-builder software and the eNeuro atlas: resources for developmental biology and neuroscience

Ellie S. Heckscher^{1,¶}, Fuhui Long^{2,¶}, Michael J. Layden^{1,*}, Chein-Hui Chuang¹, Laurina Manning¹, Jourdain Richart¹, Joseph C. Pearson³, Stephen T. Crews³, Hanchuan Peng^{2,‡}, Eugene Myers^{2,§} and Chris Q. Doe^{1,**}

ABSTRACT

A major limitation in understanding embryonic development is the lack of cell type-specific markers. Existing gene expression and marker atlases provide valuable tools, but they typically have one or more limitations: a lack of single-cell resolution; an inability to register multiple expression patterns to determine their precise relationship; an inability to be upgraded by users; an inability to compare novel patterns with the database patterns; and a lack of three-dimensional images. Here, we develop new 'atlas-builder' software that overcomes each of these limitations. A newly generated atlas is three-dimensional, allows the precise registration of an infinite number of cell type-specific markers, is searchable and is openended. Our software can be used to create an atlas of any tissue in any organism that contains stereotyped cell positions. We used the software to generate an 'eNeuro' atlas of the Drosophila embryonic CNS containing eight transcription factors that mark the major CNS cell types (motor neurons, glia, neurosecretory cells and interneurons). We found neuronal, but not glial, nuclei occupied stereotyped locations. We added 75 new Gal4 markers to the atlas to identify over 50% of all interneurons in the ventral CNS, and these lines allowed functional access to those interneurons for the first time. We expect the atlas-builder software to benefit a large proportion of the developmental biology community, and the eNeuro atlas to serve as a publicly accessible hub for integrating neuronal attributes - cell lineage, gene expression patterns, axon/dendrite projections, neurotransmitters - and linking them to individual neurons.

KEY WORDS: Gal4 lines, Atlas, Gene expression, Interneurons, Neuronal diversity

INTRODUCTION

To understand how embryos develop, it is essential to uniquely identify cells and cell populations. In addition, the recent advent of genetic tools for manipulating gene expression, expressing vital markers, and modifying/monitoring neuronal activity becomes more useful if the cells expressing these tools are uniquely identifiable. One approach for identifying distinct cell types in

**Author for correspondence (cdoe@uoregon.edu)

Received 5 February 2014; Accepted 14 April 2014

developing embryos is the generation of gene expression atlases. For example, the Eurexpress and the Allen Brain Atlas document thousands of gene expression patterns in the mouse embryo and brain, respectively (Diez-Roux et al., 2011; Lein et al., 2007). Similarly, FlyBase has a database of gene expression patterns in the developing fly embryo (Tomancak et al., 2007), and cell typespecific atlases have been generated for different stages of the Drosophila CNS (Beckervordersandforth et al., 2008; Broadus et al., 1995; Ito et al., 1995; Landgraf et al., 1997; Rickert et al., 2011; Wheeler et al., 2006, 2009). Despite the tremendous utility of these and similar databases, they typically have one or more limitations: a lack of single-cell resolution; an inability to register multiple expression patterns to determine their precise relationship; an inability to be upgraded by users; an inability to compare novel patterns with the database patterns; and a lack of a three-dimensional (3D) representation of the patterns.

Here, we develop new 'atlas-builder' software that overcomes each of these limitations. Our software can be used to create an atlas of any tissue in any organism with stereotyped cell positions. Newlygenerated atlases are 3D, allow the precise registration of an infinite number of markers, are searchable and are open-ended because additional markers can be added. The 'atlas-builder' software is a publicly available resource (http://www.vaa3d.org). As a proof of principle, we generated an atlas of the late embryonic Drosophila CNS. We populated the atlas with eight transcription factor markers that allowed us to identify the major cell types in the CNS [motor neurons (MNs), glia, neurosecretory cells and interneurons]. To document the value of the atlas-builder software, we next added 75 Gal4 lines that are expressed in different, small subsets of neurons (fewer than 10 per hemisegment). This dramatically increases the number of identified interneurons in the CNS (to over 50%) and provides new tools for manipulating and monitoring neuronal activity. The resulting 'eNeuro atlas' is also a publicly available resource (http://www.doelab.org/eNeuro-Atlas.html).

RESULTS

Generation and characterization of the atlas-builder software

We illustrate the utility of the atlas-builder software, using the late *Drosophila* ventral CNS as reference tissue. To generate an atlas of the *Drosophila* CNS, it was essential to have a single reference marker that labeled neurons distributed in all regions of the CNS; this 'fiduciary' marker can be used as a landmark to add all subsequent markers to the atlas. The Even-skipped (Eve) transcription factor is detected in the nuclei of a well-characterized subset of neurons distributed in all regions of the CNS, thus making it an excellent candidate for a fiduciary landmark (Doe et al., 1988; Frasch et al., 1987; Fujioka et al., 1999). Here, we focus on the second thoracic, third thoracic and first abdominal (T2, T3, A1) segments of the stage 16 embryonic CNS.



¹Institute of Neuroscience, Institute of 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 275995, USA.

^{*}Present address: Whitney Laboratory for Marine Bioscience, University of Florida, Saint Augustine, FL 32080, USA. [‡]Present address: Allen Institute for Brain Science, Seattle, WA 98103, USA. [§]Present address: Max-Planck Institute for Molecular Cell Biology and Genetics, Pfotenhauerstrasse 108, Dresden 01307, Germany. [¶]These authors contributed equally to this work

Each of these segments has a set of unpaired midline cells plus bilaterally symmetric left and right 'hemisegments'. We chose these three segments because they contain much of the known diversity in the ventral CNS: T2, T3 and A1 have important segment-specific differences; and A1 serves as a proxy for the closely related posterior A2-A7 segments (Schmid et al., 1999). We developed computational tools to map the position of Eve+ nuclei from multiple samples (see Fig. 1A-C). These tools generate a 'single-marker' atlas (supplementary material Data File S1) and a list of associated statistics (supplementary material Data File S2). Each neuron is represented by a sphere, the size and position of which represents the average size and position of all data collected on that neuron. We found some variability in neuron position between embryos. Importantly, however, the average position of an Eve+ neuron was highly reproducible with an average s.d. of up to 1.0 cell diameter in all three axes (Fig. 1D).

Next, we tested whether all CNS cell types have equally reproducible cell body positions. To identify distinct neural subtypes, we generated single-marker atlases for additional transcription factor (TF) patterns: phospho-Mothers against Dpp (pMad; all MNs), Bar (subset MNs), Dimmed (Dimm; most neurosecretory cells), Reversed polarity (Repo; glia), Apterous (Ap; subset interneurons, subset neurosecretory cells), Engrailed (En; subset interneurons, subset glia) and Hb9 (Exex – FlyBase; subset interneurons, subset MNs) (Fig. 2A-G; supplementary material Data File S3) (Benveniste et al., 1998; Bourgouin et al., 1992; Broihier and Skeath, 2002; Garces et al., 2006; Layden et al., 2006; Lundgren et al., 1995; Mccabe et al., 2003; O'Connor, 2005; Odden et al., 2002; Park et al., 2004, 2008; Patel et al., 1989; Xiong et al., 1994). Notably, we built our atlas at a time point slightly before a subset of neuropeptide (Dimm+) neurons would begin to project out of the ventral nerve cord (VNC; Allan et al., 2003; Miguel-Aliaga, 2004), thus rendering our Dimm + neurons pMad–. We found that all neuronal markers had highly reproducible average nuclear positions (Fig. 2H). In contrast, glia often had highly variable nuclear positions (Fig. 2H), even though their cellular positions are predictable (Beckervordersandforth et al., 2008; Ito et al., 1995). We concluded that interneurons, MNs and neurosecretory cells have highly stereotyped nuclear positions. Importantly, the same steps can be used to create an atlas of any marker in any organism, as long as cell or nuclear positions are stereotyped.

