Published in final edited form as:

Cell Rep. 2016 May 17; 15(7): 1580–1596. doi:10.1016/j.celrep.2016.04.046.

Cell Type-Specific Transcriptome Analysis in the *Drosophila* Mushroom Body Reveals Memory-Related Changes in Gene Expression

Amanda Crocker[1,5], **Xiao-Juan Guan**[1], **Coleen T. Murphy**[2,3,4], and **Mala Murthy**[1,2,*]

[1]Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544

[2]Department of Molecular Biology, Princeton University, Princeton, NJ 08544

[3]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544

[4]Paul F. Glenn Laboratories for Aging Research, Princeton University, Princeton, NJ 08544

Summary

Learning and memory formation in *Drosophila* rely on a network of neurons in the mushroom bodies (MB). While numerous studies have delineated roles for individual cell types within this network in aspects of learning or memory, whether or not these cells can also be distinguished by the genes they express remains unresolved. In addition, the changes in gene expression that

*correspondence: mmurthy@princeton.edu.
[5]Present Address: Neuroscience Program, Middlebury College, Middlebury, VT 05753

Author Contributions:

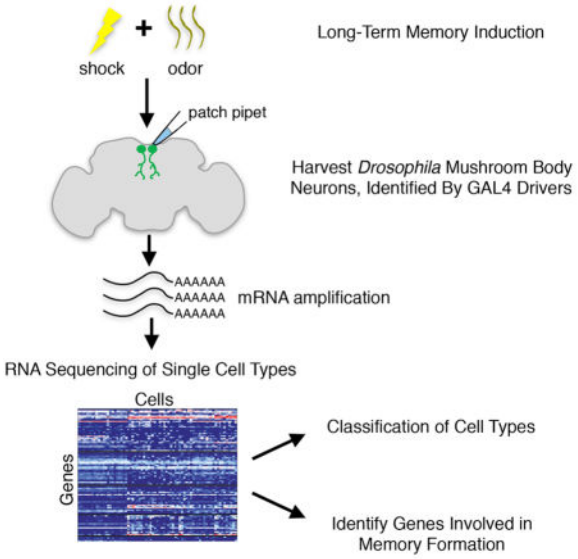
AC and MM designed the study, AC collected data with assistance from XG, AC analyzed the data, CTM advised on experimental design and data analysis, AC and MM wrote the paper.

Accession Numbers:

GSE74989,GSM1939831,GSM1939832,GSM1939833,GSM1939834,GSM1939835,GSM1939836,GSM1939837,GSM1939838,GSM1939839,GSM1939840,GSM1939831,GSM1939832,GSM1939833,GSM1939834,GSM1939835,GSM1939836,GSM1939837,GSM1939838,GSM1939839,GSM1939840,GSM1939841,GSM1939842,GSM1939843,GSM1939844,GSM1939845,GSM1939846,GSM1939847,GSM1939848,GSM1939849,GSM1939850,GSM1939851,GSM1939852,GSM1939853,GSM1939854,GSM1939855,GSM1939856,GSM1939857,GSM1939858,GSM1939859,GSM1939860,GSM1939861,GSM1939862,GSM1939863,GSM1939864,GSM1939865,GSM1939866,GSM1939867,GSM1939868,GSM1939869,GSM1939870,GSM1939871,GSM1939872,GSM1939873,GSM1939874,GSM1939875,GSM1939876,GSM1939877,GSM1939878,GSM1939879,GSM1939880,GSM1939881,GSM1939882,GSM1939883,GSM1939884,GSM1939885,GSM1939886,GSM1939887,GSM1939888,GSM1939889,GSM1939890,GSM1939891,GSM1939892,GSM1939893,GSM1939894,GSM1939895,GSM1939896,GSM1939897,GSM1939898,GSM1939899,GSM1939900,GSM1939901,GSM1939902,GSM1939903,GSM1939904,GSM1939905,GSM1939906,GSM1939907,GSM1939908,GSM1939909,GSM1939910,GSM1939911,GSM1939912,GSM1939913,GSM1939914,GSM1939915,GSM1939916,GSM1939917,GSM1939918,GSM1939919,GSM1939920,GSM1939921,GSM1939922,GSM1939923,GSM1939924,GSM1939925,GSM1939926,GSM1939927,GSM1939928,GSM1939929,GSM1939930,GSM1939931,GSM1939932,GSM1939933,GSM1939934,GSM1939935,GSM1939936,GSM1939937,GSM1939938,GSM1939939,GSM1939940,GSM1939941,GSM1939942,GSM1939943,GSM1939944,GSM1939945,GSM1939946,GSM1939947,GSM1939948,GSM1939949,GSM1939950,GSM1939951,GSM1939952,GSM1939953,GSM1939954,GSM1939955,GSM1939956,GSM1939957,GSM1939958,GSM1939959,GSM1939960,GSM1939961,GSM1939962,GSM1939963,GSM1939964,GSM1939965,GSM1939966,GSM1939967,GSM1939968,GSM1939969,GSM1939970,GSM1939971,GSM1939972,GSM1939973,GSM1939974,GSM1939975,GSM1939976,GSM1939977,GSM1939978,GSM1939979,GSM1939980,GSM1939981,GSM1939982,GSM1939983,GSM1939984,GSM1939985,GSM1939986,GSM1939987,GSM1939988,GSM1939989,GSM1939990,GSM1939991,GSM1939992,GSM1939993,GSM1939994,GSM1939995,GSM1939996,GSM1939997,GSM1939998,GSM1939999,GSM1940000,GSM1940001,GSM1940002,GSM1940003,GSM1940004,GSM1940005,GSM1940006,GSM1940007,GSM1940008,GSM1940009,GSM1940010,GSM1940011,GSM1940012,GSM1940013,GSM1940014,GSM1940015

accompany long-term memory formation within the MBs have not yet been studied by neuron type. Here we address both issues by performing RNA sequencing on single cell types (harvested via patch pipets) within the MB. We discover that the expression of genes that encode cell surface receptors is sufficient to identify cell types, and that a subset of these genes, required for sensory transduction in peripheral sensory neurons, are not only expressed within individual neurons of the MB in the central brain, but are also critical for memory formation.

Graphical Abstract



Introduction

The *Drosophila* mushroom bodies (MB) play an important role in olfactory learning and memory formation (de Belle and Heisenberg, 1994; McGuire et al., 2001; Pascual and Preat, 2001). The MB contains both intrinsic neurons (termed Kenyon cells or KCs) and extrinsic neurons (comprising output neurons (MBONs), neurons postsynaptic to the KCs, and several types of modulatory neuron) (Aso et al., 2014a; Tanaka et al., 2008). The ~2500 KCs in each brain hemisphere receive inputs from olfactory projection neurons, encode odors both sparsely and specifically (Murthy et al., 2008; Turner et al., 2008), and send their outputs into the MB lobes. The MB lobes can be subdivided broadly into 3 separate groups (α/β, α′/β′, and γ), based on where KCs target their axons (Crittenden et al., 1998; Lin et al., 2007). Previous studies have suggested that KCs innervating each lobe have distinct roles in learning and memory formation (short-term memory (STM) does not rely on new gene expression, whereas long-term memory (LTM) does (Lynch, 2004)). In general, these data point to γ KCs being important for STM, α′/β′ KCs playing a role in memory consolidation, and α/β KCs being important for LTM (Blum et al., 2009; Krashes et al., 2007; Tomchik and Davis, 2009; Trannoy et al., 2011; Yu et al., 2006) – although there are exceptions to this general picture (Aso et al., 2014b; Bouzaiane et al., 2015; Owald et al., 2015). Traces of memory formation (either molecular changes or changes in neural activity) have been found at MB intrinsic and extrinsic neuron terminals (Chen et al., 2012; Gervasi

et al., 2010; Owald et al., 2015; Pai et al., 2013; Placais et al., 2013; Sejourne et al., 2011; Tomchik and Davis, 2009; Yu et al., 2006; Yu et al., 2005; Zars, 2010). In addition, MBONs and dopaminergic neurons (encoding either reward or punishment signals) innervate specific domains or compartments of each of the MB lobes (Aso et al., 2014a; Perisse et al., 2013). Recent work has now demonstrated compartment-specific changes in synaptic strength at KC-to-MBON synapses following paired optogenetic activation of dopaminergic neurons with KC activation (Cohn et al., 2015; Hige et al., 2015a). Thus, a picture has emerged of where memory formation occurs (largely within the MB lobes, and at KC-to-MBON synapses in specific compartments), but we know much less regarding the molecular changes that underlie the formation of long-term memories within this structure.

To address this question, we require knowledge of gene expression within neurons of the MB. This relates to a broader problem in neuroscience – whether individual neurons of a single brain network can be differentiated based on gene expression (Luo et al., 2008). It is now possible to address this question with next generation sequencing methods (RNAseq) applied to single cells or cell types (Tang et al., 2009; Tasic et al., 2016). But for the networks studied so far, such as mouse sensory neurons, hippocampus, or cortex, it is not yet known if identifiable cell types exist (Macosko et al., 2015; Tasic et al., 2016; Usoskin et al., 2015; Zeisel et al., 2015). The small number of identifiable cell types in the *Drosophila* MB (e.g., ~21 MBON types) makes this an ideal brain structure for addressing this question. Moreover genetic enhancer lines reliably label identifiable neurons for cell harvesting. Here, we perform transcriptome analysis of *Drosophila* MB neurons by cell type, and we combine this method with single fly olfactory learning and memory assays, in order to also identify changes in gene expression that accompany memory formation.

Forward genetic screens in *Drosophila* have so far implicated >100 genes in memory formation (Aceves-Pina and Quinn, 1979; Boynton and Tully, 1992; Comas et al., 2004; Didelot et al., 2006; Dubnau et al., 2003; Dudai et al., 1976; Dura et al., 1993; Heisenberg et al., 1985; Quinn et al., 1979; Skoulakis and Davis, 1996; Tomchik and Davis, 2013; Walkinshaw et al., 2015). Such screens discovered genes confirmed to play important roles in memory formation across systems (Kesner and Martinez, 2007). Other strategies for identifying genes involved in memory formation in *Drosophila* include microarray studies on whole brains following the induction of LTM (Dubnau et al., 2003) or examining the downstream targets of the memory-related transcription factor CREB (Chen et al., 2012; Miyashita et al., 2012). For many of these genes, we either do not know the cell types of the MB each gene is expressed in, or it has not been confirmed that changes in gene expression are correlated with memory formation. We further hypothesize that the genes identified so far represent only a fraction of the genes involved in LTM. By applying our method of RNAseq by cell type, we focus on KCs and MB extrinsic neurons implicated in LTM, and we look within these specific neuron types for gene expression changes.

Results

Combining single-fly learning and memory assays with cell type-specific RNAseq

Our study had two goals: 1) to determine whether identified cell types within a single network of the *Drosophila* brain could be distinguished by transcriptional signatures, and 2)

to identify the transcriptional changes that underlie long-term memory (LTM) formation. We accomplished these separate but related goals by combining single-fly olfactory learning and memory assays and mRNA sequencing on *Drosophila* mushroom body (MB) neurons of a single type. We focused on a subset of intrinsic neurons (α/β and γ Kenyon cells (KC)) and extrinsic neurons (4 types of mushroom body output neuron (MBON) and the DAL neuron; Figure 1A). These neurons are identifiable by characterized genetic enhancer lines (Figure S1A) and were previously implicated in LTM formation or retrieval (Aso et al., 2014b; Chen et al., 2012; Pai et al., 2013; Placais et al., 2013; Sejourne et al., 2011). This includes the following MBON cell types (numbers of neurons per hemisphere in parentheses): i) V2 (7 neurons), ii) MBONα3 (2 neurons), iii) DAL (1 neuron), iv) MBON γ5β′2a (1 neuron), and v) MBON-β2β′2a (1 neuron). The V2 and MBONα3 dendrites innervate the α and α′ lobes, while the MBON γ5β′2a and MBON-β2β′2a dendrites innervate the γ, β, and β′ lobes (Aso et al., 2014a); consistent with receiving input from odor-responsive KCs, all four MBON types display odor responses (Hige et al., 2015b; Sejourne et al., 2011). By recording from the DAL neuron, via patch clamp methods, we observed inhibitory responses to odors, suggesting that it may also receive olfactory input in the MB (Figure S1B–C).

