A State-Dependent Memory for Sodium Chloride in *Drosophila melanogaster* and Comparative Neural Transcriptomic Analyses of *Drosophila* and *Apis mellifera*

Isabel C Morgan

Green Templeton College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2025

Acknowledgements

First, I would like to thank my supervisors, Scott Waddell and David Sims, for their support and guidance throughout various stages of my DPhil. Scott, thank you for allowing me to explore my research interests and for integrating me so well into the Waddell group. It was a privilege to learn how to become a true scientist within your lab, and I was grateful for the times we discussed science. David, I am so appreciative of your time and patience in helping me learn computational biology, and for taking me on as an additional DPhil student in your group. Your wisdom and mentorship helped me gain confidence in my computational ability, and your encouragement to develop transferable skills has prepared me well for the next stage of my professional career. Although the issues with my health have been unfortunate, I could not have been more fortunate to be fully immersed in two fields of research. I have learnt more than I could have hoped for during my DPhil. So, thank you.

To the Waddell group, thank you for the kindness, support and friendship. Special thanks to Bhagyashree, for being an excellent team mate. I always admired your incredible knowledge of the literature and your ambitious nature. To Charly, for sharing your knowledge, for helping shape the single-cell RNA sequencing experiments and for collecting and processing the samples. To the Sims group, thank you so much for taking me under your wing and helping me grow into a more confident computational scientist. I am grateful to have attended the OBDS course and to have had the continued support from the Sims group. Lena, I am thankful for the time that you dedicated to my projects and for helping me grasp a general understanding of the principles behind the computational tools I was using.

To my friends and family, I could not be more lucky. Thank you for being with me on this journey. To my parents, Terry and Claire, to my sister Jess, her partner Jack and their dog (for emotional support) Dina. I couldn't put into words how appreciative I have been to have your support and kindness over the last 4 years. To my parents, for always teaching me that there is no task too large and just to "eat the elephant with the teaspoon", and to Jess and Jack who have always been around for a pep talk and an adventure. To my friends, Ellen, Chloe, Jenny, Claire, Katherine, Brittany, the Continuity family and the climbing community. You bring me so much joy and remind me that there is more to life than keeping my head in my books.

Abstract

Associative learning and memory, combined with internal state signals, optimise the behaviours of foraging animals. The Mushroom Bodies (MBs) of insects coordinate associative learning and are comprised of intrinsic neurones called Kenyon Cells (KCs). KCs represent sensory information in a parallel stream. Valence information, like rewarding nutrient signals, is provided to the MB by dopaminergic neurones (DANs). DANs direct learning-relevant plasticity of connections between KCs and downstream MB output neurones. This thesis investigates a state-dependent NaCl-memory in *Drosophila melanogaster* and explores midbrain transcriptomic changes in response to NaCl-feeding. In addition, transcriptomic profiles of *Apis mellifera* KCs and DANs were characterised and compared to those of *Drosophila* to understand conserved transcriptomic cell profiles between the two species.

Using an olfactory associative learning paradigm, with a naively appetitive (low) or naively aversive (high) concentration of NaCl as the unconditioned stimulus, we tested the existence of a state-dependent NaCl memory in *Drosophila*. Nutrient-deprived *Drosophila* form an immediate, but short-lived, low NaCl memory. A long-lasting memory was found for high NaCl; the valence of this memory depended on nutrient status. Nutrient-deprived *Drosophila* approached the conditioned odour, NaCl-fed *Drosophila* showed no detectable preference for either the conditioned nor alternative odour and satiated *Drosophila* avoided the conditioned odour. This aversive behaviour was more pronounced in females. To understand the transcriptional changes that occur in the brain in response to NaCl feeding, we analysed single-cell RNA sequencing (scRNA-seq) data of nutrient-deprived or NaCl-fed *Drosophila* midbrains. However, the quality of this data was impacted by a high level of adapter sequence contamination in sequenced reads, but preliminary analysis identified potential nutrient-deficit signals.

ScRNA-seq analysis of twelve pollen-foraging *Apis* dorsal protocerebra identified nine KC clusters. Transcriptomic comparison of *Drosophila* and *Apis* KC clusters found little similarity; however, marker genes for *Drosophila* reward and punishment-encoding DANs annotated two discrete *Apis* DAN clusters.

