Nuclear Transcriptomes of the Seven Neuronal Cell Types That Constitute the Drosophila Mushroom Bodies

Meng-Fu Maxwell Shih,*† Fred Pejman Davis,†‡ Gilbert Lee Henry,‡§ and Josh Dubnau*,†‡

*Department of Anesthesiology, Stony Brook School of Medicine, Stony Brook, NY; †Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA; National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD; ‡Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA; Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, and §Department of Anesthesiology, Stony Brook School of Medicine; Department of Neurobiology and Behavior, Stony Brook University, Stony Brook, NY

ORCID IDs: 0000-0002-2976-9054 (M.-F.M.S.); 0000-0001-8294-1610 (F.P.D.); 0000-0003-1590-323X (G.L.H.); 0000-0002-9285-7444 (J.D.)

ABSTRACT The insect mushroom body (MB) is a conserved brain structure that plays key roles in a diverse array of behaviors. The Drosophila melanogaster MB is the primary invertebrate model of neural circuits related to memory formation and storage, and its development, morphology, wiring, and function has been extensively studied. MBs consist of intrinsic Kenyon Cells that are divided into three major neuron classes (γ, α′/β′ and α/β) and 7 cell subtypes (γd, γm, α′/β′ap, α′/β′m, α/βp, α/βs and α/βc) based on their birth order, morphology, and connectivity. These subtypes play distinct roles in memory processing, however the underlying transcriptional differences are unknown. Here, we used RNA sequencing (RNA-seq) to profile the nuclear transcriptomes of each MB neuronal cell subtypes. We identified 350 MB class- or subtype-specific genes, including the widely used α/β class marker Fas2 and the α′/β′ class marker trio. Immunostaining corroborates the RNA-seq measurements at the protein level for several cases. Importantly, our data provide a full accounting of the neurotransmitter receptors, transporters, neurotransmitter biosynthetic enzymes, neuropeptide receptors expressed within each of these cell types. This high-quality, cell type-level transcriptome catalog for the Drosophila MB provides a valuable resource for the fly neuroscience community.

Drosophila melanogaster is a powerful model system for behavioral neuroscience. The fly model takes advantage of a relatively simple brain that expresses homologous suites of genes and orchestrates a conserved yet highly diverse and elaborate suit of behaviors. Behavioral genetics in Drosophila affords the means to identify individual genes that function within identified neuronal cell types, whose connectivity and functional roles in behavior can be elucidated. The ability to form memories of past experience and to orchestrate adaptive and plastic changes in behavioral responses is an example of a fundamental field of behavioral neuroscience where Drosophila neurogenetics has made major contributions (Heisenberg 2003; Davis 2005; Margulies et al. 2005; Keene and Waddell 2007). Memory research in flies has led to the identification of fundamental cellular mechanisms of memory such as cAMP signaling and CREB-mediated gene transcription (Yin and Tully 1996; Heisenberg 2003; Davis 2005; Margulies et al. 2005; Keene and Waddell 2007), and also has contributed to our understanding of how memories are processed in a complex neural circuit. A primary site of associative learning in insects is the mushroom body (MB) (Strausfeld et al. 1998; Heisenberg 2003; Davis 2005; Margulies et al. 2005; Keene and Waddell 2007; Menzel 2012; Farris 2013), a paired brain structure that in Drosophila is comprised of approximately 2000 intrinsic Kenyon Cells (KCs) per hemisphere. MBs in fruit flies are critical sites of olfactory, visual and gustatory learning (Heisenberg 2003; Davis 2005; Margulies et al. 2005; Keene and Waddell 2007; Vogt et al. 2014; Masek and

KEYWORDS

Drosophila mushroom body transcriptome RNA-seq
Keene 2016), and also play important roles in other behavioral contexts such as temperature preferences (Hong et al. 2008), sleep (Artiushin and Sehgal 2017) and responses to ethanol exposure (Kaun et al. 2011).

MB-dependent plasticity is one of the most intensely studied aspects of invertebrate neurobiology. The morphology and developmental lineage of the neurons that populate the MB in *Drosophila*, as well as the identity and morphology of most of their neuronal inputs and outputs, have been fully characterized (Ito et al. 1998; Jefferis et al. 2002; Aso et al. 2014a; 2014b).

Many functional manipulations of both neural activity and signaling pathways relevant to plasticity have been conducted within each of the identified neuronal cell types in this circuit (Connolly et al. 1996; Zars et al. 2000; Dubnau et al. 2001; McGuire et al. 2001; Isabel et al. 2004; Krashes et al. 2007; Blum et al. 2009; Trannoy et al. 2011; Qin et al. 2012; Huang et al. 2012; Cervantes-Sandoval et al. 2013; Perisse et al. 2013; Bouzaine et al. 2015). Functional imaging studies have established neuronal activity correlates in behaving animals (Davis 2011). Together, these studies support the conclusion that the neurons of the MB play unique roles in memory acquisition, storage and retrieval. Moreover, memory performance correlates in behaving animals (Davis 2011). Together, these