We harvested cells for RNAseq from flies (carrying the GAL4 constructs that label the aforementioned MB cell types) demonstrating learned odor avoidance (Figure 1A). Our single-fly assay is similar to previous designs (Chabaud et al., 2010; Claridge-Chang et al., 2009; Steck et al., 2012), but additionally facilitates recovery of the fly after acquiring a learning score and thereby permits the study of LTM expression. *Drosophila* require spaces of roughly 10 min between training sessions to produce transcription-dependent LTM (Yin et al., 1995). In our protocol, flies were first group trained to associate an odor with a foot shock, while controls received both the odor and foot shock unpaired in time (similar to (Krashes and Waddell, 2011)), and then assayed individually for learning (10 minutes following spaced training) or memory (24 hours after spaced training), 16 flies at a time (Figure 1B and Figure S2A–B). For testing, we delivered 4-methylcyclohexanol (4-MCH) or 3-octanol (3-OCT) to the two sides of each tube in which fly position was monitored. We defined learning and memory scores as the percent of five minutes spent on the trained versus untrained odor side (Figure S2A,F; thus a negative score indicates aversive learning (Figure 1C–D)). Each genetic enhancer (GAL4) fly line used in our study showed normal learning and memory scores (Figure S2G–H). Flies carrying a mutation for the *rutabaga* gene (which encodes a calcium/calmodulin adenylate cyclase (Levin et al., 1992)) did not exhibit learning in our assay (Figure 1C), consistent with previous studies (McGuire et al., 2003). In addition flies treated with cyclohexamide, a potent inhibitor of protein synthesis, failed to show LTM (Figure 1C–D). Finally, performance indices improve when learning and memory scores are calculated by subtracting the number of flies that biased towards the paired odor stream from the number of flies that biased towards the unpaired odor stream, divided by the total number of flies (Figure S2C); this is similar to the way scores are calculated from population T-maze assays (Krashes and Waddell, 2011; Tully et al., 1994).

Within 30 minutes of training, and following testing for learning, we harvested KCs or MB extrinsic neurons for RNAseq (Figure 1E). Neurons were identified by GAL4-driven GFP expression (Figure S1A), and cell bodies were extracted via patch clamp electrodes on an electrophysiology rig. We pooled ~100 neurons from a single fly for each KC sample and

~4–14 neurons from 1–2 flies for each MB extrinsic neuron sample. The use of patch clamp pipets for cell harvesting (Morris et al., 2011) avoided both cell damage and the use of detergents, and therefore decreased mRNA degradation (Figure S3), enabling us to use the poly-T primer based SMARTer technology for mRNA amplification (Ramskold et al., 2012). We modified this protocol by limiting the volume of harvesting buffer and reducing the number of PCR cycles for amplification to decrease CG sequence biases (Saliba et al., 2014). Despite the low amplification, we consistently generated enough cDNA from each sample for sequencing (Figure 1E and Figure S3). In total, we amplified and sequenced (at a depth of 30 million reads) 143 samples. We removed 36 of these samples because they failed to contain at least 4 million total uniquely mapped gene counts (a gene count of 1 corresponds to one read that is mapped to the genome and falls within the exon or exon-exon junction of a gene product; mapped gene counts were calculated using HTseq-counts (Anders et al., 2015)). Another 7 samples were removed for quality issues (see Experimental Procedures). This left us with 47 MB extrinsic neuron samples (10 DAL, 10 V2, 10 MBONα3, 9 MBON-β2β′2a, and 8 MBON-γ5β′2a), 24 KC samples (12 α/β (one technical replicate included) and 12 γ), 2 whole brain samples, and 27 whole fly samples (see Methods).

Mushroom body cell type-specific RNAseq is reliable and enriches for neuron-specific genes

Drosophila neurons are small (cell diameters from 1–10 microns (Echalier, 1997)), and because we pooled few cells per type, we first needed to address whether our sequencing method was reliable for such small starting volumes and concentrations of cDNA (~10–90 femtograms; Figure S3). We examined the correlation in counts for each gene present between different samples, and found a strong correlation in gene counts between technical replicates (the same sample of α/β KCs split into two tubes prior to cDNA synthesis, but post cell lysis; Figure 2A), indicating that our protocol is reliable. However, we found weaker correlations in low-expressing genes, consistent with the effects of noise in the PCR amplification (Brennecke et al., 2013; Efroni et al., 2015; Islam et al., 2011). We then went on to correlate gene expression counts from one of the α/β KC samples in Figure 2A with other sample types. We observed excellent correlation between biological replicates, which are α/β KC samples collected from two different brains (Figure 2B; $r^2 = 0.86$). We further observed similarity in gene expression between α/β and γ KC samples (Figure 2C; $r^2 = 0.82$), which was expected due to the limited number of genes that have been shown to be different between these two populations (Perrat et al., 2013). Gene counts were less well correlated between the α/β KC sample and an MBON sample (Figure 2D), and even less well correlated when comparing this sample to whole brain (Figure 2E) or whole fly (Figure 2F) samples, even though the whole fly sample came from the same genetic background as the KC sample. This last result is consistent with previous work demonstrating differences in gene lists between single-cell versus whole-animal RNAseq (Saliba et al., 2014).

Results from comparing gene counts for a given α/β KC sample to individual samples of other cell types (Figure 2C–F) indicated that we should be able to distinguish cell types based on gene expression profiles, one goal of this study. We performed principal components analysis (PCA) using the same gene list as in Figure 2A–F, and showed that

gene expression profiles from each of our neuronal samples were more similar to the gene expression profiles of 2 whole brain samples (from flies containing the UAS-eGFP transgene only) versus 27 whole fly samples from paired genetic backgrounds (flies carrying both GAL4 and UAS-eGFP) (Figure 2G). Thus, differences in gene expression between cell types are not likely to simply reflect differences in genetic backgrounds. Moreover, the first two principal components, which accounted for 45.74% of the variance, were not dominated by a small number of genes (Table S1). Next, we showed that α/β and γ KCs clustered separately using PCA (Figure 2H), consistent with microarray experiments (Perrat et al., 2013). This analysis further suggests that more than 3 biological replicates are required to distinguish between cell types.

We examined gene ontology (GO) term enrichment of our samples (Lyne et al., 2007), looking for biological processes (described by their GO term) that were over-represented in the high-expressing genes from our neuronal and whole fly samples. We compared 4233 genes with mapped reads > 50 CPM total gene counts (this corresponds to a minimum of 200 counts for a specific gene) from neuronal samples and 2097 genes from whole fly samples. Not surprisingly, our neuronal samples (n = 72) contained more nervous system-specific GO terms than those of the whole fly samples when looking at the top 50 most significantly enriched terms for each group (adj. p-value <0.01) (Figure 2I and Table S2). For example, the neuronal samples showed high expression of cell-cell signaling, neuron morphogenesis, and response to stimulus genes. We obtained a similar result when we used the same list of genes and compared with existing microarray data on various *Drosophila* tissues, including the brain, thoracic ganglion, ovaries, and testes (Figure S4A). We then compared the number of genes that were unique in each of our cell types with respect to the genes in whole brain or whole fly samples (Figure 2J). In all cases, there were more genes found in the neuronal samples that were not found in the whole fly sample (2643 genes), versus the whole brain sample (2006 genes). Thus, RNAseq of individual neuronal cell types increases the ability to detect nervous system-specific genes.

A Minimal Gene List That Identifies Cell Types

We next asked if the gene expression profile of a neuron allowed us to differentiate between cell types, even within a single network of the brain, and how much of the transcriptome was required for this identification. We computed a pairwise distance matrix for expression profiles of all samples (of 7 different cell types; see Figure 1A); based on our results comparing technical replicates (Figure 2A), distances were determined only on gene expression > 2 CPM in half the samples of each cell type, which included a total of 8,345 out of ~16,000 genes in the fly genome (Figure 3A). We then used a nearest neighbor classifier to assign samples with similar expression profiles to a cell type (Clemens et al., 2011). Results of this classification can be visualized as a confusion matrix, which tabulates the probability with which samples of a given cell type were classified correctly (Figure 3B). We found that the classifier successfully grouped samples by cell type based on the full gene list. The diagonal of the confusion matrix contains information on the overall similarity between decoded and actual cell types, and the mean of the diagonal represents the overall percent correct. This information score (see Experimental Procedures) using our full gene list is 2.68 (perfect score = 2.8).

To identify the minimal gene list with known biological function that identifies a cell type, we used the top GO terms pulled out for all of our neuronal samples (Figure 2I). By using GO terms we hoped to discover biological processes that identify neuron types, rather than focusing on differences in gene expression specific to our samples. For this same reason, we removed any GO term lists with < 100 genes (leaving 476 lists). We found that two gene lists contained as much information as the full list: cell surface receptor signaling genes (GO:00077166) and its sub-group, G-protein coupled receptor signaling (GO:0007186) (Figure 3C and Table S3). Cell surface receptor signaling also came up when we searched for GO terms that contained any genes expressed in only 1/7 cell types (Table S4). PCA based on the 434 genes under this GO term versus all 8345 genes revealed strong clustering by cell type (Figs. 3D–E). Included in this list are genes for receptors and neuropeptides, which drive the differences along the 1st principal component (Table S5). If we chose 434 genes at random (over multiple iterations) from any of the genes present in these samples, we were not able to cluster by cell type (Figure 3F–G). In addition, this result is not an artifact of the different genetic backgrounds of the flies (Figure S4B).

Cell Surface Receptor Gene Expression in KCs and MB extrinsic neurons

We next examined the expression of genes within the cell surface receptor signaling category by cell type (Figure 4). For this analysis, we pooled across the samples of a given cell type (both from flies that had formed associative memories and from control flies that had not). The cell surface receptor signaling gene list contains genes for neurotransmitter synthesis, transporters and receptors (Figure 4A), peptides and their receptors (Figure 4B), and sensory transduction (Figure 4C). Gene expression levels were normalized to account for differences in sample size, and were z-scored by gene (in other words, across each row in Figure 4) to highlight relative expression levels between cell types. Actual gene expression counts for each individual sample are plotted in Figure S5.