Contents

List of Abbreviations

AL . . . . . Antennal Lobe.
CS . . . . . Conditioned Stimulus.
DANs . . . . . Dopaminergic Neurones.
DGE . . . . . Differential Gene Expression.
GtACR1 . . . . *Guillardia theta* Anion Channelrhodopsin 1.
INSO . . . . . Internal Sodium-sensing Neurones.
KC . . . . . Kenyon Cell.
MAD . . . . . Median Absolute Deviation.
MB . . . . . Mushroom Body.
MBONs . . . . Mushroom Body Output Neurones.
NP . . . . . Neuropeptide.
NPR . . . . . Neuropeptide Receptor.
NT . . . . . Neurotransmitter.
NTR . . . . . Neurotransmitter Receptor.
OL . . . . . Optic Lobe.
OPNs . . . . . Olfactory Projection Neurones.
PAM . . . . . Protocerebral Anterior Medial.
PCA . . . . . Principle Component Analysis.
PER . . . . . Proboscis Extension Response.
PI . . . . . Performance Index.
PPL . . . . . Protocerebral Posterior Lateral.
SAG . . . . . Sex peptide Abdominal Ganglion.
scRNA-seq . . Single-cell RNA sequencing.
SER . . . . . Sting Extension Reflex.
SEZ . . . . . Subesophageal Zone.
SPR . . . . . Sex Peptide Receptor.
SPSNs . . . . . Sex Peptide Sensory Neurones.
UMAP . . . . . Uniform Manifold Approximation and Projection.
UMI . . . . . Unique Molecular Identifier.
US . . . . . Unconditioned Stimulus.

Chapter 1

Introduction

Contents

1.1 Motivation

Obtaining sufficient nutrition from one's environment is essential for normal organismal function. Memories of sensory cues associated with essential nutrition or harmful substances help animals optimise foraging behaviours. Examples of nutrition-related associative learning have been shown across insect species. For instance, the *Heliconius* butterfly can use visual colour cues to guide feeding behaviours towards rewarding nutrient sources (Young et al. 2024), locusts can learn to seek out odours previously paired with food rewards (Simões et al. 2011), and honey bees will reduce reward-seeking behaviours to a previously rewarding sucrose-associated cue when paired with malaise (Wright et al. 2010). Social insects such as the honey bee, *Apis mellifera*, exhibit more complex behaviours compared to non-social insects, such as the fruit fly *Drosophila melanogaster*. Social insects use both private and social information to guide foraging behaviours (Grüter et al. 2011), to meet both personal and colony nutritional requirements. The specialised brain structures, called the Mushroom Bodies (MBs), coordinate associative learning in insects (Heisenberg 1998). The MB circuitry has been most extensively studied in *Drosophila* due to the wealth of genetic tools and connectomic information

that has allowed scientists to probe the learning and memory circuits (Owald et al. 2015b; Scheffer et al. 2020; Li et al. 2020a).

This thesis outlines the research contributing to the scientific literature by the identification of a novel nutrient-memory to sodium chloride in *Drosophila*, the transcriptomic characterisation of cells associated with learning and memory in *Apis mellifera*, and performing a comparative analysis of these cells to those of *Drosophila melanogaster*.

1.2 The Circuits Involved in Learning and Memory

Associative olfactory learning protocols in insects have been indispensable for understanding mechanisms for learning and memory. Protocols for associative olfactory learning involve pairing a novel odour (conditioned stimulus, CS) with a rewarding (such as sweet sugars) or punishing (such as bitter tastes) unconditioned stimulus (US) (Krashes et al. 2008; Huetteroth et al. 2015; Burke et al. 2012; Das et al. 2014; Quinn et al. 1974). Memory is tested by recording behavioural performance after presentation of the CS odour without the US. In *Drosophila*, associative learning is often carried out using a T-maze where groups of flies ($\sim$100) are exposed to an unpaired novel odour (CS-), then exposed to a US paired odour (CS+). The testing phase of the protocol involves presenting the flies with a binary choice of the CS+ and CS- odours without the US. Flies that have an appetitive memory will approach the CS+ odour, but flies with an aversive memory will avoid the CS+ odour (Tempel et al. 1983; Krashes et al. 2008; Krashes et al. 2009; Tully et al. 1985; Krashes et al. 2011a; Krashes et al. 2011b). Flies that show no memory performance will be distributed randomly between the CS+ and CS- odours. In *Apis*, the proboscis extension response (PER) assay is commonly used to test associative olfactory learning (Takeda 1961; Bitterman et al. 1983). The training phase of the PER assay starts with exposing a restrained *Apis* to a CS odour paired with the US (like sucrose) delivered to the antenna, which causes the proboscis to extend, and then to the proboscis. The testing phase includes exposing the *Apis* to CS odour and measuring the PER; in untrained *Apis*, odour exposure does not elicit PER (Bitterman et al. 1983). Both the T-maze assay with *Drosophila* and the PER assay with *Apis* can

be used to test memory persistence (Hammer et al. 1995; Wright et al. 2007; Eisenhardt 2006).

Associative learning experiments, such as the T-maze and PER assays, have been used to identify cells that coordinate associative learning. For example, cooling of the MB in *Apis* and experiments using *Drosophila* MB mutants revealed the importance of the MBs in associative learning and memory in these insects (Erber et al. 1980; Heisenberg et al. 1985). In *Drosophila*, dopaminergic neurones (DANs) provide the teaching signals to reinforce appetitive and aversive memories. Replacing the US with optogenetic activation of DAN clusters, like the protocerebral posterior lateral 1 (PPL1) or protocerebral anterior medial (PAM) DAN clusters, during associative olfactory training is sufficient to create an aversive or appetitive memory, respectively (Lin et al. 2014; Liu et al. 2012; Burke et al. 2012; Yamagata et al. 2015; Huetteroth et al. 2015; Claridge-Chang et al. 2009; Aso et al. 2012). In *Apis*, evidence suggests that octopaminergic signalling is required for appetitive memories, but dopaminergic signalling is required for aversive memories. VUMmx1 is an octopaminergic neurone that innervates the *Apis* MB and depolarisation of the VUMmx1 neurone is sufficient to act as a US to reinforce appetitive memories (Hammer 1993). Aversive learning in *Apis* is coordinated by dopaminergic signalling. Dopaminergic antagonists impaired the sting extension response that occurs in *Apis* when presenting the CS odour that was previously paired with electric shocks (Vergoz et al. 2007). Given that the comparative analysis between *Apis* and *Drosophila* in this report was focused towards the intrinsic cells of the MB and DANs, I will focus the rest of the introduction on these cell types.