Single-marker atlases can be merged to identify marker co-expression

A single-marker atlas, as we have generated for each of the eight TFs, is of limited usefulness. But if cell body positions are sufficiently stereotyped, we should be able to use Eve to register, or merge, all atlases into the same 3D volume. This would identify cells co-expressing two or more markers, as well as allow us to determine the percentage of cells in the CNS that express at least one marker (or 'coverage'). We created computational tools to merge atlases, and used these tools to generate a single 'TF atlas' displaying all eight TFs expressed in 714 cells (Fig. 3A; supplementary material Data File S4).



Fig. 1. Atlas building pipeline. (A) Image acquisition. A 3D rendering of a stage 16 CNS stained with anti-Even-skipped (Eve). Segments T2-A1 shown, anterior upwards, ventral view, rotated 45° about anterior-posterior axis. (B) Segmentation. Image from A after segmentation, with each nucleus as a distinct object and assigned a random color. (C) Annotation. Cells in A1R (white box in B). Each segmented object is given a user-defined name/marker prefix (e.g. U4/eve) and a hemisegment suffix (e.g. A1R). Anterior is towards the top, z-projection. Scale bar: 10 µm. (D) Registration. A 3D rendering of cells from nine segmented, annotated images. Stacks are overlaid to illustrate the variation in position between embryos for each Eve+ cell. (D') The average s.d. in cell body position for all Eve+ cells is a measure of the variance in position for each set of nuclei. (E,F) User interface. In the final Eve single marker atlas, each nucleus is shown by a single sphere where diameter represents mean nuclear size and location represents mean x,y,z position of each nucleus. A view of the Eve pattern in the atlas viewer was obtained by dragging and dropping the '.apo' file (supplementary material Data File S1) into the V3aaD window. See supplementary material Data File S9 for a description of functions 1-21.



Fig. 2. Cell bodies of neurons and a subset of glia have predictable positions in the embryonic CNS. (A-G) In vivo transcription factor expression patterns (green, names at top) co-stained with Even-skipped (Eve, magenta). Maximum intensity projections; co-expressing cells appear white. Scale bar: 10 µm. FlyBase full names: pMad, phospho-Mothers against Dpp; Dimm, Dimmed; Repo, Reversed polarity; Bar, BarH1/BarH2; Hb9, Extra-extra; Ap, Apterous; En, Engrailed. (A'-G') Single-marker atlases for each pair of markers shown in A-G. Co-expressing cells are represented by overlapping green/magenta spheres, because single marker atlases do not account for co-expression. (See below for merged atlas showing co-expression.) (H) Histogram showing the distribution of average s.d. of nuclear positions grouped by gene expression. Only glia (Repo+) have variable nuclear positions. (I) The average s.d. in nuclear position of the uniquely identifiable Eve+, FoxO+ neuron (yellow) is similar to that of the Eve+ lateral cluster of neurons (red) and of other laterally positioned neurons (black).

Importantly, the TF atlas merges overlapping cells that express two or more markers into a single cell.

To test the accuracy of the TF atlas - i.e. the number of false positives (two markers falsely predicted to co-express in a cell) or false negatives (two markers incorrectly assigned to separate cells when they are actually co-expressed) – we experimentally determined the actual overlap of all eight markers from our own stainings or from the literature (e.g. Fig. 3B,C). When we predicted co-expression solely based on proximity between cell body positions, we correctly identified 65% of all examples of in vivo co-expression (although the correct percentage rose to over 80% for neurons located in the medial CNS) (Fig. 3D; supplementary material Data File S5). When we added Rule 1, "if a pair of cells in one hemisegment is predicted to co-express a marker, then these pairs also co-express the marker in the other hemisegments", we increased correct predictions to 96%, but this resulted in a high number of false positives (Fig. 3E). Ultimately, we settled on a set of conditions called Rule 2 that maximized true positives and negatives, while minimizing false positives and negatives (Fig. 3E; see Materials and Methods). Thus, we used the experimentally determined TF colocalization data as 'ground truth' to develop algorithms that allowed us to optimally determine whether two markers - stained in different embryos - are co-expressed or adjacent within the CNS.

The eNeuro atlas: a single atlas containing eight transcription factors and 75 Gal4 lines

We wanted to add a large number of new markers to generate a comprehensive eNeuro atlas. This would help us to validate our atlasbuilder software and to provide new markers for previously

uncharacterized neurons (mostly interneurons). We chose to add Gal4 lines expressed in 1-10 neurons per hemisegment, which gave us the dual benefit of increasing known molecular markers for interneurons and of providing functional access to these neurons for the first time. From a previous large-scale screen of 5000 Gal4 lines (Manning et al., 2012), we selected 75 lines with strong, regular expression patterns at late embryonic stages. We used each Gal4 line to express a nuclearlocalized GFP, and created a single-marker atlas for each line (Fig. 4; supplementary material Data File S6). The 75 single marker atlases contained a total of 1998 cells (pre-merge) within the T2-A1 CNS, and 91% of the Gal4+ neurons had highly predictable positions (Fig. 4). We merged the Gal4 atlases with the TF atlas into a single, comprehensive 'eNeuro' atlas containing all 83 markers (supplementary material Data File S7). The eNeuro atlas contained fewer cells (1362) than the premerge numbers due to some neurons expressing multiple Gal4 lines, which were collapsed into a single neuron in the eNeuro atlas. Notably, we were able to experimentally validate 50% of the predicted co-expression in the eNeuro atlas (e.g. Fig. 5; supplementary material Data File S5). Thus, the eNeuro atlas integrated the co-expression of 83 molecular markers at single-cell resolution.

In the rodent spinal cord, over 20 different molecular subtypes of interneurons have been identified, but it is likely that interneuron diversity is much greater (Alaynick et al., 2011; Arber, 2012). Here, we used the eNeuro atlas to analyze the extent of interneuron diversity in the Drosophila embryonic ventral CNS. First, we assessed the coverage of the interneuron population. We defined interneurons as being negative for all motor neuron (pMad), neurosecretory (Dimm) and glial (Repo) markers (supplementary material Data File S8). The eNeuro atlas identified 59-60% of the interneuron population in



Fig. 3. Atlases can predict a majority of in vivo-determined co-expression. (A) The TF atlas is a merge of the eight TF single marker atlases. Anterior is upwards. (A') Cross-section of neurons in A1 of the TF atlas; clear zone is neuropil; dorsal is upwards. (A") The pan-neuronal marker Elav+ (Robinow and White, 1988) shows CNS morphology matching that of the atlas. Single optical cross-section (xz view) shown; clear zone is the neuropil; dorsal is upwards. Scale bar: 10 µm. (B) Atlas-predicted co-expression (white) for Hb9 (green) and pMad (magenta). Data used to build each atlas were from separate populations of embryos. (C) An embryo co-stained for Hb9 (green) and pMad (magenta) showing in vivo co-expression (white). Maximum intensity projection. Scale bar: 10 µm. (D) Scoring the accuracy of co-expression prediction based solely on distance between nuclei for co-expression among all TF single marker atlases ('all'), or for neurons with cell bodies located in the medial ('M'), intermediate ('I') or lateral ('L') one-third of the CNS. (E) Scoring the accuracy of co-expression prediction among all TF single-marker atlases, when additional assumptions are invoked (see Materials and Methods for definitions of rules). Assumptions can be changed to maximize true positives (Rule 1) or minimize false positives (Rule 3).

T2-A1 depending on the method of counting total neuron numbers (Fig. 6A,B). Analysis of the eNeuro atlas shows interneurons to be extremely diverse. We describe interneuron diversity for A1, but similar results were observed in thoracic hemisegments (Fig. 7; supplementary material Table S1 and Data File S8). The interneurons in A1 can be subdivided into 80 molecularly distinct subtypes, where a subtype is defined by a unique combination of markers (Fig. 7A,C). Furthermore, 43 combinations of markers label a unique single interneuron (Fig. 7B,C). Most of the molecular subtypes found in A1 were also found in other segments (Fig. 7D). Some molecular subtypes were found only in A1, and some were unique to thoracic segments (Fig. 7D). This indicates that at different anterior-posterior positions along the nerve cord there are molecularly distinct interneuron subtypes. Thus, addition of new markers to the eNeuro atlas, which is possible when using our atlas-builder software (see Discussion), should rapidly increase the number of molecularly distinct interneurons. We conclude that interneurons are characterized by extreme molecular diversity, and predict that each interneuron will have a unique, reproducible, transcriptional profile.