Through this analysis, we identified the putative neurotransmitter used by the different MB cell types (Figure 4A). We found that KCs are cholinergic (as in (Barnstedt et al., 2016; Perrat et al., 2013)), the DAL neuron is GABAergic (previously shown by (Chen et al., 2012)), both the MBON-γ5β′2a and MBON-β2β′2a neurons are glutamatergic (consistent with the findings of (Aso et al., 2014b)), and the MBONα3 and V2 neurons are cholinergic (previously shown by (Placais et al., 2013) and (Sejourne et al., 2011)). We detected evidence of potential octopamine synthesis in the MBONα3 neurons and dopamine synthesis in the V2 neurons. We were also able to identify the neurotransmitter receptors dominant in each of these cell types (Figure 4A). While we confirmed that KCs express DopR, DopR2, OAMB, and 5-HT1A and 1B receptors (Han et al., 1998; Han et al., 1996; Qin et al., 2012; Silva et al., 2014; Yuan et al., 2006), we discovered that these receptors are also expressed in MBONs, but at lower levels than in the KCs (see also Figure S5). This indicates that both the KCs and MBONs have the ability to respond to the dopamine signal (carrying information about punishment or reward) innervating each compartment of the MB lobes. α/β KCs express a large number of neurotransmitter receptors, including TyrR (tyramine receptor), Octbeta2R, Glu-R1B, GluRIIA, and GABA-B-R1 and R2. In addition, previous work demonstrated a role for the rdl (GABA) receptor in memory formation (Liu et al., 2007), and we found this receptor is expressed strongly in both the KCs and V2

MBONs. Nicotinic AchR calcium currents and fast excitatory synaptic transmission are known to be modulated in *Drosophila* and honeybees by octopamine and dopamine signaling (Dupuis et al., 2012; Leyton et al., 2014) but determining the expression of specific nAchR subunits has been difficult. Here, we observed a diverse range of subunits expressed in different MB neurons. For instance the nAchRa7E (also known as a3) is enriched in both the V2 neurons and γ KCs. nAchRalpha-80B (also known as a4) is almost absent from the KCs but expressed in the MBONα3, MBON-γ5β′2a′ and V2 neurons, and nAchR-18C (also known as a7) is expressed at higher rates in γ KCs, V2, and MBONα3. Finally, we found that NMDA receptors, known to be involved in *Drosophila* memory formation (Xia et al., 2005), are expressed in both the KCs and MB extrinsic neurons.

Many of the MB neurons also express neuropeptides and neuropeptide receptors, potent neuromodulators (Figure 4B). We found that sNPF is expressed at very high levels in α/β and γ KCs as well as in the V2 cluster, while all MB extrinsic neurons sampled express the sNPF receptor. As a whole, the MB extrinsic neurons have a more diverse expression of neuropeptides and neuropeptide receptors than the KCs. For instance MBONα3 expressed relatively higher levels of capa, Dh31 (Diuretic Hormone 31), nplp3 (neuropeptide like precursor 1), pigment dispersing factor (Pdf), Leucokinin (LK), hugin (Hug), juvenile hormone like 21 and CCHamide-1 (CCHa1). In contrast, the DAL neuron expressed the relatively highest level of neuropeptide F (npf), nplp1, insulin like peptide 7 (ilp7), FMRFamide (FmrF), Proctolin (Proct), and Juvenile hormone like 26. These two neuronal cell types also expressed relatively different levels of peptide receptors. We also found that most of the MB cells expressed the insulin receptor (InR) at varying levels, with relatively higher expression in the V2, MBONa3 and γ KCs. This receptor is required for synaptic plasticity in sensory neurons (Root et al., 2011). In addition, insulin signaling in the mushroom bodies has been previously implicated in both learning and memory formation (Chambers et al., 2015; Naganos et al., 2012). Finally, the MB is required for normal sleep and wake in the fly, but how it receives information from the circadian system is unknown (Joiner et al., 2006). We found that the pdf receptor is most strongly expressed in the DAL and MBON-γ5β′2a neurons, which could permit integration of the circadian signal into the output of the MB.

To our surprise, we also found many sensory transduction genes in the MB neurons (Figure 4C). We detected the expression of olfactory genes, such as odorant receptors (ORs and IRs) and odorant binding proteins, visual genes, such as rhodopsins and other genes involved in light detection, and gustatory genes, such as the gustatory receptors (Figure 4 and Figure S5). Our *in vivo* dissection for cell harvesting leaves the sensory organs of the fly intact, and thus we do not think that the detection of these RNAs in the MB reflects an issue with contamination. These genes were also expressed in a cell type-specific manner. For example, Rh2 (rhodopsin 2) was expressed more highly in the MBON-β2β′2a and MBON-γ5β′2a neurons, with almost no expression in the MBONα3 neuron. Rh3 (rhodopsin 3) was expressed at relatively higher levels in the DAL, MBONα3 and MBON-β2β′2a neurons, but mostly absent from the MBON-γ5β′2a neuron (Figure S5). This diversity of expression also applied to the KCs: for example, the α/β KCs expressed higher levels of Ir47a and Ir56a relative to the γ KCs, while the γ KCs expressed a higher level of Ir93a. While the role of

these genes in the MB is not yet known, there is precedence for detecting genes in these categories in the central brain (see Discussion).

Differential gene expression following memory formation

We next determined if we could detect differential gene expression following memory formation by comparing paired (flies that received paired odor and shock using a spaced training protocol) and unpaired (odor and shock unpaired in time) samples from the same day and circadian harvest time. We harvested samples within 30 minutes of the end of the spaced training (Yin et al., 1995), and only from flies that showed a strong learning score (a score of < −50) following the paired protocol (control flies showed no strong odor preference (a score of ~0) following the unpaired protocol). There were no differences in the abilities of flies to form memories with either odor used (Figure 1C–D). We found that 3 cell types (DAL, V2, and MBONα3) showed abundant differential expression following memory induction (Figure 5A–C). These cell types had been previously found to be necessary for aversive LTM expression (Chen et al., 2012; Pai et al., 2013; Sejourne et al., 2011). Minimal differential expression was seen in the other 2 MBONs – these innervate the γ, β, or β′ lobes, which were implicated in appetitive, but not aversive, LTM (Owald et al., 2015). We also did not observe significant differential expression in the KCs (Figure S6, Table S6), likely because we pooled ~100 KCs per sample, independent of which odors (3-OCT or 4-MCH) they encode.

Focusing on the DAL, V2, and MBONα3 cell types, we found a total of 235 genes up-regulated and 155 genes down-regulated with q-value < 0.05. We calculate the q-value instead of a Bonferroni adjusted p-value because of the large number of t-tests performed (Storey, 2015). The complete differential gene expression list is found in Table S6. To validate this list, we performed qPCR on new cell isolates, following associative learning, and correlated gene counts between RNAseq and qPCR for a subset of 23 genes (Figure S6E; correlations ranged from an r^2 of 0.45 to 0.7). We then chose 5 genes from this list that showed differential expression in the MBONα3 cell type when comparing flies that had received paired versus unpaired training. We asked if these genes also showed differential expression using qPCR analysis: one of these genes (*ninaC*) showed clear differential expression (between trained and untrained samples) by qPCR (Figure S6F). Table 1 shows the genes we found to be up-regulated following memory formation in our assay (with q-value < 0.1) that were also previously implicated in memory formation in other studies. We next looked for genes found in similar biological processes (specifically GO term enrichment) that were differentially expressed following learning and memory across all 3 cell types (Table 2). Of the 12 processes we identified (including neurogenesis, metabolic processes, and nervous system development), three of these related to the cellular response to light. Thus, not only are such genes expressed outside of the eye (Figure 4C), but we have uncovered a potential role for them in learning and memory.

To test whether these light-sensing genes play a role in memory formation, we screened a handful of available mutants in our single-fly behavioral assay. We tested mutants for *NinaC*^{5}, *pinta*^{1}, *Rh3*^{1} and *Rh4*^{1} (Senthilan et al., 2012; Shen et al., 2011; Vasiliauskas et al., 2011; Wang and Montell, 2005), which were all found to be differentially expressed after

LTM formation (Figure 5A–C). We found a role for all of these genes in memory formation, and 2 out of 4 were also required for learning (Figure 5D–H). This phenotype was not due to a defect in locomotion, odor acuity, or shock reactivity (Figure S7), nor was it due to a defect in vision, as chemotaxis assays were conducted in the dark. As a control, we found that *NinaE*^{8} mutants, despite not locomoting well, were still able to form long-term memories; this gene, although involved in light-sensing and detected in our neuron samples, was not found to be differentially expressed following LTM formation (Figure 5A–B and Figure S7). These results validate our RNAseq method for identifying, by cell type, new genes involved in learning and memory formation.

Discussion

Here we present a method for RNAseq by single cell type in *Drosophila* (using harvesting via patch pipets), and apply this method to the study of the mushroom body, a brain structure important for learning and memory. We amplified and sequenced RNAs effectively from very small starting material (just a few *Drosophila* neurons per sample); our method not only enriched for nervous system genes, but it also permitted the study of genes whose products are known to be in low abundance within neurons, such as cell surface receptors. In fact, we found that genes belonging to the cell surface receptor signaling category are highly informative for separating cell types within the MB. Previous studies demonstrated that cells can be identified based on subsets of genes, but these genes were not typically linked in terms of cellular function (Tasic et al., 2016; Usoskin et al., 2015; Zeisel et al., 2015).

Our results also lend support to the notion that combinatoric expression of a small subset of genes can nonetheless give rise to a number of functionally different cell types (Hobert et al., 2010). Our method also identified the specific genes that are expressed in the *Drosophila* KCs and MB extrinsic neurons, which will facilitate building more specific reagents to target these neurons. We found that acetylcholine is the major neurotransmitter expressed in the KCs (consistent with (Barnstedt et al., 2016)). The α/β and γ KCs expressed the serotonin, dopamine, and octopamine receptors at higher levels than the MBONs we sampled, suggesting that neuromodulation acts largely presynaptically in the MB circuit (consistent with (Cohn et al., 2015) and (Hige et al., 2015a)). The γ KC samples expressed higher levels of many of the peptide receptors than the α/β KCs, including the insulin receptor, proctolin receptor, ecdysone receptor, pdf receptor, CCKLR-17D1 and 3, and the SIFamide receptor. An exception to this was expression of the fly growth hormone receptors (DH44-R1 and 2), which were expressed at higher levels in the α/β KCs. Increased expression of the peptide receptors in γ KCs may reflect developmental gene expression patterns, because the γ KCs are the only KCs present during larval stages and through metamorphosis (Armstrong et al., 1998). It also may reflect the more diverse role the γ KCs have on behavior including courtship, appetitive memory and sleep (Aso et al., 2014b; Joiner et al., 2006; Keleman et al., 2012; Owald et al., 2015).

We also found that the MBONs express a wide variety of both neurotransmitter and peptide receptors. This allowed us to distinguish between cell types, even when they expressed the same fast-acting neurotransmitter, such as the MBON-γ5β′2a and MBON-β2β′2a neurons, which are both glutamatergic. These two cells differ in their expression of the serotonin

receptors, with the MBON-γ5β′2a expressing higher levels of the 5-HT7 receptor and MBON-β2β′2a expressing higher levels of the 5-HT2 receptor. This difference may highlight other roles these neurons play either in either courtship (Becnel et al., 2011) or feeding and aggression (Gasque et al., 2013; Johnson et al., 2009). These two cell types also differed in their expression of the pdf and leucokinin receptors, with MBON-γ5β′2a expressing higher levels of both receptors; these receptors are known to be important for feeding and circadian behaviors (Al-Anzi et al., 2010; Peschel and Helfrich-Forster, 2011). Thus, these data suggest a mechanism for the established relationships between memory formation and circadian or feeding rhythms (Joiner et al., 2006; Krashes et al., 2009).