### MATERIALS AND METHODS

#### Fly stocks

Because quantitative traits such as gene expression profiles can be sensitive to genetic background, we created a newly isogenic subsline of *w^1118* (isoC1), which was itself derived from Canton-S wild type as an inbred line many years ago (Yin et al. 1994). To generate the isogenic strain, we used ten-generations of single male and female sibling crosses to generate 10 independent isogenic strains. MJ2 was selected based on its ability to form comparable olfactory short-term memory performance to the parental strain in the standard Pavlovian task (Supplemental Material, Figure S1). The nuclear envelope epitope tagged transgene P[5XUAS-unc84::2XGFP]attP40, the pJFRC28 strain P[10XUAS-IVS-GFP-p10]attP2 and each of the split-GAL4 inserts were backcrossed into this new MJ2 wild type strain for five generations to equilibrate each to this isogenic background. For each split-GAL4 combination, we separately backcrossed the GAL4 activating domain and DNA-binding domain components, and then combined the two hemidrivers as a split-GAL4 line in the MJ2 background thereafter using standard balancer chromosomes that had themselves been equilibrated to the MJ2 strain. The UAS-WM P[5XUAS-myr::GFP-VS-p2A-His2B: mCherry-HA] reporter strain was generated using standard approaches (Chang et al. 2018). Flies were cultured on standard cornmeal food using the standard cornmeal recipe from Bloomington stock center.

Food was supplemented with antibiotics.

For imaging to characterize expression with each split-GAL4 strain that had been reconstituted in the MJ2 background, we used 2–5 day old male flies. For RNA-seq sample preparation, each split-GAL4 line was crossed to the P[5XUAS-unc84::2XGFP]attP40. 2–5 day old adult progeny for each genotype were collected and frozen in liquid nitrogen between 10 AM and 7 PM.

#### TAPIN purification of nuclei

Fly heads from a mixed population of male and female flies were first isolated with a customized sieve. 400 frozen heads were added to 6 mL of 20 mM sodium acetate pH 8.5, 2.5 mM MgCl₂, 250 mM sucrose, 0.5% NP-40, 0.6 mM spermidine, 0.2 mM spermine, 1 mM DTT, 1× complete protease inhibitor (Sigma: 5056489001), 0.5 mg/mL torula RNA (ThermoFisher: AM7118), 0.6 mg/mL carboxyl coated Dynabeads (ThermoFisher: 14306D) and 1.6 μg anti-GFP antibody (ThermoFisher: G10362). Homogenization was carried out on ice by 50 tractions in a Dounce homogenizer using the tight pestle followed by filtration over either a 10 or 20 μm cup filter (Partec: 040042315, 040042315).

Released chromatin and broken nuclei were adsorbed to carboxyl coated magnetic beads for 30 min at 4°C with constant rotation. Unbound antibody was removed by incubating the sample on ice for 20 min with 100 μL of washed UNosphere SUPra resin (Bio-Rad: 1560218). After the resin was removed on a 10 μm cup filter and the carboxyl beads on a magnet stand, the nuclei-containing supernatant was mixed with an equal volume of 500 mM sodium acetate pH 8.5, 250 mM sucrose, 6 mM EGTA, 6 mM EDTA, 0.6 mM spermidine, 0.2 mM spermine, 1 mM DTT, 1× complete protease inhibitor, 0.25 mg/mL torula RNA and 30 μL Protein A Dynabeads (ThermoFisher: 10002D). A 2-hour incubation on ice with occasional agitation was used to recover tagged nuclei. Bead-bound nuclei were then recovered on a magnet stand and washed twice with 250 mM sodium acetate pH 8.5, 250 mM sucrose.
and 0.1% NP-40. Nuclei were then released at 37°C for 1 hr by incubation in 50 μL of 10 mM Tris pH 7.5, 2.5 mM MgCl₂, 0.5 mM CaCl₂, 250 mM sucrose, 0.1% NP-40, 1 mg/mL torula RNA, 40 units RNAse A (Promega: N2515), 2 units DNaseI (NEB: M0303L), 320 units IdeZ protease (NEB: P0770S). The sample was diluted to 100 μL with 10 mM Tris pH 7.5, 2.5 mM MgCl₂, 0.5 mM CaCl₂, 250 mM sucrose and 0.1% NP-40, EGTa was added to 1 mM and the suspension was rapidly triturated 100 times. After returning the sample to a magnet stand, 90 μL of buffer containing released nuclei was removed and added to 1.5 μL of Protein G Dynabeads (ThermoFisher: 10004D) that were previously resuspended in 10 μL of 10 mM Tris pH 7.5, 2.5 mM MgCl₂, 0.5 mM CaCl₂, 250 mM sucrose and 0.1% NP-40. The second binding reaction was run for 1–3 hr on ice with occasional agitation, followed by 2x 250 μL washes in 10 mM Tris pH 7.5, 2.5 mM MgCl₂, 0.5 mM CaCl₂, 250 mM sucrose and 0.1% NP-40. Prior to the last wash a 5 μL aliquot was removed for quantitation and the remainder of the sample was solubilized in Arcturus Picopure RNA extraction buffer (ThermoFisher: KIT0204).

**RNA-seq library construction**

Nuclear RNA was DNaseI treated and purified using the Arcturus PicoPure system exactly as instructed by the supplier. Purified RNA was mixed with a 1:100,000 dilution of ERCC standard RNA (ThermoFisher: 4456740) and amplified using the Nugen Ovation v2 system (Nugen: 7102-32). cDNA was then blunt ligated, barcoded and linkers (Nugen: 0319-32, 0320-32) and sequenced in paired-end mode on an Illumina HiSeq2500 to 125 nt read lengths.