The eNeuro atlas accurately identifies all known MNs and midline cells, and provides new genetic tools to characterize neuronal diversity

Although our primary focus was on interneurons, we also used the eNeuro atlas to analyze midline cells and MNs. These populations are well characterized, which helps us to determine the accuracy of the eNeuro atlas. Midline cells contain the progeny of single-minded (sim)+ precursors, dorsal median cells and channel glia (Chiang et al., 1994; Ito et al., 1995; Kearney et al., 2004; Wheeler et al., 2006). We found that the position of midline cells in the eNeuro atlas strongly resembled the in vivo position of the midline cells (Fig. 8A,B; supplementary material Table S2 and Data File S8). Furthermore, we identified five Gal4 lines labeling subsets of midline cells that had not previously been distinguished at the molecular level (Fig. 8C; supplementary material Table S2). MNs include a set of 35 somatic MNs per hemisegment that can be molecularly subdivided based on Eve, Hb9 and Bar expression as well as visceral MNs (Landgraf and Thor, 2006; Miguel-Aliaga et al., 2008). The eNeuro atlas identified all Eve+, Hb9+ or Bar+ somatic MNs (Fig. 8D), and a group of poorly characterized Eve/ Hb9/Bar triple-negative MNs (Fig. 8D). Furthermore, we found unique molecular profiles for the majority of MNs. In A1R alone we identified 23 subtypes of MNs (Fig. 8E; supplementary material Table S3 and Data File S8), whereas previously only seven molecularly distinct MN subtypes were known (Landgraf and Thor, 2006). We conclude that the eNeuro atlas (1) accurately identifies all known MNs and midline cells; (2) reveals unprecedented molecular diversity in MNs and midline cells; and (3) provides novel tools for studying cells that were previously genetically inaccessible (see Discussion).

DISCUSSION

To our knowledge, atlases generated with our 'atlas-builder' software are unique in allowing the integration of an unlimited number of new markers. For example, the Allen Brain Atlas has an infinitely expandable number of pages for gene expression profiles, but the expression patterns are not linked, and it is unclear which fine structures co-express different markers. In contrast, we use a single fiduciary marker to register new patterns to the existing atlas; thus, an unlimited number of markers can be co-stained with the fiduciary marker, added to the atlas and their overlap predicted with high accuracy. A second important feature of an atlas generated with our software is that users can perform a 'BLAST'-like search to predict which known cells express a new marker. Although perfect accuracy is probably impossible due to existing biological variation, the ranking of nearby candidate cells will greatly facilitate identification of cells that express a new marker. Our tools are publicly available, and we provide detailed instructions to guide new users in creating and 'BLASTing' atlases for new markers (supplementary material Data File S9).

Previously, only a handful of Drosophila embryonic interneurons were individually identifiable using a combination of molecular marker expression and cell body position – e.g. pCC, EW1-EW3, Apterous+ neurons and Eve+ lateral interneurons (Doe et al., 1988; Higashijima et al., 1996; Park et al., 2004; Thomas et al., 1984; Tsuji et al., 2008). The eNeuro atlas expands the number of identified interneurons to 958, each of which has been assigned a unique name, and each of which is described by quantitative information on cell body position, size and frequency of appearance, as well as its Gal4 line expression profile (supplementary material Data File S8). Nevertheless, the eNeuro atlas is not yet 'saturated' for interneurons: $\sim 40\%$ of interneurons do not express any marker (Fig. 5). Importantly, it is simple to add new markers to the eNeuro atlas (see below). We anticipate that additional markers will allow the unique identification of every interneuron in the ventral CNS. In the future, we will provide updated marker atlas files and add atlas

2527

2 Ш

 \bigcirc

در دور در مرتبع المرتبي مرتبع المرتبي	20 20 20 20 20 20 20 20 20 20	200 200 200 200 200 200 200 200 200 200	مد می مدر مد کر کردر ور مرکز در	2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000		ور در اور در در اور در در اور در در اور در در
A227	R10D02	R10D07	R11F02	R11F04	R12A10	R12B12
	- 20 - 20			4.5 ¹⁷ ₂₀ + 2 ¹⁷ 2 2.52 ¹⁷ 20 + 2 ¹⁷ 2 5.5 ¹⁹ 2 2.5 ¹⁹ 2 2	$\begin{array}{c} & & & & & & & & & & & & & & & & & & &$	
R12E03	R12E07	R13D08	R13E06	R14E12	R14H08	R15D07
	200 2017 2017 2017 2017 2017 2017 2017 2	222222 3122232 3122232 322234 322234 322234 322234 322232 32232 323 32 32	200 - 200 2012 - 200 2012 - 200 2010 - 200 2010 2010 - 200 2010 2010 2010 2010 2010 2010 2010 2010 2	200 - 200 200 - 200 200 200 - 200 200 200 - 200 200 200 - 200 200 200 - 200 200 200 200 200 200 200 200 200 200		
R15E11	R15F01	R17C08	R17F04	R18A04	R19A07	R19H09
	د دور در مرد در در مرد مرد در در مرد مرد مرد مرد مرد مرد مرد مرد مرد مرد مرد	ور مرکز در در میگرد مرد مرد میگرد در	$\begin{array}{c} & & & & & & & & & & & & & & & & & & &$	د دولاد ه ماد د ولارد وال ماد د ولارد وال ماد م ولارد وال	مر می اور مر می ورود در اور مرد و و و و و و و و و و و و و و و و و و و	من مد ¹⁰⁰ مد ¹⁰⁰ مد ¹⁰⁰ در ¹⁰⁰ مالا ¹⁰⁰ مراجع مالا ¹⁰⁰ مرجع
R20A12	R21A04	R21B12	R21C06	R22C10	R23A02	R23G10
ر دولارده مروز مروز مروز مروز مروز مروز مروز مروز مروز مروز مروز	2000 100 100 1000 100 100 1000 100 100 1000 100 1	20 30 30 30 30 30 30 30 30 30 3	10 - 10 - 10 - 10 10 - 10 - 10 - 10 10 - 11 - 10 10 - 11 - 10 - 10 - 11 - 10 - 10 -	- 200 - 200 - 200 - 200 - 200 - 200 - 200		20 00000000000000000000000000000000000
R23G11	R25C03	R25D09	R25G03	R25H11	R26B03	R26H02
10.432 - 20 10.132 - 20 10.132 - 20 10.132 - 20	20 20 20 20 20 20 20 20 20 20 20 20 20 2	20 20 20 20 20 20 20 20 20 20 20 20 20 2	20 20 20 20 20 20 20 20 20 20 20 20 20 2	هو دېلو د مو د هم د ځارو د مولاد د مې ولو د مولاد د مې ولو د مولاد		
R27A09	R29A12	R30A02	R30E03	R31G06	R33E02	R33G07
ر در دور مرد مرود در مرد مرود در مرد مرود در مرد مرود در مرود مرود در مرود در مرود در مرود در مرود مرود در مرود مرود مرود در مرود مرود مرود مرود مرود مرود مرود مرود	"""""""" """""""" """"		20 20 20 20 20 20 20 20 20 20 20 20 20 2	<u>مروم میں میں میں میں میں میں میں میں میں می</u>		
R34D08	R35D09	R35D11	R37G07	R38C08	R38E07	R45D07
		200 - 200 200 - 200 200 - 200 - 200 - 200 - 200 	ر د در	ده، در می مدر در می مرد در می در می مرد در می در مرد مرد در می در می	20 20 20 20 20 20 20 20 20 20 20 20 20 2	2
R46A06	R47H04	R48G05	R50B02	R52B03	R53B03	R53E03
در نواور ه مرور روز مرور روز مرور مرور مرور مرور	۵۵ ۵۵ می می می ۱۹۹۵ می می می ۱۹۹۵ می می			59 89 99 99 99 99 99 99 99 99 99 99 99 99	یں میں کروٹ در مرکز کر	200 - 200 200 - 200 200 - 200 200 - 200 - 200 - 200 -
R53F07	R54B03	R54E06	R54E08	R55C06	R55E04	R59D01
200 - 20 200 - 100 - 620 200 - 100 - 620 200 - 100 - 200 200 - 200 - 200 200 - 200 - 200 200 - 200 - 200	در میں			م می میں م می میں میں میں میں میں میں میں میں	20 20 20 20 20 20 20 20 20 20 20 20 20 2	30 30 30 30 30 30 30 30 30 30 30 30 30 3
R64A12	R64B05	R64B06	R65B10	R65D03	R65H11	R66C01
2000 2013 2014 2014 2014 2014 2014 2014 2014 2014	-20 -20 -20 -20 -20 -20 -20 -20	2000 2000 2000 2000 2000 2000 2000 200	22 2455 - 247 - 25 2455 - 247 - 25 2456 - 247 - 25 2456 - 252 - 2556 -	20 20 20 20 20 20 20 20 20 20 20 20 20 2		
R71D07	R71F10	R74B12	R75G10	9-58		
average stdev (cell diameters)						
	more	9%				