We also discovered that classical sensory transduction genes are expressed throughout the MB, with each cell type we sampled expressing a different subset of these genes. There is precedence for ORs (Flegel et al., 2013; Otaki et al., 2004) being expressed in the brain of other systems including humans, although their function is unknown. In addition, several GRs (including Gr28b, identified in our study) have been detected outside of the primary gustatory organs of *Drosophila*, including in the CNS (Thorne and Amrein, 2008). While rhodopsin gene expression has been detected in the brains of some vertebrates (Masuda et al., 2003; Wada et al., 1998), its expression was found only in photosensitive cells. We found a surprising role for some light-sensing genes in memory formation (e.g., *NinaC, pinta, Rh3, and Rh4)*: they were differentially expressed within specific MB extrinsic neurons following LTM formation, and mutants caused defects in LTM (see Figure 5). Since our memory assays were conducted in complete darkness, these proteins must have functions outside of light sensing. In support of this notion, recent studies (in *Drosophila* peripheral sensory neurons) showed a role for this class of genes in mechanosensation (Senthilan et al., 2012) or thermosensation (Shen et al., 2011). Taken together, we hypothesize that genes previously implicated in visual signal transduction may be of broad relevance for neuronal signaling and plasticity throughout the nervous system. It will be interesting to determine, as a future direction, the expression profile of all ~21 classes of mushroom body output neuron, each which can be identified by specific intersectional driver lines (Aso et al., 2014a), to determine if expression of these genes is a general feature of MBONs. In addition, knock-down of classical sensory transduction genes in each of (or combinations of) the MBONs should shed light on their role in specific phases of LTM.

More broadly, our study demonstrates that RNAseq by cell type is an effective method for identifying new genes important for memory formation. While we only detected differential expression following induction of LTM in 3 types of MB extrinsic neuron, we might have seen differential expression in KCs by sampling from smaller subsets. Given that the tools for labeling mushroom body cell types are developing rapidly (Aso et al., 2014a) and that functional roles of specific cell types are being worked out in detail (Aso et al., 2014b; Chen et al., 2012; Cohn et al., 2015; Hige et al., 2015a; Hige et al., 2015b; Owald et al., 2015; Sejourne et al., 2011), our study provides a timely addition to the mushroom body literature – a way forward from neural circuit to molecular mechanisms. In addition, by combining cell type-specific transcriptome analysis with single fly learning and memory assays, as we have done here, future studies can determine how variability in learning or memory performance relates to variability in gene expression within individual neurons. Here, we validated only 4 of the genes we identified as being differentially expressed following

memory formation, with genetic mutants and behavioral assays. However, the long gene list we discovered should serve as an important launching point for future studies into the still mysterious molecular mechanisms underlying learning and memory.

Experimental Procedures

Single Fly Olfactory Learning and Memory Assay

Flies were group trained, but assayed individually, 16 at a time, using a modified Trikinetics' multibeam monitor that measures fly movements in single tubes. One group of flies was tested 10 minutes after the end of the 8 spaced training sessions to score immediate learning. A second group of flies were placed on food to test the next day for 24 hour memory. Fly movements were recorded for 5 minutes. Odor preference (Learning or Memory score) was calculated as the percentage of time over the 5 minutes spent in trained odor space. All learning and memory scores were normally distributed as determined by Kolmogorov-Smirnov test and a students t-test was used to determine significance. See Supplemental Experimental Procedures for more details, along with a list of all fly lines used.

Single Cell-Type RNA Sequencing

Cells labeled by GAL4 lines were harvested via patch pipets and the Clontech HV SMARTer Ultra Low RNAseq kit was used for mRNA/cDNA amplification. qPCR experiments were performed on separate cell isolates. Samples were sheared to 200bp fragments, and libraries were made using IntegenX's Apollo 324 automated library prep system. Samples were then barcoded (Bio Scientific). Libraries were run on the Illumina HiSeq2500, 12 samples per lane, and each sample run across two lanes. This resulted in a sequencing depth of 30 million reads. Sequences were mapped to the fly genome using TopHat2 with Bowtie2 (Kim et al., 2013), and processed using HTseq-count (Anders et al., 2015) to map exons and determine gene counts. See Supplemental Experimental Procedures for more details.

Gene Expression and Ontology Analysis

Samples needed to contain more than 4 million counts following HTseq-count mapping to be included in analysis. We defined high expressing genes as any gene with a count of at least 50 counts per million (CPM). Principal components analysis (PCA) was performed on normalized gene counts using DESeq2 (Love et al., 2014) (regularized log transformation of normalized data) after removing genes with less than 2 CPM in the α/β KC samples. For gene ontology (GO) analysis, the list of genes was generated by determining genes present at values >50 CPM in neuronal samples (4233 genes) and whole fly samples (2097 genes). We also determined a list of all unique genes -- present in only one cell type (576 genes), and a list of all up-regulated genes at a q-value <0.1 in the DAL, MBONα3, V2 following learning (235 genes). See Supplemental Experimental Procedures for more details.

Classification of Cell Types

We used a nearest neighbor classifier on distance matrices (city block metric) to assign each sample's gene expression profile to a cell type (Clemens et al., 2011). The results of the

nearest neighbor classifier were tabulated in a confusion matrix, from which information scores were derived. See Supplemental Experimental Procedures for more details.

Differential Gene Expression

Differential expression was determined using DESeq2 (version 1.1.3)(Love et al., 2014), which took into account batch effects. We determined q values from these data. We removed extreme outliers or highly variable genes using Cooks distance, used trimmed means to replace a single outlier, and removed data with no information from the analysis. These results were verified using the single cell to CT kit and Taqman assays (Life Technologies) to quantify gene abundance by qPCR. See Supplemental Experimental Procedures for more details.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dave Robinson (Storey Lab) and Jan Clemens (Murthy lab) for assistance with data analysis, Gregory Pil and Sarady Merghani (Crocker Lab) for technical assistance with fly genetic experiments, and Amita Seghal and Vanisha Lahkina for feedback on the manuscript. AC was funded by a T32 training grant and an NRSA postdoctoral fellowship. MM is funded by the Alfred P. Sloan Foundation, the Human Frontiers Science Program, an NSF CAREER award, the McKnight Endowment Fund, the Klingenstein Foundation, an NIH New Innovator award, and an NSF BRAIN Initiative EAGER award. RNAseq experiments were funded by the Glenn Centers for Aging Research, Coleen T. Murphy Director.

References Cited

Aceves-Pina EO, Quinn WG. Learning in normal and mutant Drosophila larvae. Science. 1979; 206:93–96. [PubMed: 17812455]

Al-Anzi B, Armand E, Nagamei P, Olszewski M, Sapin V, Waters C, Zinn K, Wyman RJ, Benzer S. The leucokinin pathway and its neurons regulate meal size in Drosophila. Curr Biol. 2010; 20:969–978. [PubMed: 20493701]

Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015; 31:166–169. [PubMed: 25260700]

Armstrong JD, de Belle JS, Wang Z, Kaiser K. Metamorphosis of the mushroom bodies; large-scale rearrangements of the neural substrates for associative learning and memory in Drosophila. Learn Mem. 1998; 5:102–114. [PubMed: 10454375]

Aso Y, Hattori D, Yu Y, Johnston RM, Iyer NA, Ngo TT, Dionne H, Abbott LF, Axel R, Tanimoto H, et al. The neuronal architecture of the mushroom body provides a logic for associative learning. Elife. 2014a; 3:e04577. [PubMed: 25535793]

Aso Y, Sitaraman D, Ichinose T, Kaun KR, Vogt K, Belliart-Guerin G, Placais PY, Robie AA, Yamagata N, Schnaitmann C, et al. Mushroom body output neurons encode valence and guide memory-based action selection in Drosophila. Elife. 2014b; 3:e04580. [PubMed: 25535794]

Barnstedt O, Owald D, Felsenberg J, Brain R, Moszynski JP, Talbot CB, Perrat PN, Waddell S. Memory-Relevant Mushroom Body Output Synapses Are Cholinergic. Neuron. 2016; 89:1237–1247. [PubMed: 26948892]

Becnel J, Johnson O, Luo J, Nassel DR, Nichols CD. The serotonin 5-HT7Dro receptor is expressed in the brain of Drosophila, and is essential for normal courtship and mating. PLoS One. 2011; 6:e20800. [PubMed: 21674056]

Blum AL, Li W, Cressy M, Dubnau J. Short- and long-term memory in Drosophila require cAMP signaling in distinct neuron types. Curr Biol. 2009; 19:1341–1350. [PubMed: 19646879]

Bouzaiane E, Trannoy S, Scheunemann L, Placais PY, Preat T. Two independent mushroom body output circuits retrieve the six discrete components of Drosophila aversive memory. Cell Rep. 2015; 11:1280–1292. [PubMed: 25981036]

Boynton S, Tully T. latheo, a new gene involved in associative learning and memory in Drosophila melanogaster, identified from P element mutagenesis. Genetics. 1992; 131:655–672. [PubMed: 1321066]

Brennecke P, Anders S, Kim JK, Kolodziejczyk AA, Zhang X, Proserpio V, Baying B, Benes V, Teichmann SA, Marioni JC, et al. Accounting for technical noise in single-cell RNA-seq experiments. Nat Methods. 2013; 10:1093–1095. [PubMed: 24056876]

Chabaud MA, Preat T, Kaiser L. Behavioral characterization of individual olfactory memory retrieval in Drosophila melanogaster. Front Behav Neurosci. 2010; 4:192. [PubMed: 21258642]

Chambers DB, Androschuk A, Rosenfelt C, Langer S, Harding M, Bolduc FV. Insulin signaling is acutely required for long-term memory in Drosophila. Front Neural Circuits. 2015; 9:8. [PubMed: 25805973]

Chen CC, Wu JK, Lin HW, Pai TP, Fu TF, Wu CL, Tully T, Chiang AS. Visualizing long-term memory formation in two neurons of the Drosophila brain. Science. 2012; 335:678–685. [PubMed: 22323813]

Claridge-Chang A, Roorda RD, Vrontou E, Sjulson L, Li H, Hirsh J, Miesenbock G. Writing memories with light-addressable reinforcement circuitry. Cell. 2009; 139:405–415. [PubMed: 19837039]

Clemens J, Kutzki O, Ronacher B, Schreiber S, Wohlgemuth S. Efficient transformation of an auditory population code in a small sensory system. Proc Natl Acad Sci U S A. 2011; 108:13812–13817. [PubMed: 21825132]

Cohn R, Morantte I, Ruta V. Coordinated and Compartmentalized Neuromodulation Shapes Sensory Processing in Drosophila. Cell. 2015; 163:1742–1755. [PubMed: 26687359]

Comas D, Petit F, Preat T. Drosophila long-term memory formation involves regulation of cathepsin activity. Nature. 2004; 430:460–463. [PubMed: 15269770]

Crittenden JR, Skoulakis EM, Han KA, Kalderon D, Davis RL. Tripartite mushroom body architecture revealed by antigenic markers. Learn Mem. 1998; 5:38–51. [PubMed: 10454371]

de Belle JS, Heisenberg M. Associative odor learning in Drosophila abolished by chemical ablation of mushroom bodies. Science. 1994; 263:692–695. [PubMed: 8303280]

Didelot G, Molinari F, Tchenio P, Comas D, Milhiet E, Munnich A, Colleaux L, Preat T. Tequila, a neurotrypsin ortholog, regulates long-term memory formation in Drosophila. Science. 2006; 313:851–853. [PubMed: 16902143]

Dubnau J, Chiang AS, Grady L, Barditch J, Gossweiler S, McNeil J, Smith P, Buldoc F, Scott R, Certa U, et al. The staufen/pumilio pathway is involved in Drosophila long-term memory. Curr Biol. 2003; 13:286–296. [PubMed: 12593794]