1.2.1 The Mushroom Bodies

The MBs of insects are a pair of dorsal protocerebral neuropil structures important for integrating sensory information for associative learning. The intrinsic neurones of the MB are called Kenyon Cells (KCs); these cells receive sensory information from projection neurones, mostly olfactory information in *Drosophila* from the antennal lobes,

but also some visual inputs from the optic lobes (Li et al. 2020a; Ganguly et al. 2024; Li et al. 2020b; Vogt et al. 2016). The *Apis* KCs receive olfactory information from the antennal lobes via olfactory projection neurones (OPNs), which travel via the medial antenno-glomerular tract or the lateral antenno-glomerular tract to innervate the MB calyces (Malun 1998). The bee MB also receive visual information from the optic lobes' medulla and lobula that innervate the MB calyces' collar region (Malun 1998).

The MBs of *Apis* vary in size, morphology and proposed cell types compared to *Drosophila*. There are about 2000 KCs per hemisphere in *Drosophila* compared to about 170,000 KCs per hemisphere in *Apis* brains, making up about 4% and 34% of the total neurones in the brain, respectively (Witthöft 1967; Aso et al. 2014; Schatton et al. 2017). The enlarged MB structures in *Apis* are likely a reflection of the greater need to perform more complex tasks as eusocial insects (O'Donnell et al. 2004).

1.2.2 The Kenyon Cells

Drosophila melanogaster

The MBs of *Drosophila* comprise three main types of KCs and can be identified by their anatomical location. The cell bodies of the KCs are located posteriorly in the brain and the KC dendrites form the MB calyx. A structure called the peduncle carries the KC neurites to the MB lobes (as reviewed by Lin 2023, see Figure 1.1a). The γ KCs innervate the MB horizontal lobes, and the $\alpha\beta$ and the $\alpha'\beta'$ KCs innervate both the vertical and horizontal MB lobes (Aso et al., 2014; Crittenden et al. 1998, Figure 1.1a). The different KC populations have distinct transcriptomic profiles and form three separate clusters in single-cell RNA sequencing (scRNA-seq) data of *Drosophila* brains (Croset et al. 2018; Park et al. 2022). The KC cluster subtypes were identified using the expression of marker genes such as, *short neuropeptide F* (*sNPF*), *Ca-alpha1T* and *Eip93F* for α/β KCs, *trio* and *CG8641* for $\alpha'\beta'$ cells and *sNPF*, *trio* and *ab* for the γ KCs (Awasaki et al. 2000; Johard et al. 2008; Park et al. 2022; Shih et al. 2018). *sNPF* is required for appetitive olfactory learning in *Drosophila* (Knapek et al. 2013). *Ab* and

trio are involved in axonal development and KC temporal specification (Lin, 2023).

Apis mellifera

The MB of *Apis* consists of two cup-like structures, the medial and lateral calyces, which sit adjacent to one of two classes of KC bodies, Class I. Class II KC somata are positioned outside of the calyx, next to the basal ring and collar. The two peduncles carry the neurites from the KCs to the medial, vertical and γ lobes of the MB (Figure 1.1b, Malun 1998; Strausfeld 2002). I will refer to the MB lobes as described in Strausfeld, (2002). Initially, the *Apis* MB lobes were called α and β (Malun 1998; Mobbs et al. 1982), which comprise the vertical and medial lobes, respectively; Strausfeld, however, renamed the lobes to prevent confusion between α and β lobes of *Drosophila*. *Apis* Class I KC axons bifurcate to the vertical and medial lobes, whereas the Class II KCs extend axons to populate the γ lobes (Strausfeld 2002). The γ lobes were named as such due to their developmental and anatomical similarities to the *Drosophila* γ lobes.