Fig. 4. Single-marker atlases for 75 Gal4 lines expressed in small subsets of neurons. Single-marker atlases for 75 Gal4 lines. Gal4+ nuclei are shown in black, with the name underneath; Eve+ nuclei are white; anterior is upwards; ventral view with 45° rotation about the anterior-posterior axis. For a list of all individual cells and statistics, see supplementary material Data File S5. Histogram showing the distribution of average s.d. of nuclear position for cells labeled by Gal4 lines.

files from other users with their permission. In addition, we plan to add markers that identify individual neuroblast lineages.

Until now, few genetic tools existed to manipulate specific subtypes of embryonic interneurons. Here, we have mapped 75 Gal4 lines that each label small, reproducible subsets of embryonic interneurons (Fig. 4). Some lines show the same pattern in larvae while others change pattern over time (data not shown). Importantly, the use of 'immortalization' genetics (Buckingham and Meilhac, 2011) can make many of these Gal4 markers useful at later stages of development or in the adult. Furthermore, many interneurons in the eNeuro atlas were labeled by multiple Gal4 lines (Fig. 6A,C; supplementary material Fig. S1), thereby raising the possibility of using intersectional genetics to target single interneurons, because the vast majority of these lines were made with a defined

cis-regulatory DNA fragment (Pfeiffer et al., 2010). Thus, this work provides a significant advance in our ability to characterize and test the function of individual identified interneurons.

We also provide important new tools for the study of MNs and midline cells (Figs 7,8; supplementary material Tables S2,S3). These can be used to trace the axon projection of a previously uncharacterized Eve/Hb9/Bar triple-negative motor neuron (supplementary material Fig. S2). In the future they could be used to study the specification, axon pathfinding, target recognition and role in motor behavior. We also made a novel discovery regarding the well-studied Hb9+ MNs: although most Hb9+ MNs project out of the intersegmental nerve (ISN), some have been reported to project out of the segmental nerve root (SN) to transverse somatic muscles, similar to the projection pattern of the Bar+ MNs (Broihier



Fig. 5. Single marker atlases can be unified in atlases. (A,B') Atlaspredicted and *in vivo*-validated co-expression (A) R35D09 and Hb9 are predicted to co-express (white) in dMP2 (arrowhead) and vMP2 (arrow). (B,B') *In vivo* validation of R35D09 and Hb9 co-expression in dMP2 and vMP2. Anterior is upwards. Scale bar: 10 µm. (C-E) Atlaspredicted and *in vivo*-validated co-expression. (C) R12E07 and Repo are predicted to co-express (white) in the B-SPG (arrow) glial cell. (D,D') *In vivo* validation of R12E07 and Repo co-expression in B-SPG. (E) R12E07-driving membrane GFP reveals the characteristic subperineural glia morphology in the B-SPG. Scale bars: 10 µm. Anterior is upwards.

and Skeath, 2002; Garces et al., 2006). Our atlas did not contain any Hb9+Bar+ co-expressing neurons, which was confirmed by experiments showing no Hb9+ neurons projecting out of the SN to transverse muscles (supplementary material Fig. S3). We conclude that the Hb9+ MNs do not co-express Bar or project out of the SN, at least not at embryonic stage 16. Thus, the eNeuro atlas provides tools for the developmental and functional analysis of individual and poorly characterized groups of MNs.

Limitations of the atlas approach include potential annotation errors that exaggerate positional variability, and stochasticity in Gal4 expression that results in some cells being unrepresented or unilaterally represented. It is also possible that the eNeuro atlas contains errors in cell merging. Errors are likely to be rare, however, because we examined 171 of 243 predicted Gal4/TF co-expressing cells and found only one error, which was manually corrected (supplementary material Data File S5); the remaining Gal4/Gal4 merge predictions are likely to have a similarly high accuracy.

The genetic analysis of neural development in *Drosophila* has led to many fundamental discoveries that also apply to mammalian neural development. The study of interneuron specification and



Fig. 6. eNeuro atlas labels a majority of interneurons. (A) All interneurons in the eNeuro atlas. Each neuron is color-coded according to the number of markers it expresses: 1 (gray), 2 (blue), 3 (teal), 4 (green), 5-6 (yellow), >6 (red). Top panel, anterior is upwards; bottom panel, dorsal is upwards. (B) eNeuro atlas identifies a majority of interneurons in the ventral CNS. White bars indicate coverage based on 270 interneurons per hemisegment (Rickert et al., 2011). Black bars indicate coverage based on total number of Elav+pMad–Dimm– neurons (see Materials and Methods). (C) A fraction of the interneurons in the eNeuro atlas is multiply labeled. Colors as in A.

function is in its infancy in both *Drosophila* and mammals. We suggest that this atlas of interneuron diversity and tools for functional analysis will facilitate the study of *Drosophila* interneurons, and that this will have a direct impact on the study of mammalian interneurons. This is an important area of research, as interneuron dysfunction is thought to underlie autism, epilepsy and other debilitating human neural disorders.



Fig. 7. eNeuro atlas subdivides interneurons into 112 distinct molecular subtypes. (A) All interneurons included in the eNeuro atlas. Each molecularly different subtype is represented by a different color (supplementary material Data File S8 and Table S1). Top panel, anterior is upwards; bottom panel, dorsal is upwards. (B) The subset of interneurons (from 6A) for which a single neuron in each hemisegment has a unique molecular profile. View as in A. (C) eNeuro atlas reveals interneuron molecular diversity. White bars indicate molecularly distinct interneuron subtypes (in A). Black bars indicate molecular yunique single interneurons (in B). (D) There are 112 distinct molecular subtypes in the eNeuro atlas: 80 subtypes in A1 (green), 82 in T2 (purple) and 86 in T3 (red).



Fig. 8. Midline and motor neuron molecular diversity in the eNeuro atlas. (A,A') Composite image of a stage 16 *sim-Gal4 UAS-tau-GFP* embryo stained with anti-GFP (green) and anti-Sim (magenta). (A') Sagittal view. (B,B') Atlas view of Sim+ (magenta) and Sim– (green) midline cells. (B) Dorsal view; anterior is upwards. View as in A,A'. (C) Atlas sagittal view of midline cells with each subtype represented as a different color (supplementary material Table S3). (D) pMad+ MNs subdivided in four molecular subtypes: Bar+pMad+ cells (red); Hb9+pMad+ (green) cells; Eve+pMad+ (blue); and poorly characterized triple-negative pMad+ (white) MNs. There is no overlap between Hb9+, Bar+ and Eve+ cells. Anterior is upwards. (e) The eNeuro atlas increases the number of molecularly distinct motor neuron subtypes in A1-23, each of which indicated by a different color (supplementary material Table S3). Anterior is upwards.