Dudai Y, Jan YN, Byers D, Quinn WG, Benzer S. dunce, a mutant of Drosophila deficient in learning. Proc Natl Acad Sci U S A. 1976; 73:1684–1688. [PubMed: 818641]

Dupuis J, Louis T, Gauthier M, Raymond V. Insights from honeybee (Apis mellifera) and fly (Drosophila melanogaster) nicotinic acetylcholine receptors: from genes to behavioral functions. Neurosci Biobehav Rev. 2012; 36:1553–1564. [PubMed: 22525891]

Dura JM, Preat T, Tully T. Identification of linotte, a new gene affecting learning and memory in Drosophila melanogaster. J Neurogenet. 1993; 9:1–14. [PubMed: 8295074]

Echalier, G. Drosophila Cells in Culture. New York: Academic Press; 1997. Drosophila Cells in Culture; p. 107

Efroni I, Ip PL, Nawy T, Mello A, Birnbaum KD. Quantification of cell identity from single-cell gene expression profiles. Genome Biol. 2015; 16:9. [PubMed: 25608970]

Flegel C, Manteniotis S, Osthold S, Hatt H, Gisselmann G. Expression profile of ectopic olfactory receptors determined by deep sequencing. PLoS One. 2013; 8:e55368. [PubMed: 23405139]

Gasque G, Conway S, Huang J, Rao Y, Vosshall LB. Small molecule drug screening in Drosophila identifies the 5HT2A receptor as a feeding modulation target. Sci Rep. 2013; 3:srep02120. [PubMed: 23817146]

Gervasi N, Tchenio P, Preat T. PKA dynamics in a Drosophila learning center: coincidence detection by rutabaga adenylyl cyclase and spatial regulation by dunce phosphodiesterase. Neuron. 2010; 65:516–529. [PubMed: 20188656]

Han KA, Millar NS, Davis RL. A novel octopamine receptor with preferential expression in Drosophila mushroom bodies. J Neurosci. 1998; 18:3650–3658. [PubMed: 9570796]

Han KA, Millar NS, Grotewiel MS, Davis RL. DAMB, a novel dopamine receptor expressed specifically in Drosophila mushroom bodies. Neuron. 1996; 16:1127–1135. [PubMed: 8663989]

Heisenberg M, Borst A, Wagner S, Byers D. Drosophila mushroom body mutants are deficient in olfactory learning. J Neurogenet. 1985; 2:1–30. [PubMed: 4020527]

Hige T, Aso Y, Modi MN, Rubin GM, Turner GC. Heterosynaptic Plasticity Underlies Aversive Olfactory Learning in Drosophila. Neuron. 2015a; 88:985–998. [PubMed: 26637800]

Hige T, Aso Y, Rubin GM, Turner GC. Plasticity-driven individualization of olfactory coding in mushroom body output neurons. Nature. 2015b; 526:258–262. [PubMed: 26416731]

Hobert O, Carrera I, Stefanakis N. The molecular and gene regulatory signature of a neuron. Trends Neurosci. 2010; 33:435–445. [PubMed: 20663572]

Islam S, Kjallquist U, Moliner A, Zajac P, Fan JB, Lonnerberg P, Linnarsson S. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res. 2011; 21:1160–1167. [PubMed: 21543516]

Johnson O, Becnel J, Nichols CD. Serotonin 5-HT(2) and 5-HT(1A)-like receptors differentially modulate aggressive behaviors in Drosophila melanogaster. Neuroscience. 2009; 158:1292–1300. [PubMed: 19041376]

Joiner WJ, Crocker A, White BH, Sehgal A. Sleep in Drosophila is regulated by adult mushroom bodies. Nature. 2006; 441:757–760. [PubMed: 16760980]

Keleman K, Vrontou E, Kruttner S, Yu JY, Kurtovic-Kozaric A, Dickson BJ. Dopamine neurons modulate pheromone responses in Drosophila courtship learning. Nature. 2012; 489:145–149. [PubMed: 22902500]

Kesner, RP.; Martinez, JL. Neurobiology of learning and memory. 2. Boston, MA: Butterworth-Heinemann, an imprint of Elsevier; 2007.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013; 14:R36. [PubMed: 23618408]

Krashes MJ, DasGupta S, Vreede A, White B, Armstrong JD, Waddell S. A neural circuit mechanism integrating motivational state with memory expression in Drosophila. Cell. 2009; 139:416–427. [PubMed: 19837040]

Krashes MJ, Keene AC, Leung B, Armstrong JD, Waddell S. Sequential use of mushroom body neuron subsets during drosophila odor memory processing. Neuron. 2007; 53:103–115. [PubMed: 17196534]

Krashes MJ, Waddell S. Drosophila aversive olfactory conditioning. Cold Spring Harb Protoc. 2011; 2011 pdb prot5608.

Levin LR, Han PL, Hwang PM, Feinstein PG, Davis RL, Reed RR. The Drosophila learning and memory gene rutabaga encodes a Ca2+/Calmodulin-responsive adenylyl cyclase. Cell. 1992; 68:479–489. [PubMed: 1739965]

Leyton V, Goles NI, Fuenzalida-Uribe N, Campusano JM. Octopamine and Dopamine differentially modulate the nicotine-induced calcium response in Drosophila Mushroom Body Kenyon Cells. Neurosci Lett. 2014; 560:16–20. [PubMed: 24334164]

Lin HH, Lai JS, Chin AL, Chen YC, Chiang AS. A map of olfactory representation in the Drosophila mushroom body. Cell. 2007; 128:1205–1217. [PubMed: 17382887]

Liu X, Krause WC, Davis RL. GABAA receptor RDL inhibits Drosophila olfactory associative learning. Neuron. 2007; 56:1090–1102. [PubMed: 18093529]

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15:550. [PubMed: 25516281]

Luo L, Callaway EM, Svoboda K. Genetic dissection of neural circuits. Neuron. 2008; 57:634–660. [PubMed: 18341986]

Lynch MA. Long-term potentiation and memory. Physiol Rev. 2004; 84:87–136. [PubMed: 14715912]

Lyne R, Smith R, Rutherford K, Wakeling M, Varley A, Guillier F, Janssens H, Ji W, McLaren P, North P, et al. FlyMine: an integrated database for Drosophila and Anopheles genomics. Genome Biol. 2007; 8:R129. [PubMed: 17615057]

Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 2015; 161:1202–1214. [PubMed: 26000488]

Masuda T, Iigo M, Mizusawa K, Aida K. Retina-type rhodopsin gene expressed in the brain of a teleost, ayu (Plecoglossus altivelis). Zoolog Sci. 2003; 20:989–997. [PubMed: 12951405]

McGuire SE, Le PT, Davis RL. The role of Drosophila mushroom body signaling in olfactory memory. Science. 2001; 293:1330–1333. [PubMed: 11397912]

McGuire SE, Le PT, Osborn AJ, Matsumoto K, Davis RL. Spatiotemporal rescue of memory dysfunction in Drosophila. Science. 2003; 302:1765–1768. [PubMed: 14657498]

Miyashita T, Oda Y, Horiuchi J, Yin JC, Morimoto T, Saitoe M. Mg(2+) block of Drosophila NMDA receptors is required for long-term memory formation and CREB-dependent gene expression. Neuron. 2012; 74:887–898. [PubMed: 22681692]

Morris J, Singh JM, Eberwine JH. Transcriptome analysis of single cells. J Vis Exp. 2011

Murthy M, Fiete I, Laurent G. Testing odor response stereotypy in the Drosophila mushroom body. Neuron. 2008; 59:1009–1023. [PubMed: 18817738]

Naganos S, Horiuchi J, Saitoe M. Mutations in the Drosophila insulin receptor substrate, CHICO, impair olfactory associative learning. Neurosci Res. 2012; 73:49–55. [PubMed: 22342328]

Otaki JM, Yamamoto H, Firestein S. Odorant receptor expression in the mouse cerebral cortex. J Neurobiol. 2004; 58:315–327. [PubMed: 14750145]

Owald D, Felsenberg J, Talbot CB, Das G, Perisse E, Huetteroth W, Waddell S. Activity of defined mushroom body output neurons underlies learned olfactory behavior in Drosophila. Neuron. 2015; 86:417–427. [PubMed: 25864636]

Pai TP, Chen CC, Lin HH, Chin AL, Lai JS, Lee PT, Tully T, Chiang AS. Drosophila ORB protein in two mushroom body output neurons is necessary for long-term memory formation. Proc Natl Acad Sci U S A. 2013; 110:7898–7903. [PubMed: 23610406]

Pascual A, Preat T. Localization of long-term memory within the Drosophila mushroom body. Science. 2001; 294:1115–1117. [PubMed: 11691997]

Perisse E, Yin Y, Lin AC, Lin S, Huetteroth W, Waddell S. Different kenyon cell populations drive learned approach and avoidance in Drosophila. Neuron. 2013; 79:945–956. [PubMed: 24012007]

Perrat PN, DasGupta S, Wang J, Theurkauf W, Weng Z, Rosbash M, Waddell S. Transposition-driven genomic heterogeneity in the Drosophila brain. Science. 2013; 340:91–95. [PubMed: 23559253]

Peschel N, Helfrich-Forster C. Setting the clock--by nature: circadian rhythm in the fruitfly Drosophila melanogaster. FEBS Lett. 2011; 585:1435–1442. [PubMed: 21354415]

Placais PY, Trannoy S, Friedrich AB, Tanimoto H, Preat T. Two pairs of mushroom body efferent neurons are required for appetitive long-term memory retrieval in Drosophila. Cell Rep. 2013; 5:769–780. [PubMed: 24209748]

Qin H, Cressy M, Li W, Coravos JS, Izzi SA, Dubnau J. Gamma neurons mediate dopaminergic input during aversive olfactory memory formation in Drosophila. Curr Biol. 2012; 22:608–614. [PubMed: 22425153]

Quinn WG, Sziber PP, Booker R. The Drosophila memory mutant amnesiac. Nature. 1979; 277:212–214. [PubMed: 121760]

Ramskold D, Luo S, Wang YC, Li R, Deng Q, Faridani OR, Daniels GA, Khrebtukova I, Loring JF, Laurent LC, et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol. 2012; 30:777–782. [PubMed: 22820318]

Root CM, Ko KI, Jafari A, Wang JW. Presynaptic facilitation by neuropeptide signaling mediates odor-driven food search. Cell. 2011; 145:133–144. [PubMed: 21458672]

Saliba AE, Westermann AJ, Gorski SA, Vogel J. Single-cell RNA-seq: advances and future challenges. Nucleic Acids Res. 2014; 42:8845–8860. [PubMed: 25053837]

Sejourne J, Placais PY, Aso Y, Siwanowicz I, Trannoy S, Thoma V, Tedjakumala SR, Rubin GM, Tchenio P, Ito K, et al. Mushroom body efferent neurons responsible for aversive olfactory memory retrieval in Drosophila. Nat Neurosci. 2011; 14:903–910. [PubMed: 21685917]

Senthilan PR, Piepenbrock D, Ovezmyradov G, Nadrowski B, Bechstedt S, Pauls S, Winkler M, Mobius W, Howard J, Gopfert MC. Drosophila auditory organ genes and genetic hearing defects. Cell. 2012; 150:1042–1054. [PubMed: 22939627]

Shen WL, Kwon Y, Adegbola AA, Luo J, Chess A, Montell C. Function of rhodopsin in temperature discrimination in Drosophila. Science. 2011; 331:1333–1336. [PubMed: 21393546]

Silva B, Goles NI, Varas R, Campusano JM. Serotonin receptors expressed in Drosophila mushroom bodies differentially modulate larval locomotion. PLoS One. 2014; 9:e89641. [PubMed: 24586928]

Skoulakis EM, Davis RL. Olfactory learning deficits in mutants for leonardo, a Drosophila gene encoding a 14-3-3 protein. Neuron. 1996; 17:931–944. [PubMed: 8938125]

Steck K, Veit D, Grandy R, Badia SB, Mathews Z, Verschure P, Hansson BS, Knaden M. A high-throughput behavioral paradigm for Drosophila olfaction - The Flywalk. Sci Rep. 2012; 2:361. [PubMed: 22511996]

Storey, J. R package version 2.0.0. 2015. qvalue: Q-value estimation for false discovery rate control.