**RNA-seq data analysis**

We trimmed RNA-seq reads (5nt from the 5’ end of the forward read, using seqtk option "trim -b 5"; https://github.com/lh3/seqtk) to remove non-transcript sequences introduced by the NuGen Ovation kit and then pseudo-aligned these to the Dro sophila transcriptome (ENSEMBL release 91, BDGP6) using kallisto (Bray et al. 2016) to estimate transcript abundances. We also included sequences for the synthetic ERCC spike-in species and TAPIN reporter in the transcriptome index. After pseudo-alignment, we removed ERCC, TAPIN reporter, and ribosomal RNA entries and renormalized the transcript abundance matrix to units of transcripts per million (TPM). To visualize TAPIN-seq signal across the genome, we also aligned trimmed reads to the whole genome (BDGP6, dmb) using STAR (Dobin et al. 2013), created bigWig genome tracks (deetools; Ramirez et al. 2016), and visualized them in the IGV genome browser (Robinson et al. 2011).

To identify class- and subtype-enriched genes, we performed differential expression analysis using the estimated counts from kallisto as input to limma (Ritchie et al. 2015), voom (Law et al. 2014), and quantile normalizing the expression levels to account for differences in the number of genes detected in each sample (Table S1). We used criteria of at least 10 TPM abundance in one sample, at least twofold difference in expression, and 5% false discovery rate to identify differentially expressed genes.

**Immunohistochemistry**

Immunohistochemistry was performed essentially as in a previous report (Wu et al. 2013). Brains were dissected in isotonic PBS and immediately transferred to 4% paraformaldehyde in PBS for a 30-min fixation at room temperature. Fixed brain samples, were rinsed with isotonic PBS and incubated in PBS containing 2% Triton X-100, 10% normal goat serum (NGS; Penetration & Blocking Buffer) while being subjected to a degassing procedure (Chiang et al. 2011). Brain samples were agitated in the same buffer at 4°C overnight. Brains were then transferred to primary antibodies diluted with PBS containing 0.25% Triton X-100, 1% NGS (Dilution Buffer) and agitated at 4°C for 1–3 day. After primary antibody incubation, the brain samples were washed in PBS containing 1% Triton X-100, 3% NaCl (Washing Buffer) three times before they were moved to the 1:250 diluted secondary antibodies for one day agitation at 4°C. For the biotin amplification staining (Figure 4, S3 & S4A), samples were washed three times and agitated in the 1:500 diluted Alexa Fluor 635 streptavidin (Thermo Fisher Scientific, USA: S-32364) at 4°C for 1 day. Finally, the immunolabeled brain samples were washed three times, cleared and mounted in a drop of FocusClear (CellExplorer Labs, Taiwan: FC-101) between two coverslips separated by a spacer ring of ~200 μm thickness, so the brain samples were not flattened. The Penetration & Blocking Buffer and Dilution Buffer contain additional 0.02% Sodium Azide as a preservative. For GALa line characterization, 1:100 dilution of mouse anti-digl (4F3, deposited to the DSHB, USA by Goodman, C.) plus 1:250 dilution of rabbit anti-GFP (Thermo Fisher Scientific, USA: A-6455) primary antibody and 1:250 dilution of secondary antibody of Alexa Fluor 633-conjugated goat anti-mouse (Thermo Fisher Scientific, USA: A-21052) and Alexa Fluor 488-conjugated F(ab’2)2 goat-anti-rabbit (Thermo Fisher Scientific, USA: A-11070) were used. For the MB marker gene confirmation, a 1:4000 dilution of rabbit anti-sNPFP (Johard et al. 2008) or 1:20 dilution of mouse anti-Fas2 (1D4, deposited to the DSHB, USA by Goodman, C.) or 1:20 dilution of mouse anti-trio (9.4A, deposited to the DSHB, USA by Hama, C.) primary antibody, 1:250 dilution of secondary antibody of biotin-conjugated goat anti-rabbit (Thermo Fisher Scientific, USA: 65-6140) or biotin-conjugated goat anti-mouse (Thermo Fisher Scientific, USA: D-20691) were used. For the GABAergic identification staining, a 1:250 dilution of mouse anti-GFP (MilliporeSigma, USA: 11814460001) together with 1:250 dilution of rabbit anti-GABA (MilliporeSigma, USA: A2052) or 1:500 dilution of rabbit anti-Gad1 (Featherstone et al. 2000) or 1:400 dilution of rabbit anti-VGAT (Fet et al. 2010) primary antibodies, and 1:250 dilution of secondary antibody of Alexa Fluor 488-conjugated goat anti-mouse (Thermo Fisher Scientific, USA: A-11029) together with biotin-conjugated goat anti-rabbit or Alexa Fluor 647-conjugated goat anti-rabbit (Thermo Fisher Scientific, USA: A-21244) were used.

**Confocal imaging and post-imaging processing**

Brain samples were imaged under a Zeiss LSM 800 confocal microscope with a 40X C-Apochromat water-immersion objective lens (N.A. value 1.2). The settings for scanning were manually adjusted. To overcome the limited field of view when imaging the GALa expression patterns, we scanned each brain in two parallel stacks of confocal images with some overlap between the two brain hemispheres, with a voxel size of 0.31 X 0.31 X 1.25 μm. We then stitched the two image stacks into a single data set with ‘Pairwise stitching’ function in Fiji (Preibisch et al. 2009; Schindelin et al. 2012), segmented the brain region based on the dlg1 staining channels with 3D Slicer (https://www.slicer.org Kikinis et al. 2013), and made a ‘Z projection’ with Fiji (Schindelin et al. 2012). The MB subtype models were constructed from the GFP channel of the confocal images used for projections, by using the 3D Slicer to segment, show 3D, and conduct smoothing.