KCs of *Apis* MB are split into five main groups based on anatomical and gene expression data (Suenami et al. 2018; Kubo et al. 2012). Class-I KCs divide into large-KCs (lKCs) with their cell bodies at the edge of the calyces, small-KCs (sKCs) and middle-KCs (mKCs) whose cell bodies are close to the centre of the MB calyces (Figure 1.1b). lKCs are the first to develop from the Class-I KCs (Suenami et al. 2016). The lKCs express the marker gene *mushroom body large-type Kenyon-cell specific protein-1* (*Mblk-1*), an ecdysone signalling-related transcription factor (Takeuchi et al. 2001; Matsumura et al. 2022). *Mblk-1* is suggested to target genes related to synaptic plasticity in the MBs (Matsumura et al. 2022). sKCs develop after the lKCs and can be identified by their expression of *E74*, a transcription factor also involved in the ecdysteroid-signalling pathway (Suenami et al. 2016; Paul et al. 2005). mKCs are the last to develop from the Class-I KCs (Suenami et al. 2016). They were identified by their expression of *middle-type KC-preferential arrestin-related protein* (*mKast*). *mKast* is similar to mammalian ARRDCs (arrestin domain-containing protein) and is suggested to have a

role in neuronal receptor regulation (Kaneko et al. 2013). An additional KC subtype was identified with the expression of *FoxP* (FoxP-KCs). *FoxP* is a transcription factor that serves a function in decision-making in *Drosophila* (Das et al. 2014). *Apis* FoxP-KCs cell bodies are positioned adjacent to mKCs, evidenced by expression of *FoxP* and *mKast* in the *Apis* adult brain (Schatton et al. 2017). Class-II KC bodies are located on the MB calyx's outer surface (Fahrbach 2006); these cells do not have any specific marker genes. A scRNA-seq analysis comparing cells of the *Apis* brain between different castes (Queen, nurse and pollen-forager) found clusters associated with lKCs, sKCs and FoxP-KCs. However, this analysis did not identify mKC or Class-II KCs (Zhang et al. 2022). The authors found additional KC clusters that they left unannotated. The transcriptomic profiles of mKCs and class-II KCs remain incomplete.

1.2.3 The Dopaminergic Neurones

Drosophila melanogaster

As mentioned above, DANs are required for signalling value to reinforce memories during associative learning. Two major DAN clusters are involved in olfactory learning and memory: the PAM DANs, which innervate the horizontal lobes, the PPL1 DANs, which innervate the vertical lobe, the peduncle and the site where the horizontal and vertical lobes diverge (Waddell 2013). Per hemisphere, there are ∼100 PAM neurones and 12 PPL1 neurones (Mao et al. 2009). Subsets of DANs innervate discrete non-overlapping compartments of the MB (Aso et al., 2014a; Li et al. 2020a). The release of dopamine from specific DANs within the corresponding MB compartments is suggested to reduce the synaptic strength between the odour-activated KCs and the MB output neurones (MBONs) associated with this compartment. MBONs coordinate approach or avoidance behaviours; DANs associated with reward signalling share compartments with avoidance MBONs, therefore, during reward learning, dopamine release from reward-encoding DANs causes the depression between the odour activated KCs and avoidance MBONs, which causes a skew in MBON output and leads to more approach

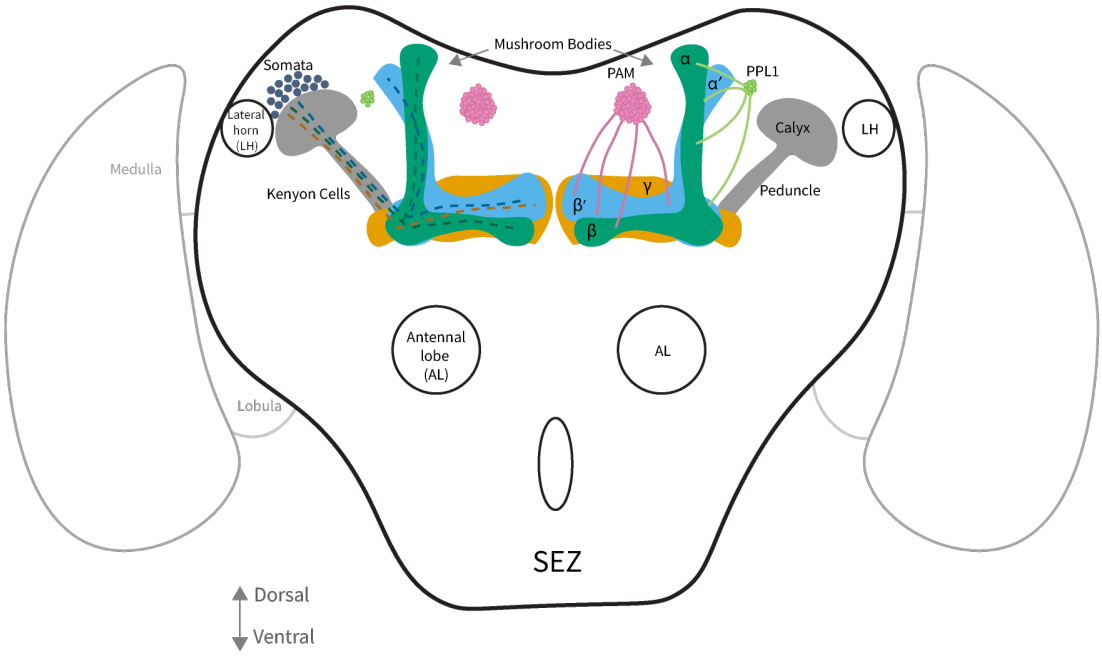
behaviours (Owald et al. 2015a; Perisse et al. 2016; Felsenberg 2021; Hige et al. 2015; Cohn et al. 2015; Handler et al. 2019).