Tanaka NK, Tanimoto H, Ito K. Neuronal assemblies of the Drosophila mushroom body. J Comp Neurol. 2008; 508:711–755. [PubMed: 18395827]

Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A, et al. mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods. 2009; 6:377–382. [PubMed: 19349980]

Tasic B, Menon V, Nguyen TN, Kim TK, Jarsky T, Yao Z, Levi B, Gray LT, Sorensen SA, Dolbeare T, et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat Neurosci. 2016; 19:335–346. [PubMed: 26727548]

Thorne N, Amrein H. Atypical expression of Drosophila gustatory receptor genes in sensory and central neurons. J Comp Neurol. 2008; 506:548–568. [PubMed: 18067151]

Tomchik SM, Davis RL. Dynamics of learning-related cAMP signaling and stimulus integration in the Drosophila olfactory pathway. Neuron. 2009; 64:510–521. [PubMed: 19945393]

Tomchik, SM.; Davis, RL., editors. Drosophila Memory Research through Four Eras: Genetic, Molecular Biology, Neuroanatomy, and Systems Neuroscience. Elsevier; 2013.

Trannoy S, Redt-Clouet C, Dura JM, Preat T. Parallel processing of appetitive short- and long-term memories in Drosophila. Curr Biol. 2011; 21:1647–1653. [PubMed: 21962716]

Tully T, Preat T, Boynton SC, Del Vecchio M. Genetic dissection of consolidated memory in Drosophila. Cell. 1994; 79:35–47. [PubMed: 7923375]

Turner GC, Bazhenov M, Laurent G. Olfactory representations by Drosophila mushroom body neurons. J Neurophysiol. 2008; 99:734–746. [PubMed: 18094099]

Usoskin D, Furlan A, Islam S, Abdo H, Lonnerberg P, Lou D, Hjerling-Leffler J, Haeggstrom J, Kharchenko O, Kharchenko PV, et al. Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. Nat Neurosci. 2015; 18:145–153. [PubMed: 25420068]

Vasiliauskas D, Mazzoni EO, Sprecher SG, Brodetskiy K, Johnston RJ Jr, Lidder P, Vogt N, Celik A, Desplan C. Feedback from rhodopsin controls rhodopsin exclusion in Drosophila photoreceptors. Nature. 2011; 479:108–112. [PubMed: 21983964]

Wada Y, Okano T, Adachi A, Ebihara S, Fukada Y. Identification of rhodopsin in the pigeon deep brain. FEBS Lett. 1998; 424:53–56. [PubMed: 9537514]

Walkinshaw E, Gai Y, Farkas C, Richter D, Nicholas E, Keleman K, Davis RL. Identification of genes that promote or inhibit olfactory memory formation in Drosophila. Genetics. 2015; 199:1173–1182. [PubMed: 25644700]

Wang T, Montell C. Rhodopsin formation in Drosophila is dependent on the PINTA retinoid-binding protein. J Neurosci. 2005; 25:5187–5194. [PubMed: 15917458]

Xia S, Miyashita T, Fu TF, Lin WY, Wu CL, Pyzocha L, Lin IR, Saitoe M, Tully T, Chiang AS. NMDA receptors mediate olfactory learning and memory in Drosophila. Curr Biol. 2005; 15:603–615. [PubMed: 15823532]

Yin JC, Del Vecchio M, Zhou H, Tully T. CREB as a memory modulator: induced expression of a dCREB2 activator isoform enhances long-term memory in Drosophila. Cell. 1995; 81:107–115. [PubMed: 7720066]

Yu D, Akalal DB, Davis RL. Drosophila alpha/beta mushroom body neurons form a branch-specific, long-term cellular memory trace after spaced olfactory conditioning. Neuron. 2006; 52:845–855. [PubMed: 17145505]

Yu D, Keene AC, Srivatsan A, Waddell S, Davis RL. Drosophila DPM neurons form a delayed and branch-specific memory trace after olfactory classical conditioning. Cell. 2005; 123:945–957. [PubMed: 16325586]

Yuan Q, Joiner WJ, Sehgal A. A sleep-promoting role for the Drosophila serotonin receptor 1A. Curr Biol. 2006; 16:1051–1062. [PubMed: 16753559]

Zars T. Visualizing PKA dynamics in a learning center. Neuron. 2010; 65:442–444. [PubMed: 20188649]

Zeisel A, Munoz-Manchado AB, Codeluppi S, Lonnerberg P, La Manno G, Jureus A, Marques S, Munguba H, He L, Betsholtz C, et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science. 2015; 347:1138–1142. [PubMed: 25700174]

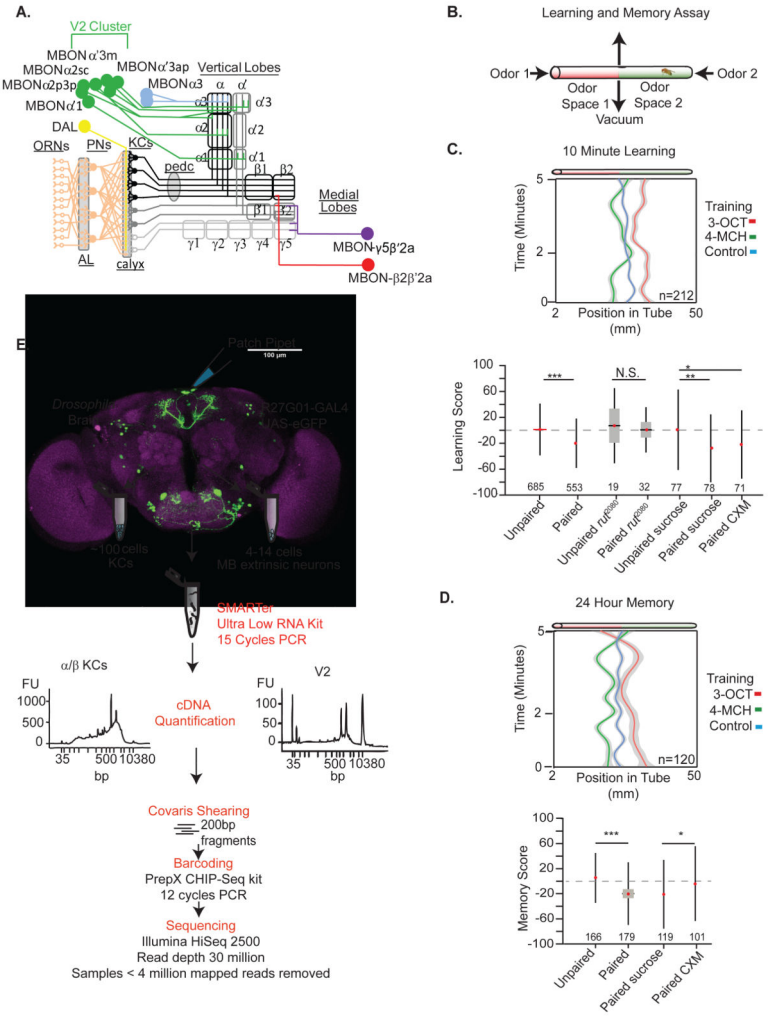


Figure 1. Combining single-fly learning and memory assays with cell type-specific RNAseq

A The wiring diagram of the olfactory pathway to the MB (afollowing the schematic of Aso et al. (eLife 2014a)). The mushroom body cell types we focused on in this study are cartooned; they each innervate distinct compartments of the MB calyx or lobes. Kenyon cells (KCs) are defined by the lobes their axons innervate. This color code is used throughout the paper to identify each cell type. Black=α/β KCs, dark gray=α′/β′ KCs, light gray=γ KCs, yellow=DAL, green=V2 cluster, blue=MBONα3, purple=MBON-γ5β′2a, red=MBON-β2β′2a.

B Schematic of the learning and memory assay. Individual flies are loaded into 53mm length tubes with odor inlets on each end and a vacuum port in the center. Each tube is inserted into a DAM (Drosophila Activity Monitor) device with 17 infrared beams spanning the length of the tube to track fly position.

C Population trajectory for one genotype (*R71D08-GAL4*; V2) (top) and learning scores (bottom) taken ten minutes after training. Trajectory data reveal the position of flies over a five minute period (mean and SEM are shown). For learning scores, mean (red dot), SEM (box) and standard deviation are plotted for all flies of 7 different genotypes (***p<0.0001). *rutabaga* mutants (*rut*^{2080}) do not learn. Flies (*R71D08, UAS-eGFP*) fed cyclohexamide

(versus only sucrose) showed normal learning (*p<.05, **p<.01). Numbers of animals in each experiment are indicated.

D Population trajectory for one genotype (*R71D08-GAL4*)(V2) (top) and memory scores (bottom) taken 24 hours after training. Memory scores are plotted for all flies of 7 genotypes (***p<0.0001). Flies (*R71D08, UAS-eGFP*) fed cyclohexamide (versus only sucrose) failed to form a long-term memory (*p<.05). Numbers of animals in each experiment are indicated.

E Protocol for cell harvesting and mRNA amplification. Cells were harvested using patch clamp pipets, pooled, and mRNA was amplified and sequenced (see Experimental Procedures). Quantification of cDNA (mRNA amplification generates cDNA) is shown for one KC sample (100 α/β KCs from one fly pooled) and one V2 sample (14 cells from 1 fly pooled).