**Cell counting**

The confocal images for cell counting were acquired with a voxel size of 0.16 X 0.16 X 1.00 μm. The GFP channel was first used to identify KCs, and then the ‘Cell Counter’ plugin in Fiji was used to count all detectable nuclei in the mCherry channel (Preibisch et al. 2009). For each line, we counted three hemispheres from three different animals.
**RESULTS**

**TAPIN-seq profiling of MB neuronal cell subtypes**

To label MB subtypes, we used seven split-GAL4 lines: MB607B (yd), MB131B (α+/β+), MB370B (α/B+/β+), MB418B (α/β+/m), MB371B (α/βp), MB185B (α/βs) and MB594B (α/βc) (Aso et al. 2014a). We first backcrossed all the split GAL4 hemi-drivers and the nuclear-tag reporter (Henry et al. 2012) into MJ2, an isogenic Canton-S derivative (Methods). Because of the change in genetic background, we re-characterized the expression pattern of each split-GAL4 combination to confirm that the expected cell subtype-specific pattern had not been altered. We used a novel membrane-GFP-P2A-nuclear-mCherry dual label reporter (Watermelon or WM for short; see Methods) (Chang et al. 2018), which expresses membrane-tethered GFP and nuclear mCherry from a single transcript via a viral ribosome skip peptide (P2A; Daniels et al. 2014). The GFP label revealed neuronal morphology, thereby confirming MB subtype specificity, and the nuclear-mCherry marked each nucleus, which permitted an accurate cell count. Using WM labeling, we found that each split-GAL4 combination yielded limited expression in only small numbers of neurons outside the MB, but exhibited strong expression in the annotated MB cell subtype (Figures 1A – 1G, 1A′ – 1G′ & Files S1 – S7) (Aso et al. 2014a). Using high-resolution imaging (Figures 1A ′ – 1G′ & Files S8 – S14), we were able to count the total number of MB KCs of each subtype that is labeled by a given split-GAL4 combination (Table 1). The numbers of labeled cells for each MB KC subtype are in general agreement with the previous report (Aso et al. 2014a) and it appears that most if not all of the neurons of each subtype are labeled by these Split-GAL4 combinations. In fact the total number of KCs labeled by each of the MB KC neuronal subtypes are labeled with these Split-GAL4 combinations.

To profile the nuclear transcriptome of each MB subtype, we used TAPIN-seq (Davis et al. 2018), a modification of INTACT (Henry et al. 2012) that yields improved selectivity by use of a two-step purification method. The 7 characterized split-GAL4 combinations were used to express the nuclear membrane protein UNC84 fused with 2 copies of GFP (Figure 2A). For each split-GAL4 combination, tagged nuclei of a given MB KC subtype were purified from ~400 fly heads. Nuclear RNA was then extracted and used to generate RNA-seq libraries (Figure S2B). We generated libraries from two independent biological replicates for each MB KC subtype, and paired-end sequenced them (Table S1). We estimated transcript abundances in each library using kallisto (Bray et al. 2016; see Methods). The sequencing reads and estimated abundances are available in NCBI GEO (GSE119629).
this protein distribution, we identify signal in the (Figure 4A). Fas2 has been reported to exhibit strong immunoreactive a detected in the cell bodies of KCs (Figure S3), consistent with a previous report and is consistent with the TAPIN-seq results. We observe no reactivity is localized in the MB (20 TPM in m + d, a’/b’ m (Ets96B, otp), a’/b’ (br, brus3, dl, disco-r, ind, rm), and a’/b’c (Tet) subtypes. Although the abundance and significance of DNA methylation in Drosophila is unclear, enrichment of the Tet DNA methyltransferase is intriguing given the role of DNA methylation in memory formation in other insects (Biergans et al. 2012) and mammals (Day and Sweat 2010). The subtype-enriched expression could reflect a remnant of a functional expression pattern from an ancestral species with DNA methylation.

Several genes encoding cell-surface molecules were also differentially expressed across the subtypes. These genes included members of gene families previously implicated in specifying synaptic connectivity, including the defective proboscis extension response (Dpr) as well as the Dpr-interacting protein (DIP). Interactions between proteins from these families have been previously documented and shown to underlie synaptic connectivity in circuits including the visual system (Özkan et al. 2013; Carrillo et al. 2015). The expression patterns we observed for several cell-surface molecules suggest they might be involved in specifying class- and subtype-specific connectivity.

**Neurotransmitter output**

To explore the functional utility of our TAPIN-seq measurements, we next focused on genes related to the input and output properties of the MB cell subtypes. Specifically, we examined genes encoding neurotransmitter biosynthetic enzymes, neurotransmitter transporters (Figure 5A), neurotransmitter receptors (Figure 5B), neuropeptides (Figure 5C), neuropeptide and protein hormone receptors (Figure 5D), and gap junction components (Figure 5E). A recent report has established that the primary neurotransmitter in the MB is acetylcholine (Barnstedt et al. 2016). Because our method profiled each of the seven MB neuronal subtypes, we re-assessed this conclusion at a higher resolution. We confirmed that all seven MB cell subtypes express high levels of both choline acetyltransferase (ChAT; 75 – 198 TPM) and vesicular acetylcholine transporter (VACHT; 257 – 517 TPM).