MATERIALS AND METHODS

Fly genetics

We used standard methods for propagating fly stocks. Transcription factor patterns were determined in wild-type flies (*yellow white*). Gal4 patterns were determined by crossing Gal4 [BN-gal4 (Sato et al., 1999), Ap-Gal4 (Bloomington #3041); 9-58-Gal4 (Ulrike Heberlein, Janelia Farm Research Campus, JFRC); a227-Gal4 (Julie Simpson, JFRC); Hb9-Gal4 (Broihier and Skeath, 2002); and all other Gal4s (JRFC) (Pfeiffer et al., 2008)] males to females containing *UAS-GFP* with a nuclear localization sequence or membrane tether (Bloomington #6452, #32184, #32197).

Embryo immunostaining and staging

We used standard methods to stain *Drosophila* embryos (Manning et al., 2012). Primary antibodies were mouse Eve 2B8 (1:50), mouse En 4D9 (1:10), and rat Elav (1:10), all from the Developmental Studies Hybridoma Bank (DSHB) developed under the National Institute of Child Health and Human Development (NICHD) and maintained by the University of Iowa; guinea pig Dimm (1:100; a gift from S. Thor, Linkoping University, Sweden); rat Repo (Campbell et al., 1994) (1:1000); rabbit pMad (pS1 1:300; Peter Ten Dijke, Leiden University Medical Center, Leiden, The Netherlands); chicken GFP (1:1000; Aves, Tigard, OR); rabbit Hb9 (Odden et al., 2002) (1:1000); and guinea pig Fox (Nechipurenko and Broihier, 2012) (1:20). Secondary antibodies were from Invitrogen/Molecular Probes and were used according to the manufacturer's instructions. Stage 16 embryos were selected for imaging based on morphological criteria (Pereanu et al., 2007).

Image acquisition

Images were acquired on a BioRad Radiance, Zeiss 700 or 710 confocal microscope with a $63 \times$ objective collected at 0.3 or 0.5 μ m steps. Images were cropped in ImageJ (NIH) to include segments T2-A1. Some images

were used to generate atlases, whereas others were used to validate co-expression or for figures.

Segmentation

We developed a Vaa3D (Peng et al., 2010) (http://Vaa3d.org/) plug-in to segment cells. First, a TIFF image was smoothed with a median filter and/or Gaussian filter to reduce noise. We then separated the image into background and foreground pixels to make the later segmentation step more robust and computationally efficient. We provided three thresholding methods: adaptive thresholding, global 3D thresholding and global 2D thresholding. After thresholding, we used a watershed algorithm to segment cells based on either the intensity information or the shape information of foreground pixels. For each transcription factor we segmented and annotated a median of 10 image stacks. We segmented a median of five image stacks per Gal4 line.

Annotation

The segmented cells were annotated using VANO (Peng et al., 2009) software (http://vano.cellexplorer.org/). Where possible, neuron names were taken from the literature and had the following syntax: 'name/marker hemisegment' (e.g. pCC/eve T2R or A/19H09 T2R). Segment identity was based on neuron position at mid stage 16; the U1-U5 neurons defined the posterior boundary of each segment. Ruby scripts were used to detect annotation errors, such as typos (code is available upon request). To minimize annotation, we used only Eve+ medial neurons (aCC, pCC, RP2, U1-U5) in each segment, as this was sufficient for optimal registration (data not shown). For new markers, annotation could be carried out on just one cell of interest or on all cells.

Registration

A newly developed Vaa3D plug-in called 'AtlasBuilder' uses rigid registration to map the annotated image stacks into a standard coordinate space defined by a target stack. The target stack was automatically determined from a randomly selected subset of stacks as the one that minimized the summed cell position variations (supplementary material Data File S1). We computed an affine transformation that best aligned the EVE channel of a subject stack to that of the target stack. This transformation was then applied to all channels of an image stack. Note that this type of rigid registration does not use local warping to align the neurons. It calculates an average position for each nucleus, which acts as unique x,y,z coordinates in virtual space; it also calculates the s.d. of the variance of actual centers from the mean, which acts as a measure of the stereotypy of each cell. For a complete list of data produced by registration, see the 'cellStatistics.xls' file for each atlas (supplementary material Data Files S2, S3 and S5). 'AtlasBuilder' can build atlases with different confidence thresholds, which are calculated based on how often a given cell appears in image stacks. For transcription factor atlases, we included only cells that were found in 33% or more of the image stacks; for Gal4 atlases, we included all cells.

Co-expression prediction and confirmation

To establish *in vivo* co-expression between markers, we used data from previously published reports, or performed co-labeling experiments ourselves. We determined the individual identity of each cell based on the relative position of the cell within the marker pattern, as we did for manual annotation. We also scored for projection patterns/cell morphology and for nuclear shape where these characteristics had been reported in the literature (supplementary material Data File S5).

To predict overlap between single marker atlases, we developed a 'Detect coexpressing cells' function in AtlasBuilder. This method takes '.apo' files (single atlases and/or atlases) and determines which cells are within a user-specified distance (we used six pixels, ~ average nuclear radius). We did not use symmetry or bilateral rules, although they are available (supplementary material Data File S9 for details). This plug-in generates two output files: a '. coexpress.apo' where predicted co-expression is white, and a '.coexpress.txt' file, which lists each candidate cell followed by all cells found within the specified radius.

To calculate prediction accuracy we calculated three parameters: 'true positives' – co-expression found both by atlas prediction and *in vivo* analysis; 'false negatives' – co-expression found only by *in vivo* analysis; 'false positive' – co-expression found only by atlas prediction. True positives and false negatives were scored as a percentage of all expected (n=133) *in vivo* co-expressions (supplementary material Data File S5). False positives were scored as a percentage of total predictions.

We used the following three rules.

Rule 1: if a cell in one hemisegment is predicted to co-express, then all examples of that cell also co-express; 96% true positive rate, 4% false negative rate and 51% false positive rate.

Rule 2: if the same cell in two hemisegments is predicted to co-express, then all examples of that cell also co-express; 84% true positives, 16% false negatives and 30% false positives.

Rule 3: if the same cell in three hemisegments is predicted to co-express, then all examples of that cell also co-express; 40% true positives, 60% false negatives and 6% false positives.

Merging

The 'Merge coexpressing cells' plug-in for Vaa3d 'AtlasBuilder' takes two types of input: '.apo' files and a user-generated look-up table text file (supplementary material Data File S10). The merging plug-in produces a new atlas with both '.apo' and 'cellstatistics.txt' files containing positional information for the new cells; these new positions are calculated from the average positions of all original cells, even if they come from two separate '.apo' files. In the resulting atlas, all markers expressing in a given cell are listed in the comments column of the '.apo' files.

To create the TF atlas, we assigned each cell a new name, which was a truncated version of the original names (i.e. aCC/eve T2R and aCC/pMad T2R became aCC T2R).

To create the eNeuro atlas, we discarded any low-quality cells: cells appearing in less than 33% of the annotated Gal4 line image stacks, or cells with an average s.d. of the nuclear position that was greater than a single cell diameter. We used the 'Detect coexpressing cells' function to generate a list of predicted co-expressions among all 83 single atlases. We used custom Ruby scripts to apply 'Rule 2' (see above). We then used Ruby scripts and MATLAB to view the matrix of co-expression as biograph objects (code is available upon request). We merged cells in the following order: MPs (dMP2, vMP2, MP1), Eve+ cells, Bar+ cells, Hb9+ cells, pMad+ cells, cells predicted to label with other transcription factors, all other cells. Each subtype was assigned a name; if the cell appeared in the TF atlas, we used the same name; all others (646 cells) were assigned names using 'IN' as a prefix followed by a three-digit number, such as 'IN094' (see supplementary material Data File S10 for complete list of merged cells).