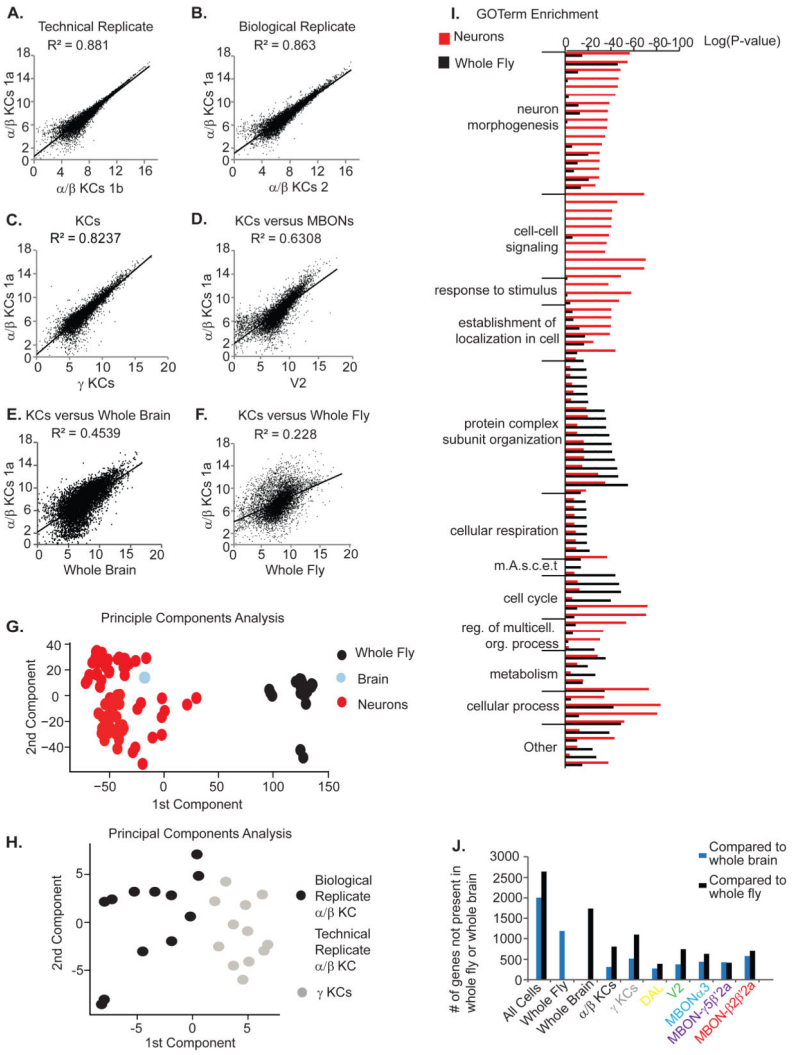


Figure 2. Reliability of single cell-type RNAseq and detection of neuron-specific transcripts

Correlation of gene counts across samples; HTseq-counts were used to determine gene counts and DESeq2 was used for normalization. Data are plotted as log normalized transcript counts. Count correlations (R^2) for technical replicates (**A**), biological replicates (**B**), related cell types (**C**), different cell types (**D**), and for one neuronal sample versus a whole brain sample (**E**) or a whole fly sample (**F**). Data from the same α/β KC sample is on the vertical axis of each plot. The fly genotype is the same for the neuronal and whole fly sample in **F**. 7234 genes were used for correlations, following removal of all genes with less than 2 CPM in the α/β KC sample.

G Principal components analysis (PCA) showing clustering of 27 different samples from whole flies of the 7 GAL4/GFP genotypes used in this study (black), 71 samples of individual neuronal cell types (red), and 2 whole brain samples from flies containing the UAS-eGFP transgene only (blue). The same 7234 genes in **A–F** were used for PCA.

H PCA of α/β KC samples (black) versus γ KC samples (gray), including the α/β technical replicate (white). Again based on the same 7234 genes as above.

I For high expressing genes in the 71 neuronal samples (red) and 27 whole fly samples (black), the top 50 Gene Ontology (GO) enrichment terms for biological processes. Adjusted p-values for enrichment are plotted (Holm-Bonferroni $p<0.05$). Abbreviated names are: m.A.s.c.e.t = mitochondrial ATP synthesis coupled electron transport; reg. of multicell. org. proc.= regulation of multicellular organismal processes. Full list of GOterms is found in Table S2.

J Comparison of the number of genes that are expressed above a 50 CPM mapped read cutoff in each neuronal sample compared to either whole brain samples (blue) or whole fly (black) samples.

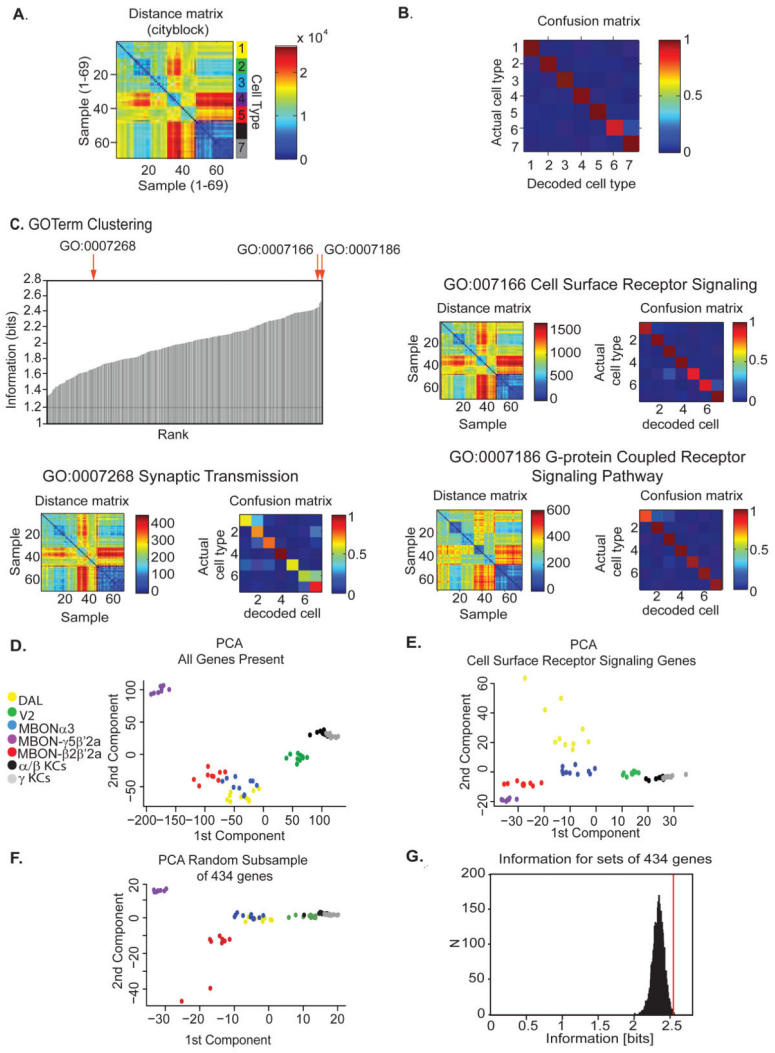


Figure 3. Classification of cell types by transcriptional profiling

A Pairwise distance matrix for each sample using all genes after removal of counts less than 2 CPM (8345 genes). Order of samples is: Yellow (1)= DAL, Green (2) = V2 Cluster, Blue (3) = MBONα3, Purple (4) = MBON-γ5β′2a, Red (5)= MBON-β2β′2a, Black (6)= α/β KCs, Grey (7)= γ KCs. **B** A confusion matrix was constructed (see Experimental Procedures) with each number representing a cell type shown in the distance matrix. A value of 1 means the classifier properly grouped all samples of a given cell type.

C Information values (see Experimental Procedures) for clustering by cell type based on gene lists from the 476 GO terms overrepresented in the neuronal samples. A distance matrix and confusion matrix are shown for genes from the three indicated GO term lists. The Cell Surface Receptor Signaling GO term (GO:0007166) contains 434 genes.

D PCA based on all 8345 genes after removing genes with counts < 2CPM, for all 69 neuronal samples. Loadings for each gene are shown in Table S5A.

E PCA on all 69 neuronal samples and based on only the 434 genes contained in Cell Surface Receptor Signaling GO term, after removing <2 CPM. Loadings for each gene are shown in Supplemental Table 5B.

F PCA on a random sub-sample of 434 genes from the full (8345) gene list. PCA based clustering of whole fly samples based on the list of Cell surface receptor signaling genes is shown in Figure S4.

G Information values (see Methods) for clustering by cell type based on genes in the Cell Surface receptor signaling gene list (red line) compared with 5,000 lists, each of 434 genes pulled at random.

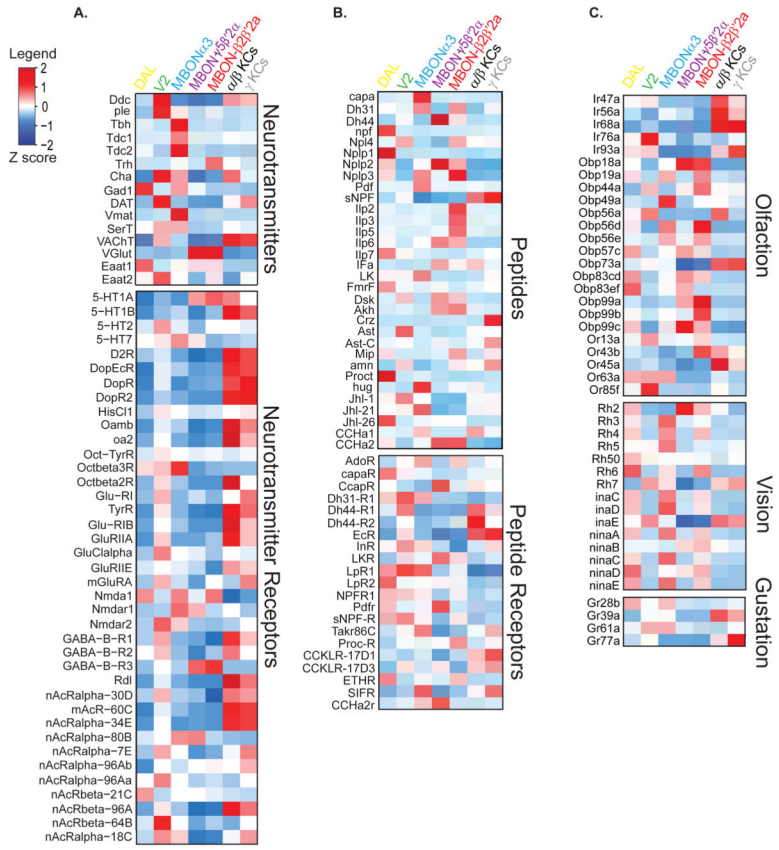


Figure 4. Analysis of gene expression by cell type

Normalized gene expression levels for each cell type (averaged across all samples of that type) then standardized across genes to highlight relative expression. See also Fig. S5 for a map of the raw counts (not z-scored) for each neuronal sample.

A The relative expression of neurotransmitter transporters and synthesis enzymes (top) and neurotransmitter receptors (bottom) for each of the 7 MB cell types.

B Peptide and peptide receptor expression for each of the 7 MB cell types.

C Expression of various sensory transduction genes, grouped by the peripheral sensory organ they are also expressed in; each gene was expressed in at least one of the cell types.

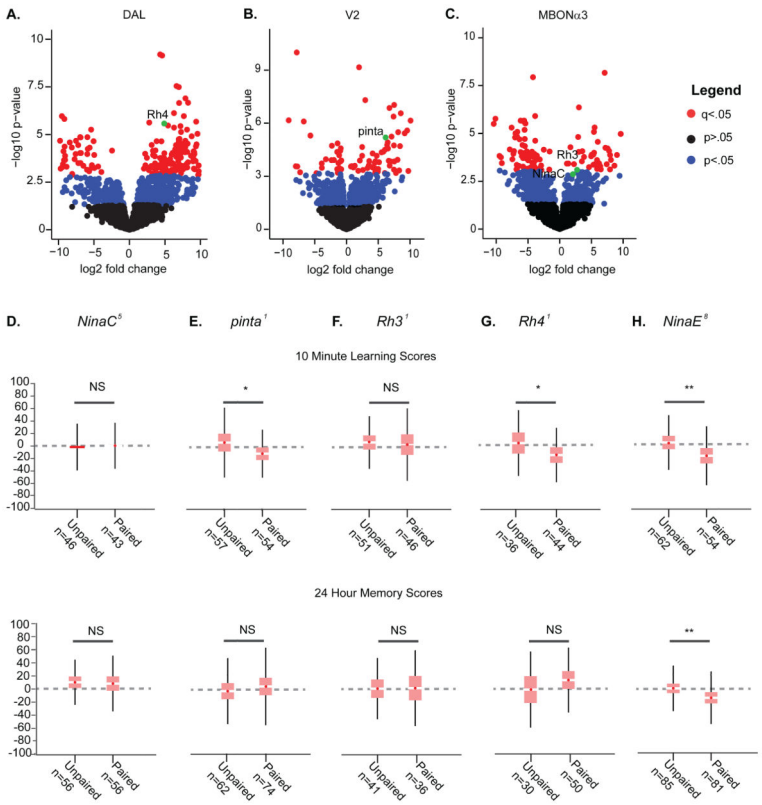


Figure 5. Learning and Memory-Related Changes in Gene Expression

Differential gene expression following memory formation for 3 MB extrinsic neuron cell types (data from other cell types is plotted in Figure S6). Red indicates a q value < 0.05, blue p < 0.05, and black denotes genes with no significant change in expression as determined by DESeq2.