We next asked whether one or more KC subtypes might co-release other small molecule neurotransmitters, but found no strong argument to support such a conclusion. In a few cases, we see moderate expression of biosynthetic enzymes for other neurotransmitters including GABA, serotonin and dopamine. But in each case, there are findings that undermine the conclusion that these neurotransmitters are consistently produced/released by any of the 7 KC subtype. For example, we do see strong expression in all MB cell subtypes of vesicular GABA transporter (VGLUT; 152 – 279 TPM), which is an essential transporter that is responsible for packaging the neurotransmitter GABA into synaptic vesicles (Fei et al. 2010). On the other hand, the gene for GABA biosynthetic enzyme, Gad1, is moderately expressed in only the a’/b’ (121 TPM) and a’/b’c subtypes of KCs (54 TPM) and at lower levels in the remaining cell subtypes (4 – 14 TPM). In principle, moderate Gad1 expression might be consistent with the hypothesis that GABA is released from a fraction of a’/b’p and/or a’/b’c subtype neurons. To examine this possibility, we conducted immunofluorescence experiments.

---

**Table 1 List of split-GAL4 lines labeling MB subtype (n = 3)**

<table>
<thead>
<tr>
<th>DRIVER</th>
<th>MB SUBTYPE</th>
<th>NUMBER OF CELLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MB607B</td>
<td>γcl</td>
<td>132.00 ± 6.03</td>
</tr>
<tr>
<td>MB131B</td>
<td>γm+d</td>
<td>574.33 ± 2.73</td>
</tr>
<tr>
<td>MB370B</td>
<td>a’/b’ ap+m</td>
<td>405.33 ± 18.80</td>
</tr>
<tr>
<td>MB418B</td>
<td>a’/b’ m</td>
<td>152.67 ± 3.71</td>
</tr>
<tr>
<td>MB371B</td>
<td>a’/b’p</td>
<td>77.33 ± 2.85</td>
</tr>
<tr>
<td>MB185B</td>
<td>a’/b’ c</td>
<td>294.33 ± 16.68</td>
</tr>
<tr>
<td>MB594B</td>
<td>a’/b’c</td>
<td>505.33 ± 9.67</td>
</tr>
</tbody>
</table>
with antibodies against GABA, Gad1, and VGAT but found no marked immunoreactivity in any \(\alpha/\beta p\) KC subtype cell bodies (Figure S4). This result suggests that the most likely explanation for our observed Gad1 pattern is that the split-GAL4 lines we used for \(\alpha/\beta p\) and \(\alpha/\beta c\) KC subtypes drive low levels of expression in some subset of GABAergic neurons outside the MB (Figure S4A).

A similar set of findings are apparent with Dopa decarboxylase (Ddc), a commonly used marker for dopaminergic or serotonergic neurons. Ddc catalyzes the decarboxylation of dopa to dopamine and 5-hydroxytryptophan to serotonin but not tyrosine to tyramine (Gramates et al. 2017). We see fairly strong Ddc expression in all MB cell types (55 – 574 TPM), especially the \(\gamma m+d\) (574 TPM) and \(\gamma d\) (339 TPM) KCs. However, we do not see expression of Tryptophan hydroxylase (Trh; 0 – 3 TPM), which provides the first and rate-limiting step in the synthesis of serotonin. Nor do we detect serotonin transporter (SerT; 0 – 2 TPM; Giang et al. 2011) in any of the 7 cell subtypes (Figure 5A). Pala (plic), which encodes tyrosine hydroxylase (TH) for dopamine synthesis, also is not highly expressed in any MB cell subtype (2 – 16 TPM; Figure 5A). Thus, we conclude that all MB KC subtypes likely release acetylcholine (Barnstedt et al. 2016; Crocker et al. 2016) and no other small molecule neurotransmitters.