Other calculations

Cell diameter was calculated using the mean volume of all cells and then solved for r (volume= $4/3\pi r^3$). The s.d. of cell position was calculated as *x*,*y*,*z* s.d. (std_x, std_y, std_z)/3 in Fig. 1. In Figs 2 and 4, average s.d. was calculated similarly, but we combined data from cell bodies in all hemisegments (e.g. aCC represents data from aCC/eve T2R, T2L, T3R, T3L, A1R and A1L).

We used two methods to estimate the number of interneurons in each segment: 540 interneurons, based on single interneuron DiI labeling (Campbell et al., 1994), which does not include unpaired midline neurons; or an average of 549 interneurons per segment (515 in T2, 542 in T3, 589 in A1), based on our own counts of Elav+ neurons minus pMad+ MNs and Dimm+ neurosecretory cells. To determine the interneuron coverage in the eNeuro atlas, we defined interneurons as triple negative for pMad (MNs), Dimm (neurosecretory cells) and Repo (medial glia), as well as having low positional variability (thus excluding lateral glia). This gave us an average of 324 interneurons per segment (Fig. 6; 337 in T2, 339 in T3, 295 in A1).

Acknowledgements

We thank Matt Clark, Sen-Lin Lai and Chris Wreden for comments on the manuscript; Maria Purice, Art Li, Andrew Reichold and Jimmy Kelly for help annotating marker patterns; and the *Drosophila* Stock Center (Bloomington, IN, USA) and Todd Laverty at the JFRC (VA, USA) for fly stocks.

Competing interests

The authors declare no competing financial interests.

Author contributions

M.J.L. and F.L. initiated the project; F.L., H.P. and E.M. developed software; J.C.P. and S.T.C. tested software prior to publication and analyzed midline cells; E.S.H., M.J.L., C.-H.C., L.M. and J.R. contributed to data acquisition, image processing, annotation and registration; E.S.H. merged atlases and performed data analysis; E.S.H. and C.Q.D. wrote the manuscript.

Funding

This work was funded by the American Heart Association [#0920025G postdoctoral fellowship to E.S.H.], the National Institutes of Health [27056 to C.Q.D.], the Howard Hughes Medical Institute (C.Q.D.), the National Institute of Neurological Disorders and Stroke [R01 NS64264 to S.T.C.] and the JFRC (F.L., H.P. and E.M.). Deposited in PMC for release after 12 months.

Supplementary material

Supplementary material available online at http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.108720/-/DC1

References

- Alaynick, W. A., Jessell, T. M. and Pfaff, S. L. (2011). SnapShot: spinal cord development. Cell 146, 178-178. e1.
- Allan, D., Pierre, S., Miguel-Aliaga, I. and Thor, S. (2003). Specification of neuropeptide cell identity by the integration of retrograde BMP signaling and a combinatorial transcription factor code. *Cell* **113**, 73-86.
- Arber, S. (2012). Motor circuits in action: specification, connectivity, and function. *Neuron* 74, 975-989.
- 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.
- Benveniste, R. J., Thor, S., Thomas, J. B. and Taghert, P. H. (1998). Cell typespecific regulation of the Drosophila FMRF-NH2 neuropeptide gene by Apterous, a LIM homeodomain transcription factor. *Development* **125**, 4757-4765.
- Bourgouin, C., Lundgren, S. and Thomas, J. (1992). apterous is a Drosophila LIM domain gene required for the development of a subset of embryonic muscles. *Neuron* **9**, 549-561.
- 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.
- Buckingham, M. E. and Meilhac, S. M. (2011). Tracing cells for tracking cell lineage and clonal behavior. *Dev. Cell* 21, 394-409.
- Campbell, G., Göring, H., Lin, T., Spana, E., Andersson, S., Doe, C. Q. and Tomlinson, A. (1994). RK2, a glial-specific homeodomain protein required for embryonic nerve cord condensation and viability in Drosophila. *Development* 120, 2957-2966.
- Chiang, C., Patel, N. H., Young, K. E. and Beachy, P. A. (1994). The novel homeodomain gene buttonless specifies differentiation and axonal guidance functions of Drosophila dorsal median cells. *Development* **120**, 3581-3593.
- Diez-Roux, G., Banfi, S., Sultan, M., Geffers, L., Anand, S., Rozado, D., Magen, A., Canidio, E., Pagani, M., Peluso, I. et al. (2011). A high-resolution anatomical atlas of the transcriptome in the mouse embryo. *PLoS Biol.* **9**, e1000582.
- Doe, C. Q., Smouse, D. and Goodman, C. S. (1988). Control of neuronal fate by the Drosophila segmentation gene even-skipped. *Nature* **333**, 376-378.
- 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.
- Garces, A., Bogdanik, L., Thor, S. and Carroll, P. (2006). Expression of DrosophilaBarH1-H2 homeoproteins in developing dopaminergic cells and segmental nerve a (SNa) motoneurons. *Eur. J. Neurosci.* 24, 37-44.
- Higashijima, S., Shishido, E., Matsuzaki, M. and Saigo, K. (1996). eagle, a member of the steroid receptor gene superfamily, is expressed in a subset of neuroblasts and regulates the fate of their putative progeny in the Drosophila CNS. *Development* 122, 527-536.
- Ito, K., Urban, J. and Technau, G. (1995). Distribution, classification, and development of Drosophila glial cells in the late embryonic and early larval ventral nerve cord. *Dev. Genes Evol.* 204, 284-307.
- 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.

- Landgraf, M. and Thor, S. (2006). Development of Drosophila motoneurons: specification and morphology. Semin. Cell Dev. Biol. 17, 3-11.
- Landgraf, M., Bossing, T., Technau, G. M. and Bate, M. (1997). The origin, location, and projections of the embryonic abdominal motorneurons of Drosophila. *J. Neurosci.* 17, 9642-9655.
- 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.
- Lein, E. S., Hawrylycz, M. J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A. F., Boguski, M. S., Brockway, K. S., Byrnes, E. J. et al. (2007). Genomewide atlas of gene expression in the adult mouse brain. *Nature* 445, 168-176.
- Lundgren, S., Callahan, C., Thor, S. and Thomas, J. (1995). Control of neuronal pathway selection by the Drosophila LIM homeodomain gene apterous. *Development.* **121**, 1769-1773.
- Manning, L., Heckscher, E. S., Purice, M. D., Roberts, J., Bennett, A. L., Kroll, J. R., Pollard, J. L., Strader, M. E., Lupton, J. R., Dyukareva, A. V. et al. (2012). A resource for manipulating gene expression and analyzing cis-regulatory modules in the Drosophila CNS. *Cell Rep.* 2, 1002-1013.
- Mccabe, B. D., Marqués, G., Haghighi, A. P., Fetter, R. D., Crotty, M. L., Haerry, T. E., Goodman, C. S. and O'Connor, M. B. (2003). The BMP homolog Gbb provides a retrograde signal that regulates synaptic growth at the Drosophila neuromuscular junction. *Neuron* 39, 241-254.
- Miguel-Aliaga, I. (2004). Independent roles of the dachshund and eyes absent genes in BMP signaling, axon pathfinding and neuronal specification. *Development* 131, 5837-5848.
- Miguel-Aliaga, I., Thor, S. and Gould, A. (2008). Postmitotic specification of Drosophila insulinergic neurons from pioneer neurons. *PLoS Biol.* 6, e58.
- Nechipurenko, I. V. and Broihier, H. T. (2012). FoxO limits microtubule stability and is itself negatively regulated by microtubule disruption. J. Cell Biol. **196**, 345-362.
- 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.
- O'Connor, M. B. (2005). Shaping BMP morphogen gradients in the Drosophila embryo and pupal wing. *Development* **133**, 183-193.
- Park, D., Han, M., Kim, Y.-C., Han, K.-A. and Taghert, P. H. (2004). Ap-let neurons–a peptidergic circuit potentially controlling ecdysial behavior in Drosophila. *Dev. Biol.* 269, 95-108.
- Park, D., Veenstra, J. A., Park, J. H. and Taghert, P. H. (2008). Mapping peptidergic cells in Drosophila: where DIMM fits in. *PLoS ONE* **3**, e1896.
- Patel, N. H., Martin-Blanco, E., Coleman, K. G., Poole, S. J., Ellis, M. C., Kornberg, T. B. and Goodman, C. S. (1989). Expression of engrailed proteins in arthropods, annelids, and chordates. *Cell* 58, 955-968.
- Peng, H., Long, F. and Myers, E. W. (2009). VANO: a volume-object image annotation system. *Bioinformatics* 25, 695-697.