Apis mellifera

There are four major clusters of dopaminergic cells in the *Apis* brain, C1-C4 (Tedjakumala et al. 2017; Schürmann et al. 1989; Schäfer et al. 1989, see Figure 1.1b). C1 - C3 dopaminergic clusters innervate the *Apis* MB vertical and medial lobes. Tedjakumala et al., (2017) suggest C1 and C2 clusters are homologous to the *Drosophila* PAM cluster because of their location and because they innervate the medial lobe of the MB. They also suggest that C3 is similar to *Drosophila* PPL clusters. MBONs have also been described in the *Apis* brain. Like *Drosophila*, dendrites from *Apis* MBONs innervate different sections of the MB lobes where KC terminals sit (Strausfeld 2002). Considering subsets of DANs innervate the MB lobes, it is possible that DAN signalling could modulate the KC-MBON connection to drive changes in MBON activity after associative learning (Strube-Bloss et al. 2011; Okada et al. 2007).

ScRNA-seq analysis of the *Apis* brain by Zhang et al., (2022) identified monoaminergic cells by the expression of *Vmat*. They found that these cells made up $\sim$10% of their data; however, *Vmat* was broadly expressed across their dataset and not in a specific cluster. *Drosophila* scRNA-seq data from Park et al., (2022) found distinct clusters for monoaminergic cells that were separate from the other cell types. Further analysis by Park et al., (2022) found additional granularity within the monoaminergic clusters based on the expression of different transmitter types (dopaminergic, serotonergic, tyraminergic, octopaminergic). The scRNA-seq data used in this report utilised cells from the dorsal protocerebrum from twelve pollen-foraging *Apis* honey bees to increase the representation of DAN and MB cell types to aid cell type identification and annotation. A transcriptomic comparison between KC and DAN cell types would be beneficial for understanding if similar cell types have similar functions between *Apis* and *Drosophila*.