**Neurotransmitter receptors**

We next examined expression profiles of small molecule neurotransmitter receptors. Dopamine has been established as a key input to MB for many behaviors including aversive and appetitive olfactory learning (Kim et al. 2007; Claridge-Chang et al. 2009; Qin et al. 2012; Liu et al. 2012; Burke et al. 2012), modulation of motivational state (Krashes et al. 2009), regulated forgetting (Berry et al. 2012; Shuai et al. 2015), sleep (Sitaraman et al. 2015b), courtship behaviors (Kuo et al. 2015; Lim et al. 2018), and temperature preference (Bang et al. 2011). A network of dopaminergic neurons innervates all MB cell subtypes (Aso et al. 2014a), and all four dopamine receptors, Dop1R1, Dop1R2, Dop2R and DopEcR, are expressed in all 7 MB KC subtypes (394 – 1152, 1152 – 2304, 2304 – 4608, 4608 – 9216, 9216 – 18432, 18432 – 36864, 36864 – 73728 TPM) (Figure 2C).
222 – 408, 444 – 1059, 3284 – 8335 TPM, respectively; Figure 5B) (Han et al. 1996; Kim et al. 2007; Qin et al. 2012; Ishimoto et al. 2013; cf. Draper et al. 2007). We confirmed that the Dopamine transporter (DAT) is preferentially expressed in both MB α’β’ap and α’β’m cell subtypes (397 and 262 TPM in α’β’ and α’β’m, respectively, vs. 0.2 – 4.3 TPM in other subtypes; Figure 5A) (Cros et al. 2018; Davie et al. 2018), suggesting that precise temporal control of dopaminergic signaling may be at play in α’β’ KCs. We also observe equivalently high levels of expression in all MB KC subtypes for all six GABA receptor genes. Among the three GABAA receptors, Resistance to dieldrin (Rdl) is strongly expressed (5,021 – 222 TPM), consistent with previous findings (Harrison et al. 1996; Liu et al. 2007), while Ligand-gated chloride channel homolog 3 (Lch3) is moderately expressed (70 – 153 TPM), and Drosophila Glycine receptor (Gri) is not detected (0.1 – 0.8 TPM). All the GABAB receptors are broadly expressed in MB except for GABA-B-3 (1 – 61 TPM; Figure 5B). These findings are consistent with a report that establishes the importance of GABAergic feedback from the anterior paired lateral neurons to MB KCs to support anesthesia-resistant memory formation (Lee et al. 2008; Sinakevitch et al. 2010). Consistent with this conclusion, we observe low VGluT expression in all α’β’ and α’β’m KCs to support anesthesia-resistant memory formation (Lee et al. 2011). A previous report observed 5-HT1A-GAL4 expression and 5-HT1B immunoreactivity in the γ KC class (Yuan et al. 2005). We did not detect nuclear 5-HT1B transcripts by TAPIN-seq (0.1 – 1.0 TPM), but we cannot rule out the possibility that higher levels of transcripts and protein are present in the cytoplasm, beyond detection in this nuclear transcriptome.

MB KCs receive nicotinic acetylcholine receptor (nAChR)-mediated synaptic transmission from antennal lobe projection neurons (Gu and O’Dowd 2006). We found the nAChR subunits nAChRalpha1, nAChRalpha4, nAChRalpha5, nAChRalpha6, nAChRalpha7, nAChRbeta1 and nAChRbeta2 are strongly expressed in all seven MB cell subtypes, but nAChRalpha3 and nAChRbeta3 transcripts are absent or undetectable (Figure 5B). Two muscarinic acetylcholine receptors (mAChRs), mAChR-A and mAChR-B, are also expressed in the MB with an enrichment in α’β’ lobe KC class (Figure 5B & Table S2).

We also examined the expression of the six known octopamine receptors (Gramates et al. 2017). We found that Oamb and CG18208 (recently characterized as Octα2R; Qi et al. 2017), which are α-adrenergic-like receptors, are class- and subtype-specific genes, respectively (Table S2). Oamb TAPIN-seq signal is strongly detected in α’β’ class KCs and far lower levels are seen in α’β’m class KCs (Figure 5B) (Crittenden et al. 1998; Kim et al. 2013). The three β-adrenergic-like receptors — Octβ1R, Octβ2R and Octβ3R — are all detected in the MB with variable levels across cell subtypes (cf. Wu et al. 2013 for Octβ2R immunolabeling).

Although transient glutamate immunoreactivity has been shown in α’β’ lobe KC class of young adult, VGluT expression has never been observed in the MB KCs (Daniels et al. 2008; Sinaikovitch et al. 2010). Consistent with this conclusion, we observe low VGluT transcript abundance (5 – 50 TPM; Figure 5A). We also confirmed that the NMDA receptors, both Nmdar1 and Nmdar2, are broadly expressed in the MB (Figure 5B) (Xia et al. 2005; Ueno et al. 2017); ionotropic receptors GluRIA and GluRIB are also expressed (Figure 5B). Flies also have a unique metabotropic glutamate receptor called nGluR, which has been

### Table 2 Number of genes enriched and depleted in individual MB classes and subtypes. Genes identified by limma/voom (q-value < 0.05, ifold change > 2x)

<table>
<thead>
<tr>
<th>GROUPTYPE</th>
<th>GROUP</th>
<th>ENRICHED</th>
<th>DEPLETED</th>
</tr>
</thead>
<tbody>
<tr>
<td>class</td>
<td>γ</td>
<td>47</td>
<td>22</td>
</tr>
<tr>
<td>class</td>
<td>α’β’</td>
<td>78</td>
<td>54</td>
</tr>
<tr>
<td>class</td>
<td>α/β</td>
<td>77</td>
<td>63</td>
</tr>
<tr>
<td>subtype</td>
<td>γd</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>subtype</td>
<td>γm+cl</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>subtype</td>
<td>α’β’ap+m</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>subtype</td>
<td>α’β’m</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>subtype</td>
<td>α’βp</td>
<td>21</td>
<td>11</td>
</tr>
<tr>
<td>subtype</td>
<td>α’βs</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>subtype</td>
<td>α’βc</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
previously observed by immunolabeling throughout the adult brain, but minimally in the MB lobes (Devaud et al. 2008). Here, with the cell type resolution of our dataset, we identified mGluR as an MB class-specific gene that is expressed in a subset of KCs—mGluR is depleted in α/β class KCs (40 TPM; Table S2) relative to both α'/β' and γ classes of KCs (504 and 612 TPM, respectively; Figure 5B).