- 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.
- Pereanu, W., Spindler, S., Im, E., Buu, N. and Hartenstein, V. (2007). The emergence of patterned movement during late embryogenesis of Drosophila. *Devel. Neurobio.* 67, 1669-1685.
- Pfeiffer, B. D., Jenett, A., Hammonds, A. S., Ngo, T.-T. B., Misra, S., Murphy, C., Scully, A., Carlson, J. W., Wan, K. H., Laverty, T. R. et al. (2008). Tools for neuroanatomy and neurogenetics in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9715-9720.
- Pfeiffer, B. D., Ngo, T.-T. B., 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.
- Rickert, C., Kunz, T., Harris, K.-L., Whitington, P. M. and Technau, G. M. (2011). Morphological characterization of the entire interneuron population reveals principles of neuromere organization in the ventral nerve cord of Drosophila. *J. Neurosci.* **31**, 15870-15883.
- Robinow, S. and White, K. (1988). The locus elav of Drosophila melanogaster is expressed in neurons at all developmental stages. *Dev. Biol.* **126**, 294-303.
- Sato, M., Kojima, T., Michiue, T. and Saigo, K. (1999). Bar homeobox genes are latitudinal prepattern genes in the developing Drosophila notum whose expression is regulated by the concerted functions of decapentaplegic and wingless. *Development* **126**, 1457-1466.
- Schmid, A., Chiba, A. and Doe, C. (1999). Clonal analysis of Drosophila embryonic neuroblasts: neural cell types, axon projections and muscle targets. *Development*.
- Thomas, J. B., Bastiani, M. J., Bate, M. and Goodman, C. S. (1984). From grasshopper to Drosophila: a common plan for neuronal development. *Nature* 310, 203-207.
- Tomancak, P., Berman, B. P., Beaton, A., Weiszmann, R., Kwan, E., Hartenstein, V., Celniker, S. E. and Rubin, G. M. (2007). Global analysis of patterns of gene expression during Drosophila embryogenesis. *Genome Biol.* 8, R145.
- 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.
- Wheeler, S. R., Kearney, J. B., Guardiola, A. R. and Crews, S. T. (2006). Singlecell 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. (2009). MidExDB: a database of Drosophila CNS midline cell gene expression. *BMC Dev. Biol.* 9, 56.
- Xiong, W. C., Okano, H., Patel, N. H., Blendy, J. and Montell, C. (1994). repo encodes a glial-specific homeo domain protein required in the Drosophila nervous system. *Genes Dev.* 8, 981-994.

Supplemental Table S1. Interneuron diversity

Supplemental Table S2. Midline diversity

Supplemental Table S3. Motor neuron diversity

Supplemental Figure 1. The eNeuro atlas provides tools to probe the function of single interneurons

(A-B) The eNeuro atlas can be used to characterize interneurons. Gal4 lines driving membranebound GFP (green) co-stained with FasII (magenta). Anterior up. (A-A') R48G05-Gal4 reveals projection patterns of dMP2 (white circles) along the DM fascicle (white arrowhead), MP1 (dashed circles) along the CI fascicle (arrow), and the previously uncharacterized IN147 (yellow arrowhead) along the lateral side of the neuropil (hollow arrow). (B-B') R21B12-Gal4 shows dMP2 and MP1. Symbols as in A. Maximum intensity projections through volume of neuropil. Scale bar 10 μ m.

(C) The eNeuro atlas can be used to design intersectional tools for obtaining Gal4 expression in single interneurons. Anterior up. (Left) R48G05-Gal4 marks dMP2, MP1 and IN147; (Middle) R21B12-Gal4 marks dMP2 and MP1. (Right) Generating R21B12-Gal80 and crossing to R21B12-Gal4 should result in Gal4 expression in a single interneuron, IN147.

Supplemental Figure 2. eNeuro atlas provides functional access to individual motor neurons

(A-B) R23G10-Gal4 driving UAS-myristolated-GFP (green) reveals the MDM2 motor neuron contralateral projections out the ISN (arrowhead), and the IN014 interneuron (dashed circle) anterograde projection along an intermediate fascicle (asterisk). FasII+ (magenta). Anterior up. Scale bar, 5 µm.

Supplemental Figure 3. Hb9+ motor neurons do not project out SNa nerve root

Hb9-Gal4 crossed to UAS-myristolated-GFP (green) counter stained with FasII+ (magenta). SNa (arrowhead), SNa nerve root; A, abdominal segment; anterior to the top; midline to the left; Scale bar 10 microns.

Supplemental Data File S1. Eve.apo

Folder containing the raw Eve image to which all other images are registered: as a TIF stack,

Quicktime movie, and EVE.apo file. The movie shows a rotation of a 3D view of the Eve+ neurons in T2-A1. The EVE.apo file can be opened in Vaa3D or using any text editor.

Supplemental Data File S2. Cell statistics data for the Eve+ neurons

Columns are for cell name (e.g. aCC/eve A1R where aCC=cell name, eve=marker, A1R=hemisegment), mean_x, mean_y, mean_z position (units are pixels; average nuclear diameter is ~10 pixels), mean_vol (nuclear volume), std_x, std_y, std_z (standard deviation along the x, y, or z axis), cellstd_vol (standard deviation of the nuclear volume), observed, expected, ratio (expected = number of embryos imaged; observed = actual number of times the cell was annotated; ratio = observed/expected).

Supplemental Data File S3. TF single marker atlases

Folder containing single marker atlases (.apo files) for individual TF markers (pMAD, Dimm, Repo, Bar, Hb9, Ap, Eve, En, [Sim—generated as a test of the software pipeline prior to publication]) plus the cell statistics.

Supplemental Data File S4. TF atlas

TF atlas apo file, plus the cell statistics, and gene expression matrix files.

Supplemental Data File S5. Validation of co-expression in TF and eNeuro atlases

A folder containing (1) wet lab stains or reference for in vivo co-expression, (2) list of references.

Supplemental Data File S6. Gal4 single marker atlases

Folder containing single apo atlas files for all 75 Gal4 lines plus the cell statistics file.

Supplemental Data File S7. eNeuro atlas

Folder containing the comprehensive atlas apo file showing all eight TF markers and 75 Gal4 markers, plus the cell statistics file and marker expression matrix.

Supplemental Data File S8. Cell type-specific atlases

Folder containing atlas files for all motor neurons (defined as pMad+), all neurosecretory cells (defined as Dimm+) and all interneurons (defined as pMad-, Dimm-, and Repo-negative), a subset of glial cells (defined as Repo+ with reproducible cell body positions), and midline cells. Note that eNeuro.apo and MidlineCells.apo depict additional AMG glia since their position can

vary due to migration.

Supplemental Data File S9. Users Manual

Instructions and additional files needed to generate, or manipulate atlases.

Supplemental Data File S10. Look up table for merging cells into the eNeuro atlas

The text file used to generate the eNeuro atlas from all 83 single marker atlases. A sample line reads: "MAC1 A1L: F/66C01 A1L,I/14E12 A1L,C/14H08 A1L,MAC/pMAD A1L", where "MAC1 A1L" is the name of new, merged cell that appears in the eNeuro atlas, and where "F/66C01 A1L,I/14E12 A1L,C/14H08 A1L,MAC/pMAD" are the name of cells in the 66C01.apo, 14E12.apo, 14H08.apo and pMAD.apo single marker atlases, respectively. Raw data from all cells in single marker atlases are averaged to make up the merged MAC1 A1L cell.