(a) *Drosophila melanogaster*



(b) *Apis mellifera*

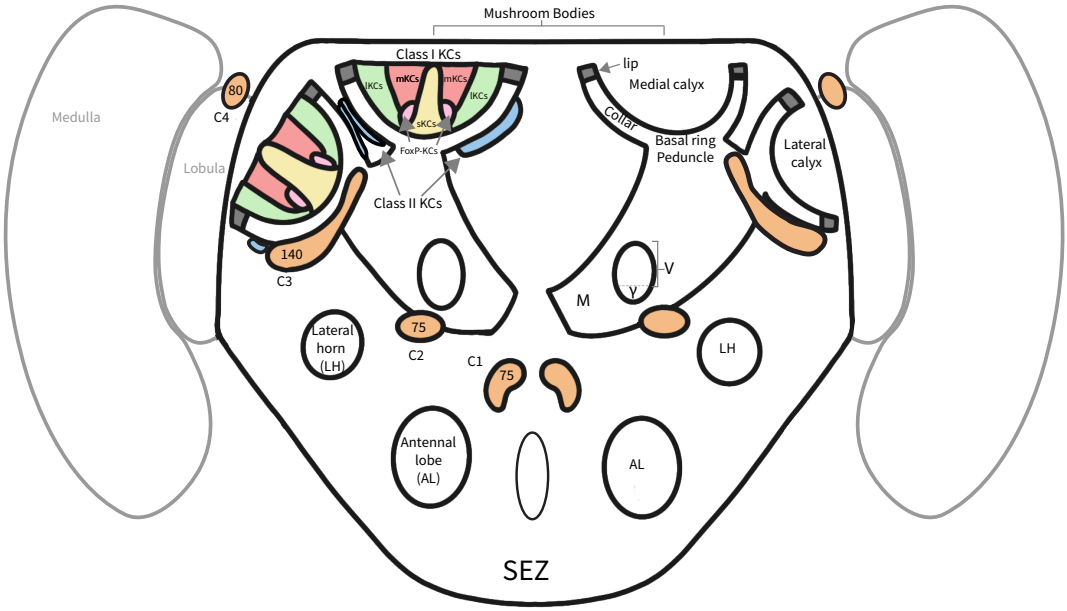


Figure 1.1: Schematic of the *Drosophila melanogaster* and *Apis mellifera* brain - with focus on the learning and memory relevant cells.
For both brains, *Drosophila melanogaster* (a) and *Apis mellifera* (b) have two neuropil structures that sit in the dorsal protocerebrum, called the mushroom bodies (MB). a. The *Drosophila* MB houses three main types of Kenyon Cells (KCs). The KC cell bodies are positioned posteriorly, with their dendrites forming the calyx and their axons bundled to form the peduncle. The $\alpha\beta$ and $\alpha'\beta'$ KCs bifurcate to innervate the α and β (green) and α' and β' lobes (blue), respectively. The γ KCs do not bifurcate; they send their projections medially to make the γ lobe (Aso et al. 2014). Two dopaminergic neurone (DAN) clusters are shown that innervate the MB; the Protocerebral Anterior Medial (PAM) DANs innervate the horizontal lobes, and Protocerebral Posterior Lateral (PPL) DANs, innervate the vertical lobe, peduncle and the region where vertical and horizonal lobes branch off (Waddell 2013). b. The *Apis* MBs comprise of two cup-like structures, the medial and the lateral calyces. Each calyx is divided into subregions: the lip, collar and basal ring. The Class I KC cell bodies are positioned in the calyces cup; Class I KCs project their axons via a peduncle to the medial (M) and vertical (V) of the MB. The cell bodies of the Class-I sKCs (yellow) and mKCs (red) are located towards the centre of the calyces cup, and the lKCs (green) are positioned to the outside of the cup. FoxP-KC cell bodies (pink) are positioned adjacent to mKCs. Class II KC cell bodies are found on the outer surface of the MB calyces (blue); Class II KCs project their axons to the γ lobe (Suenami et al. 2018; Kubo et al. 2012; Malun 1998; Strausfeld 2002). Four main clusters of dopaminergic neurones C1-C4 (orange). There are $\sim$75 dopaminergic cells in clusters C1 and C2, $\sim$140 in C3 and $\sim$80 in C4 (Tedjakumala et al. 2017).

1.3 Nutrient Preference and Learning

Nutrient preference

Animals require adequate consumption of macronutrients and micronutrients to survive. *Drosophila* will seek out nutrients during periods of deprivation or when nutritional demand increases after copulation (Corrales-Carvajal et al. 2016; Ribeiro et al. 2010; Walker et al. 2015). Honey bees and flies will self-select for sodium at different concentrations (Zhang et al. 2013; De Sousa et al. 2022), likely based on internal requirements and taste preference. For example, flies can discriminate between low and high NaCl concentrations, opting for solutions containing sucrose plus 0.05 M NaCl compared to sucrose alone. However, when the concentration of NaCl increased to 0.2 M or more, flies opted for the sucrose solution (Zhang et al. 2013). When NaCl deprived, *Drosophila* exhibited a greater PER to solutions with high NaCl concentrations compared to NaCl-fed *Drosophila* (Jaeger et al. 2018). Mated female *Drosophila* have a greater PER response to high NaCl concentrations compared to mated males and virgin females (Walker et al. 2015). The evidence above suggests that insects can discriminate and self-regulate NaCl to match internal needs.

Although NaCl learning has been reported in rats, crickets and *Drosophila* larvae (Cone et al. 2016; Robinson et al. 2013; Mizunami et al. 2010; Russell et al. 2011; Niewalda et al. 2008), there is currently no published evidence that NaCl learning occurs in adult *Drosophila*. This report explores NaCl learning and memory persistence at different concentrations in adult *Drosophila*, considering virgin and mated, males and females.

Nutrient Learning

Pathways for taste and nutrient memories, such as those of sweet and nutritious sugars, have been studied in the *Drosophila* brain. *Drosophila* will approach odours previously paired with sweet-tasting sugars, but will avoid odours previously paired with bitter tastes, like DEET (Das et al. 2014; Huetteroth et al. 2015; Burke et al. 2011).

Taste memories, like those associated with sweet-tasting, non-nutritious sugars, can be detected immediately, but are short-lived. Short-term memories of non-nutritious sweet-tasting sugars can be reinforced into long-term memories when combined with nutritious sugars without a sweet taste (Huetteroth et al. 2015; Burke et al. 2011; Musso et al. 2015). Taste valence and nutrient information are provided to the *Drosophila* MB by different subsets of DANs. Bitter tastes are signalled via the PPL DANs (Das et al. 2014). Sweet-taste and nutrient information is provided by different subsets of the PAM DANs. Blocking the output from a population of PAM-DANs, represented by the R48B04-GAL4 driver line, impaired immediate memory to sweet-tasting sugars, but not long-term memory of nutritious sugars. Blocking the output from a different subset of PAM-DANs, covered by the R15A04-GAL4 driver line, which labels $\beta 2$, $\beta' 1$, $\alpha 1$ and $\gamma 5$ neurones, impaired long-term memory to sucrose, but not short-term memory. Artificial activation of R48B04-GAL4 neurones, which covers $\beta' 2$, $\gamma 4$ and $\gamma 5$, was sufficient to create a short-term memory only, whereas activation of R15A04-GAL4 neurones implanted a long-term appetitive memory only. These results suggest that there are different subsets of neurones that supply information to the MB about taste and nutrient content of the US to reinforce different components of the memory (Huetteroth et al. 2015; Yamagata et al. 2015). The idea that there are different components to nutrient memory was further evidenced by Das et al., (2014), where the aversive tastant DEET was combined with a nutritious sugar during olfactory learning. *Drosophila* learned to avoid the conditioned odour in the short term, this short-lived aversive memory was replaced with an appetitive long term memory. This report aimed to understand if there could be a taste and a nutrient component to NaCl memory in adult *Drosophila*.