Figure 4 (A) trio is depleted in α/β KCs. Whole mount anti-trio immunostaining confirmed strong signal in the MB α'/β' lobes, moderate signal in the γ lobe, and no signal in the α/β lobes. The cell bodies of MB α'/β' KC class also showed immunoreactivity. (B) Fas2 is depleted in MB α'/β' KC class. Whole mount anti-Fas2 immunostaining confirmed stronger signal in the MB α/β and γ'd lobes. (C) sNPF is depleted in MB α'/β' KC class. Whole mount anti-sNPF precursor immunostaining confirmed no detectable signal in the MB α'/β' lobes. Among the immunoreactive α/β and γ lobes, the α/β'p lobes showed the strongest signal. In each plot the bars represent the mean TPM, and the dots represent individual replicate values. Scale bars represent 20 μm. Expression patterns of the split-GAL4 lines were reported by P(10XUAS-IVS-GFP-p10)attP2.
Neuropeptides

In addition to small neurotransmitter systems, we also examined expression of neuropeptides and neuropeptide receptors. We consistently detect expression of a small group of well-characterized and putative neuropeptides. Unexpectedly, this includes amnesiac (*amn*), which is detected at fairly robust levels in all 7 MB KC subtypes.
(52 – 175 TPM) — although this small gene resides in the intron of another gene, *Herc*, which is highly expressed (218 – 239 TPM), thus complicating the accurate estimation of *ann* abundance. Previous work has established a requirement during memory formation for *ann* expression in a single pair of DPM neurons outside MB (Waddell et al. 2000). *sNPF* expression appears as a class-specific transcript, with high levels in α/β and γ KC classes and very little expression in the α’/β’ KC class (414, 725, 20 TPM, respectively; Figure 4C). This is consistent with previous reports (Johard et al. 2008) and we further confirm this expression pattern by immunostaining (Figure 4C). Another notable class-specific neuropeptide is *Drosophila Insulin-like peptide 1 (Ilp1; Liu et al. 2016b)*, expressed at high levels in α’/β’ class of KCs (284, 382 TPM in α’/β’, α’/β’m, respectively) and α/βc subtype (220 TPM) but not in most other MB KC subtypes (0.8 – 55 TPM; Figure 5C). As with neuropeptides, we also detect a panel of neuropeptide and protein hormone receptors, some of which are robustly expressed in all MB KC subtypes (e.g., *Ec dysone receptor*), some of which are class specific (e.g., *Dh44-R1* and *h ector*), and some of which are enriched or depleted in one or more subtypes of neurons (e.g., *AstC-R1* and *ETHR*; Figure 5D).

Finally, because of a report that gap junctions may form between MB KCs of different classes and play a role in visual learning (Liu et al. 2016a), we examined the expression levels of the eight gap junction genes in MB cell subtypes. *shakl* is strongly expressed (433 – 751 TPM), and *Inx3* moderately (29 – 107 TPM), in all MB cell subtypes (Figure 5E). Although *Inx5* and *Inx6* are reportedly required in the α/β and α’/β’ KC classes for visual learning and memory (Liu et al. 2016a), we do not detect either gene in any of the 7 MB KC subtypes (0 – 0.2 TPM, 0 – 0.5 TPM, respectively; Figure 5E). We cannot rule out the possibility that functionally relevant levels of expression are below our detection limit or that cytoplasmic RNA levels are higher.

**DISCUSSION**

Our results establish a high-quality, neuronal cell type-level transcriptome for *Drosophila* MB. We identified 350 differentially expressed genes, which includes most of the previously reported MB lobe (class specific) markers and many novel class-specific or cell subtype-specific profiles of expression. In addition to the subtype level resolution of our experimental design, the TAPIN approach that we used also offers several advantages and technical differences with these prior approaches. First, because TAPIN is compatible with flash frozen tissue as the input, the method introduces minimal disturbance to the endogenous transcriptome as compared to more lengthy procedures for purification of neurons for expression profiling. Second, it may be relevant that TAPIN explicitly profiles nuclear RNAs, likely enriching for actively transcribed nascent transcripts vs. abundant ones that are stably maintained in the cytoplasm. Thus, it would be attractive to apply this method to profile transcriptional response to behavioral perturbations.

Several previous studies have used genome-wide methods to profile expression in the *Drosophila* MB (Perrat et al. 2013; Crocker et al. 2016; Croset et al. 2018; Davie et al. 2018; Jones et al. 2018). Perrat et al. used a microarray-based approach to profile expression of each of the three major classes of MB KCs (purified by flow cytometry from dissociated brains) and compared these profiles with expression in the rest of the brain. They first focused on the expression of transposons (Perrat et al. 2013), and subsequently used the same transcriptome dataset to discover that MB KCs are cholinergic (Barnstedt et al. 2016) based on expression of biosynthetic enzymes. Crocker et al. used an RNA-seq-based approach to profile expression in relatively small pools of physically isolated α/β and γ class neurons to search for memory-related changes in gene expression (Crocker et al. 2016). Most recently, Croset et al. and Davie et al. used droplet-based single cell sequencing to profile the *Drosophila* brain, and by clustering the single cells they were able to identify the three MB classes, but not the further sub-division into neuronal subtypes (Croset et al. 2018; Davie et al. 2018).