Heckscher et al., Supp. Fig.1



Heckscher et al., Supp. Fig. 2

Hb9-Gal4 axon projections



Heckscher et al., Supp. Fig. 3

Download Data Files

Supp Maths files 1-10

Molecular Subtype	T2	Т3	A1
R10D02			х
R10D07, R11F04, R15F01	х	Х	х
R11F02	Х	х	х
R11F04	Х	х	х
R12A10	х		х
R12E03, R46A06	х	х	х
R12E03	Х	х	
R12E07		х	х
R13D08, R14E12, R15D07	Х	х	х
R13D08, R15D07, R55E04	Х	х	
R13D08, R55E04			х
R13D08		х	х
R13E06, R19H09	Х	х	х
R13E06	х	х	
R14E12, R14H08, R15E11, R23G10	Х	х	х
R14E12, R14H08, R15E11	Х	х	х
R14E12, R15E11, R45D07, R46A06, R65B10, R71D07	х	х	х
R14E12, R15E11, R45D07, R46A06, R65B10	х	х	х
R14E12, R15E11, R52B03, R65D03	Х	х	
R14E12, R15E11, R52B03	х	х	х
R14E12, R15E11	Х	х	х
R14E12	Х	х	х
R15D07	х	х	х
R15E11, R65D03	х	х	х
R15E11	х	х	
R17C08, R19A07, R35D09, R35D11		х	
R17C08	Х	х	х
R17F04, R35D11, R65H11	Х	х	х
R18A04, R30A02, R35D09, R35D11, R64A12, R65D03, R66C01, R71D07	Х		х
R18A04, R30A02, R35D09, R35D11, R64A12, R65D03, R66C01		х	
R19A07, R29G10	Х	х	
R19A07, R35D09, R35D11	Х		
R19A07	Х	х	х
R19H09	Х	х	х
R20A12, R38E07			х
R20A12		х	х
R21A04			х
R21C06	х	х	х

Supplemental Table 1. There are 112 molecular subtypes of interneurons in the eNeuro atlas

R23A02	Х		
R23G10			Х
R25C03	х	х	х
R25D09			Х
R25G03, R33E02		Х	
R25G03	х	Х	х
R25H11		Х	Х
R26B03	х		
R26H02	х		Х
R27A09, R74B12	Х	Х	Х
R27A09	х	Х	
R29A12	Х	Х	Х
R29G10	Х	Х	
R30A02		Х	
R30E03	х		
R31G06			Х
R33E02	х	Х	Х
R33G07, R64B06, R65B10	Х		Х
R33G07	Х	Х	Х
R35D09	Х	Х	Х
R35D11	Х	х	Х
R37G07, R75G10	Х	х	Х
R38C08	Х	Х	Х
R38E07, R74B12	Х	х	Х
R38E07		Х	Х
R45D07	Х	Х	Х
R46A06	Х	Х	Х
R47H04, R54E08	Х	Х	Х
R47H04	Х	х	Х
R48G05	Х		
R50B02			Х
R53B03	Х		
R54B03, R55E04	х	Х	х
R54B03	Х	Х	Х
R54E06	Х		Х
R54E08	х	Х	
R55C06		Х	
R55E04	Х		х
R59D01	Х	Х	Х
R64A12	х	х	х
R64B05	Х	Х	

R64B06	х	х	х
R65B10	х	х	х
R65D03	х	х	х
R65H11	х	х	х
R66C01	Х	х	
R71D07		х	х
R71F10	Х	х	
R74B12	Х	х	
R75G10		х	х
9-58, R33G07, R64B06, R65B10		х	
9-58,			х
AP, R25D09	х	х	
AP	х	х	х
EN, R12E03, R26B03, R34D08, R65H11, R71D07	х	х	
EN, R12E03, R26B03, R65H11, R71D07			х
EN, R64B05		х	х
EN, R65H11	х	х	х
EN, R74B12	х	х	х
EN, HB9	х	х	х
EN	х	х	х
EVE, R11F02			х
EVE, R14E12, R15E11, R30A02, R35D09, R35D11, R65D03			х
EVE, R14E12, R15E11, R30A02, R35D09, R35D11	х	х	
EVE	х	х	х
HB9, R11F02	х	х	х
HB9, R12B12, R17F04, R21B12, R23A02, R26B03, R48G05, R53B03, R53E03			х
HB9, R12B12, R17F04, R21B12, R23A02, R26B03, R48G05, R53E03	х	х	
HB9, R21A04, R30E03, R35D09, R35D11, R38C08, R52B03, R53B03, R53F07, R59D01, R64B05, R66C01, R71D07		х	х
HB9, R21A04, R30E03, R35D09, R35D11, R38C08, R52B03, R53B03, R53F07, R59D01, R64B05,	х	х	
R66C01 HB9, R25C03, R31G06	x	x	
HB9, R31G06	x	x	
HB9, R33E02			x
НВ9	x	x	x
sum	83	86	80
	1	1	I

TF	Gal4 line	Neurons and Glia	Color in Fig. 7
HB9	R12B12 R17F04 R21B12 R23A02 R26B03 R48G05	MP1 L	Aqua
	R53B03 R53E03	MP1 R	
EN	R65H11	iVUM4	Strawberry
	None	iVUM5	
	R12E03 R26B03 R34D08 R65H11 R71D07	iVUM6	-
pMad	R12E07 R17F04	mVUM4	Orchid
	R12E07	mVUM5	
	R12E07 R15E11	mVUM6	
None	R65H11	MNB	Moss
EN	R64B05	MNBp1 L	Flora
		MNBp1 R	
		MNBp2 L	
		MNBp2 R	
	None	MNBp3 L	
		MNBp3 R	
None	R38C08	MNBp4 L	
		MNBp4 R	
	R15E11 R65D03	H-cell	Banana
	R17F04 R35D11 R65H11	H-cell sib	Magnesium
	R66C01	AMG DA1	Salmon
	R17F04	AMG DA2	
	R17F04	AMG DM1	
	R66C01	AMG DM2	
	R17F04	AMG VM	
EN REPO	None	D-CG L	Mocha
		D-CG R	
		V-CG L	
		V-CG R	
None	R25H11	DM1	Cayenne
		DM2	

Supplemental Table 2. There are 20 molecularly distinct subtypes of midline cells in A1.

	TF	Gal4 line	Neurons	Color in
pMad	Eve	R15E11, R26B03, R27A09, R46A06, R64B06, R65D03, R71D07, R74B12	aCC	Ocean
		9-58, R19H09, R26B03, R27A09, R64B06, R74B12	RP2	Sea foam
		R55C06, R64B06, R74B12	U1	Midnight
		R34D08, R59D01, R64B06	U2	Aqua
		None	U3, U4, U5	Blueberry
	Hb9	R22C10, R25H11	RP1	Spring
		R25H11	HLCLD	Fern
		R27A09	HLCLS	Lime
		9-58	RP5	Spindrift
		None	HLC2, HLCID, HLCM2, HM, RP3, RP4	Moss
	Bar	R19A07	BVM1	Magenta
		R17C08, R19A07	BVI1	Mocha
		R12E03, R21C06, R25H11	BVI2	Cayenne
		None	BDL1, BDL2, BDL3	Maraschino
	None	R12E07, R15E11	MVUM3	Tin
		R12E07	MVUM2	Steel
		R12E07, R17F04	MVUM1	Nickel
		R14E12, R14H08, R66C01	MAC1	Cantaloupe
		R25H11	MLC27	Salmon
		R27A09	MDM	Banana
		R31G06	MLCP	lce
		R66C01	MMC4	Bubblegum
		None	MD, MHU, MLC21, MLC22, MLC23, MLC24, MLC29, MMC1, MMCA, MMCC, MMCE, MU3A, MVU2	White

Supplemental Table 3. There are 23 molecularly distinct subtypes of motor neurons in A1