State-dependent memories in *Drosophila*

The internal state, such as thirst and hunger in *Drosophila*, influences learned behaviours (Lin et al. 2014; Senapati et al. 2019; Krashes et al. 2009; Dus et al. 2011). For example, Krashes et al., (2009) found that sugar memory performance was reduced the more satiated the animal was. Water memory is affected by thirst (Lin et al. 2014; Senapati et al. 2019), but not affected by hunger, suggesting that motivation for

the water seeking behaviour was a result of water deprivation (Senapati et al. 2019). Neuropeptides such as *Drosophila* Neuropeptide F (dNPF, ortholog of mammalian NPY) and Leucokinin (Lk) signal hunger and thirst in *Drosophila* (Krashes et al. 2009; Senapati et al. 2019). Receptors for dNPF (*npfr1*) are found on a subset of PPL1 DANs (PPL1-γ1pedc), and inhibition of these PPL1 DANs promotes sugar memory performance in satiated *Drosophila* (Krashes et al. 2009). Lk provides both a hunger and thirst signal to the *Drosophila* MB. Lk activates a subset of PAM-DANs (PAM-β'2mp) to signal hunger, but inhibits a different subset of PAM-DANs (PAM-β'2a, but also PPL1-$\gamma2\alpha'1$) in the context of thirst, gating water memory expression (Senapati et al. 2019). This report aims to understand if the internal state of adult *Drosophila* influences learnt behaviours towards different concentrations of NaCl. In addition, scRNA-seq data from adult virgin or mated, males or female *Drosophila* midbrains experiencing different NaCl feeding conditions (0 M NaCl or six hours of 0.2 M NaCl feeding) was analysed for this report to understand how NaCl feeding status is represented in the brain and what cell types might be involved.

1.4 Single-cell RNA Sequencing Technology for Cell Type Annotation and Understanding Internal State

ScRNA-seq technologies have been integral for an unbiased approach to cellular characterisation based on gene expression. ScRNA-seq experiments start with isolating single-cells from a tissue of interest. The RNA from these single-cells are captured, uniquely barcoded, converted into cDNA before amplification and library preparation (Jovic et al. 2022; Zheng et al. 2017). The unique barcoding for each RNA contains sequences unique for the cell and for the RNA (unique molecular identifier, UMI). UMIs are used to combat amplification-bias and the cell barcode allows each read (a sequenced representation of the RNA) to be associated with the cell of origin after sequencing. Following sequencing, reads are then mapped to a reference genome. Clustering analysis groups cells by transcriptional profile, and cell types are annotated by known marker genes. Identifying cell type clusters by known marker genes allows for the identification

of additional marker genes. Novel cell types could be identified if the transcriptional profile does not match known marker genes.

ScRNA-seq technology has also been used to identify changes in brain states under different conditions. For example, Park et al., (2022) used scRNA-seq data of midbrains from *Drosophila* that experienced different levels of dehydration. Park et al., found that the expression of *astray* (*aay*) and *Drat* were increased in different glial cell types in *Drosophila* experiencing greater levels of dehydration. These changes in gene expression were biologically significant as RNAi knockdown of *aay* in astrocytes reduced water consumption. Dietary consumption of D-serine, the product of *aay*-encoded phosphoserine phosphatase, caused an enhanced water memory performance at 24 hours compared to control *Drosophila*. Usually, control *Drosophila* would have a reduced water memory performance at 24 hours (Lin et al. 2014, Figure 1b).

1.5 Thesis summary

The research outlined in this thesis aimed to establish a model for NaCl learning in adult *Drosophila*, use scRNA-seq data to understand how NaCl status is represented in the *Drosophila* brain, to characterise the learning and memory-related cells in *Apis* and to perform a comparative analysis of scRNA-seq data from the KC and DAN clusters between the two species.

The research outlined in this thesis found the existence of short-term and long-lasting memory for NaCl in *Drosophila* experiencing different internal states. Short-term appetitive memory for low (0.05 M) NaCl was detected in nutrient-deprived *Drosophila*. The valence of the detected long-lasting memory for high (0.2 M) NaCl was dependent on sex and nutrient status. 6 hour memory to high NaCl was appetitive in nutrient-deprived *Drosophila*, but aversive in *Drosophila* fed a standard diet. This aversive memory in fed *Drosophila* was more pronounced in female *Drosophila*.

This thesis also aimed to use scRNA-seq technology to explore the transcriptional changes that occur in the brains of *Drosophila* that were nutrient-deprived, compared to those that were fed a 0.2 M NaCl solution. Unfortunately, the data was impacted by a high level of adapter contamination, resulting in reads mapping to the reference transcriptome by 1.5-4.9%. Despite this, the data clustered into the major cell types and the proportion of cell types was similar to the data published by Park et al., (2022). DESeq2 differential gene expression analysis of a pseudobulk version of the scRNA-seq data in this report identified one differentially expressed gene. *Myosuppressin* was up-regulated in the nutrient-deprived female *Drosophila* compared 0.2 M NaCl-fed female *Drosophila* in one cholinergic cluster.