Although we used a different profiling method and resolved transcriptions at the cell subtype rather than class level, our findings are broadly compatible with prior reports (Perrat et al. 2013; Barnstedt et al. 2016; Crocker et al. 2016; Croset et al. 2018; Davie et al. 2018). Our dataset reveals strong expression of both *ChAT* and *VAChT*, consistent with the conclusion that MBs are cholinergic (Barnstedt et al. 2016). Our findings further support the conclusion that all of the individual MB KC subtypes are cholinergic. The datasets also are consistent in the expression of known class-specific markers. One notable difference is that Crocker et al. reported high levels of expression of the 5-HT1B receptor in both α/β and γ classes of KCs (Crocker et al. 2016), and Davie et al. also observed 5-HT1B expression in α/β and γ KCs single cell clusters (Davie et al. 2018). In contrast, we see no evidence for expression of this receptor in our TAPIN-seq profiles (0.1 – 1 TPM). This difference could reflect methodology: Crocker and Davie both measured 5-HT1B receptor transcripts in the cytoplasm while we measured the levels that are actively transcribed or present in the nucleus. This technical difference could be especially relevant for neurotransmitter receptors, some of which can be translated locally at dendrites (Steward and Banker 1992).

Our dataset is the first to profile expression in this brain region at neuronal subtype resolution. This level of resolution is critical given the wealth of data on the functional differences of each MB KC subtype in *Drosophila* behaviors. Our dataset provides a full accounting within each of the MB KC subtypes of the profiles of expression of the cellular machinery to produce and receive neurotransmission, including small molecule transmitters and their receptors, neuropeptides and neuropeptide receptors and subunits of gap junctions. It is noteworthy that the TAPIN expression dataset supports the conclusions that all the adult MB KC subtypes are cholinergic, and that none of the subtypes express genes that would suggest the co-release of GABA, dopamine, glutamate, or serotonin. On the other hand, we detect expression of a spectrum of neuropeptides and their receptors. This observation is consistent with the hypothesis that MB KCs may co-release both acetylcholine and several neuropeptides (Takemura et al. 2017).

In addition to these findings with regards to the inputs and outputs, we identified 350 differentially expressed genes including many that distinguish MB KC classes or even individual cell subtypes. MB α/βp subtype showed 21 enriched genes and 11 depleted genes, contrasting with two other subtypes in the α/β class and two other classes. This uniqueness is supported by its unique odor responses (Perisse et al. 2013) and connectivity (Chen et al. 2012). Despite the limitation in the methodology that we profile the MB ym subtype using the spilt-GAL4 line MB131B that has minor expression in yd, and profile the α’/β’ap subtype using MB370B that has minor expression in α’/β’m (Figure 1 & Table 1), we still identified distinct sets of enriched/depleted genes (Figure 3 & Table S2), indicating the differences between two subtypes in MB γ or α’/β’ classes.

Our dataset provides a valuable resource for the fly neuroscience community to conduct functional studies. For example, our data provide a list of previously unknown class specific and sub-type specific transcripts, whose impact on the functional differences between these neurons are not known (Figure 3; Table S2). An arsenal of genetic tools to manipulate any gene’s function within each of these cell subtypes already exists (Caygill and Brand 2016; Kaya-Çopur and Schnorrer 2016). In addition to olfactory associative memory, MBs also play fundamental roles in other forms of memory including visual and...
gustatory (Vogt et al. 2014; Masek and Keene 2016), temperature preference (Hong et al. 2008), courtship behaviors (Kuo et al. 2015; Lim et al. 2018), stress response (Ries et al. 2017), food-seeking (Tao et al. 2018), sleep (Artiushin and Sehgal 2017) and responses to ethanol (Kaun et al. 2011). This dataset will facilitate the discovery of neural mechanisms for each of these conserved behaviors.

ACKNOWLEDGMENTS
We thank Y. Aso and G. Rubin (Janelia Research Campus) for the split-GAL4 lines and pFRC28 reporter, J. Veenstra (Université de Bordeaux 1, France) for the rabbit anti-sNPFP antibody, D. Krantz (University of California, Los Angeles) for the rabbit anti-VGAT antibody, F. Jackson (Tufts University) for the rabbit anti-Gad1 antibody and the Developmental Studies Hybridoma Bank (The University of Iowa), created by the NICHD, for the mouse monoclonal antibodies against dlg1, Fas2 and trio. We thank W.-W. Liao (Washington University in St. Louis) for technical assistance in the alignment files. We also are grateful to J. Azpurua, J. Beshel, Y.-H. Chang, R. Keegan, S. Krupp, L. Talbot for helpful discussions or comments on the manuscript. This work utilized the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov). This work was supported by US National Institutes on Deafness and Other Communication Disorders [5R01DC03071-06 to J.D.]; DART Neuroscience LLC [to J.D.]; Taiwan National Science Council [102-2917-I-564-004 to M.-F.M.S.]; and US National Institute of Arthritis and Musculoskeletal and Skin Diseases [Intramural Research Program to F.P.D.].

LITERATURE CITED
Chang, Y.-H., R. M. Keegan, L. Prazak, and J. Dubnau, 2018 Cellular labeling of endogenous virus replication (CLEVR) reveals de novo insertions of the gypsy endogenous retrovirus in cell culture and in both neurons and glial cells of aging fruit flies. bioRxiv. https://doi.org/10.1101/445221
M.-F. M. Shih et al.


Communicating editor: M. Arbeitman