A DAL neuron contains 174 differentially-expressed genes with q-value <0.05.

B The V2 neurons contain 83 differentially-expressed genes with q-value <0.05.

C The MBONα3 neuron contains 129 differentially-expressed genes with q-value < 0.05. The *NinaC* gene had a q value of <0.081and the *Rh3* gene had a q value <0.063.

D–H Learning and memory scores for five phototransduction mutants. Mean, S.E.M and S.D. are shown, as in Figure 1D. Student's t-test results *p<0.05, **p<0.01.

Table 1
Genes previously implicated in memory formation

Genes that are up-regulated following long-term memory formation (q-value<0.1) in DAL, V2 or MBONα3 neurons that were also previously implicated in long-term memory formation and/or retrieval in other studies.

***Drosophila melanogaster* Gene name**	**CG Number**	**Mouse Homolog**	**Reference**
CG10585 CG17221 wds Teh3 Taf6 CG34402 Akt1 Rbpn-5 Hph CG4743 CG5104 f-cup Brf	CG10585 CG17221 CG17437 CG18676 CG32211 CG34402 CG4006 CG4030 CG44015 CG4743 CG5104 CG9611 CG31256	Pdss2 Rtn4ip1 Wdr5 NA Taf6 NA Akt1, Akt2, Akt3 Rabep1 Egln1, Egln3 Slc25a26 Sft2d2 Lrrc40 Brf1	Walkinshaw et al. (2015) Genetics *199*
RhoGEF2 lbm CG34340	CG9635 CG2374 CG34340	Arhgef1 NA Prrxl1	Lakhina et al. (2015) Neuron *85(2)*
Zasp52 gp210	CG30084 CG7897	Ldb3 Gp210 (Human)	Dubnau et al. (2003) Curr. Bio. *13*
Ube3a	CG6190	Ube3a	Jiang et al. (2010) PlosOne *5(8)* Chakraborty et al. (2015) Biochem Biophys Res Commun *462(1)*
Actn	CG4376	Actn1, Actn2, Actn3, Actn4	Fifkova and Delay (1982) J Cell Biol *95(1)*
sNPF	CG13968		Knapek et al. (2013) J Neurosci *33(12)*

Cell Rep. Author manuscript; available in PMC 2016 October 03.

Table 2

GO term enrichment for genes up-regulated following memory formation

235 genes were analyzed using GOTerm Finder (http://go.princeton.edu/cgi-bin/GOTermFinder) to look for biological process enrichment. q-value of <0.1 was used to generate the list.

Gene Ontology term	Cluster frequency	Genome frequency	Corrected P-value	FDR	Genes annotated to the term
cellular response to light stimulus	8 of 329 genes, 2.4%	40 of 16085 genes, 0.2%	0.00154	0.00%	ninaC, Rh4, TotC, Arr2, PIP82, TotA, Fbxl4, Rh3
single-organism metabolic process	72 of 329 genes, 21.9%	2032 of 16085 genes, 12.6%	0.00210	0.00%	pgant6, CG13667, dare, su(r), EndoGI, CG4849, Pdh, Kdm4A, CG3590, Tip60, Ipk2, fh, Eogt, Cyp4ac1, CREG, w, Akt1, sNPF, CG5966, Ide, GIIIspla2,CG5009, CG14512, CG1746, CG12077, lic, CG17221, MAGE, APC4, Pepck, CG42637, CG9743, cpb, CG10802, Mtpalpha, CG15547, Alg10, Gyc76C,MP1, pinta, beta3GalTII, CG10268, CG1673, l(3)72Dp, CG1516, foxo, Cyp6a23, CG8112, Irc, Cap-D3, Ark, Jheh2, CG8525, beta4GalNAcTA, CG4842,wds, Bub3, Dhod, CG3714, asparagine-synthetase, CG33123, sl, Sgf29, Gpi1, Cyp4d2, CR31541, Aats-lys, rad50, Cyp4e3, CG18547, CG42863,CG17026
cellular response to UV	5 of 329 genes, 1.5%	12 of 16085 genes, 0.1%	0.00308	0.00%	ninaC, Rh4, TotA, TotC, Rh3
cellular metabolic process	133 of 329 genes, 40.4%	4651 of 16085 genes, 28.9%	0.00520	0.00%	l(3)72Do, EndoGI, CG4849, Kdm4A, Tip60, dgrn, Ipk2, del, Eogt, Es2, twe, CG1749, GIIIspla2, CG14512, CG11906, Pcl, Su(fu), MAGE, APC4, Optix,SmD1, CG10802, Tis11, CG5033, Alg10, msl-2, Nf-YC, pinta, m-cup, CG10268, mRpS9, Fbxl4, l(3)72Dp, foxo, RpLP1, CG8112, Brf, Psc, Jheh2, Ark,CG5525, hop, wds, Bub3, CG7849, Dhod, CG3714, asparagine-synthetase, RpLP2, HIP-R, crm, csul, CG10341, asf1, sl, ninaC, Gpi1, CG16941, rad50,CG42863, CG30005, CG17026, CstF-64, danr, Taf6, dare, su(r), CG34340, Pdh, Zpr1, RpL18, CG3590, Rae1, Cyp4ac1, w, CREG, CG9596, ci, Akt1,mRpL37, sNPF, Ide, CG5009, CG1746, CG6227, CG12077, lic, Rpb8, hpo, GstE9, CG12717, CG42637, eap, CG11859, Sox102F, cpb, Mtpalpha,Lsm10, Gyc76C, CG15547, CG18596, Elp2, MP1, gek, beta3GalTII, Sirt6, CG1673, Upf3, CG1516, Ube3a, NELF-B, MED16, l(2)37Cg, Cap-D3,CG8525, zfh1, CG4935, CG6662, ey, beta4GalNAcTA, St3, CG12746, CG10376, toy, hbs, CG3756, CG33123, Sgf29, gt, Neos, CR31541, Aats-lys, bt
nervous system development	62 of 329 genes, 18.8%	1711 of 16085 genes, 10.6%	0.00650	0.00%	CstF-64, Nf-YC, babos, lea, RanGap, danr, lbm, gek, Taf6, sec15, CG9004, CG34340, EndoGI, Rae1, Ube3a, Tip60, foxo, dgrn, olf186-F, Brf, Psc, Ark,mbo, CG10257, Slh, zfh1, ci, hop, CG5525, Cpn, Akt1, ey, if, prel, fy, CG9766, wds, CG4842, Pcl, CG12077, toy, Phax, hbs, synaptogyrin, l(2)k14710,CG17221, CG33123, MAGE, hpo, CG10341, asf1, Optix, gt, CG12717, spin, CG7338, CG11859, CG12259, CG4806, cpb, CG5033, Gyc76C
metabolic process	161 of 329 genes, 48.9%	5963 of 16085 genes, 37.1%	0.00756	0.00%	CG13667, CG31704, l(3)72Do, Gasp, EndoGI, CG4849, Kdm4A, Tip60, dgrn, Ipk2, del, fh, Eogt, Es2, twe, CG5966, CG1749, CG7791, GIIIspla2,CG14512, CG11906, Pcl, CG3534, Su(fu), MAGE, APC4, Pepck, CG15738, Optix, CG31445, SmD1, CG9743, Idgf3, CG10802, Tis11, CG5033, Alg10,CG9581, msl-2, Nf-YC, RanGap, pinta, m-cup, CG10268, CG13807, mRpS9, Fbxl4, l(3)72Dp, foxo, RpLP1, Cyp6a23, CG8112, Brf, Psc, CG40486, Irc,Jheh2, Ark, CG5355,

Gene Ontology term	Cluster frequency	Genome frequency	Corrected P-value	FDR	Genes annotated to the term
					CG5525, hop, CG4842, wds, Bub3, CG7849, Dhod, CG3714, asparagine-synthetase, HIP-R, RpLP2, crm, csul, CG10341, asf1, sl,CG8207, ninaC, Gpi1, CG16941, Cyp4e3, rad50, CG42863, CG30005, CG17026, pgant6, CstF-64, danr, Taf6, dare, CG34340, su(r), Pdh, Zpr1, RpL18,CG3590, Rae1, Cyp4ac1, w, CREG, ci, CG9596, Akt1, mRpL37, sNPF, Ide, CG5009, CG1746, CG6227, CG12077, lic, Rpb8, CG17221, hpo, GstE9,CG12717, CG42637, eap, CG11859, Sox102F, cpb, Lsm10, Mtpalpha, CG15547, ECSIT, Gyc76C, CG18596, Elp2, MP1, gek, beta3GalTII, Sirt6,CG1673, Upf3, CG1516, Ube3a, CG15661, NELF-B, MED16, Cap-D3, l(2)37Cg, CG8525, zfh1, CG4935, CG6662, ey, beta4GalNAcTA, St3, CG12746,CG10376, toy, hbs, CG3756, CG33123, Sgf29, gt, Cyp4d2, Neos, CR31541, Aats-lys, bt, CG18547
response to radiation	15 of 329 genes, 4.6%	186 of 16085 genes, 1.2%	0.00848	0.00%	ninaC, Rh4, Ttc19, pinta, TotC, sec15, Ark, Arr2, dare, PIP82, hpo, TotA, Fbxl4, cdm, Rh3
cellular response to abiotic stimulus	9 of 329 genes, 2.7%	67 of 16085 genes, 0.4%	0.01066	0.00%	ninaC, Rh4, TotC, Arr2, PIP82, TotA, CG1732, Fbxl4, Rh3
response to light stimulus	13 of 329 genes, 4.0%	151 of 16085 genes, 0.9%	0.01735	0.00%	ninaC, Rh4, Ttc19, pinta, TotC, sec15, Arr2, dare, PIP82, TotA, Fbxl4, cdm, Rh3
cellular response to radiation	8 of 329 genes, 2.4%	57 of 16085 genes, 0.4%	0.02447	0.00%	ninaC, Rh4, TotC, Arr2, PIP82, TotA, Fbxl4, Rh3
neurogenesis	53 of 329 genes, 16.1%	1469 of 16085 genes, 9.1%	0.04023	0.00%	CstF-64, Nf-YC, babos, lea, RanGap, gek, Taf6, sec15, CG9004, CG34340, EndoGI, Ube3a, foxo, dgrn, Brf, Psc, Ark, mbo, CG10257, Slh, zfh1, ci, hop,CG5525, Cpn, Akt1, ey, if, prel, fy, CG9766, wds, CG4842, Pcl, CG12077, hbs, synaptogyrin, l(2)k14710, hpo, MAGE, CG33123, CG10341, asf1, Optix,gt, spin, CG12717, CG7338, CG11859, CG12259, CG4806, CG5033, Gyc76C
regulation of cellular component size	11 of 329 genes, 3.3%	119 of 16085 genes, 0.7%	0.04125	0.00%	AnnIX, gt, gek, lic, Lst8, Akt1, cpb, sNPF, CG4030, foxo, CG42863

Cell Rep. Author manuscript; available in PMC 2016 October 03.