Finally, scRNA-seq analysis of twelve *Apis* dorsal protocerebrum identified nine KC clusters, suggesting greater granularity to the *Apis* KC subtypes compared to what was published previously in the literature. Comparing the KC clusters of *Drosophila* to those of *Apis* found little correspondence between the two species; however, marker genes for *Drosophila* PAM and PPL1-DANs annotated two *Apis* DAN clusters, this could suggest a similarity in function between the species.

In summary, this report outlines the existence of a novel memory to NaCl in adult *Drosophila* and that the neuropeptide *myosuppressin* may signal NaCl status in the brain. This report also shows that there is additional granularity to the *Apis* KCs than previously thought in the literature and that two *Apis* DAN clusters express marker genes for *Drosophila* PAM and PPL1-DANs.

Chapter 2

Methods

Contents

2.1 Fly Husbandry

Drosophila melanogaster were maintained on the standard cornmeal diet at 40-50% humidity and 25°C in 12:12 hours of a light:dark cycle. Canton-S wild-type (WT) *Drosophila* were used for all experiments.

2.2 Behavioural Experiments

20 hours before behavioural experiments, populations of $\sim$ 100 *Drosophila* were transferred to 25 mL vials containing 1% agar, 2.5 mL of Milli-Q® water, and a 20 x 60 mm sized filter paper to mineral and food deprive the *Drosophila* ('nutrient-deprived *Drosophila*'). *Drosophila* that were not nutrient-deprived before experiments were instead transferred to 25 mL vials containing 2.5 mL of standard cornmeal-agar food ('fed *Drosophila*'), with 20 x 60 mm sized filter paper. Mixed-sex populations of *Drosophila* were used for each experiment, unless otherwise stated. Male and female *Drosophila* were counted separately.

The olfactory associative learning assay protocols were adapted from those previously outlined by Krashes et al. 2008 and Krashes et al. 2009 to test immediate and long-term memory of NaCl in *Drosophila* experiencing different internal states. In brief, these experiments were performed using a T-maze (as described by Tully et al. 1994 and Yin et al. 1994), where the *Drosophila* were trained for 2 minutes in one training arm of the maze where the *Drosophila* were exposed to a 50 x 70 mm dry piece of filter paper paired with a novel odour (A) (CS-). The *Drosophila* were exposed to a 30-second stream of fresh air, before being tapped into the second arm lined with the experimental condition (1% agar plus different concentrations of NaCl (0-0.2 M NaCl) set onto a 50 x 70 mm piece of filter paper) (CS+) which was paired with a different novel odour (B) for a further 2 minutes. To test immediate memory, the *Drosophila* were directly transferred to a two-choice assay containing the training odours A and B (Figure 2.1). To test the influence of feeding status on long-lasting memory, *Drosophila* were transferred either to nutrient-deprivation or feeding vials (as previously described), or vials containing 1%

agar, 2.5 mL of Milli-Q® water, 0.1 M or 0.2 M NaCl and a 20 x 60 mm sized filter paper for 3, 6 or 24 hour time periods before entering the two-choice assay (Figure 2.2). After 2 minutes, the *Drosophila* were trapped and frozen. A performance index (PI) was calculated by subtracting the number of *Drosophila* in the CS- arm from the number of *Drosophila* in the CS+ arm and dividing by the total number of *Drosophila*. The PI calculated from one experiment was averaged with another experiment's PI where the order of odours A and B were reversed.

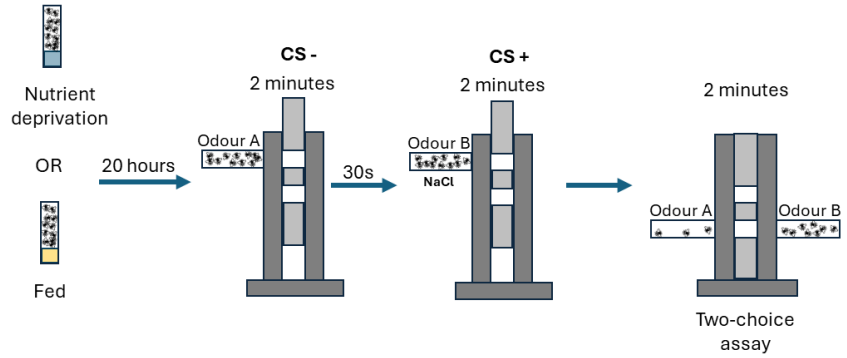


Figure 2.1: T-maze protocol for testing immediate memory to 0 - 0.2 M NaCl. Protocol for this olfactory associative learning assay adapted from Krashes et al. 2008 and Krashes et al. 2009. For this assay, $\sim$ 100 mixed-sex populations of male and female *Drosophila* were either fed a standard yellow food diet or nutrient-deprived for 20 hours before training. During training, the populations of *Drosophila* were exposed to a novel odour (A) within one arm of the T-maze for 2 minutes (CS-). The *Drosophila* were then exposed to 30 seconds of fresh air before being transferred to a second arm of the T-maze, where a second novel odour (B) was paired with the presentation of different NaCl concentrations (0 - 0.2 M NaCl) (CS+). The *Drosophila* were then transferred to a two-choice assay with both odours A and B. After two minutes, the *Drosophila* were trapped, frozen, and the *Drosophila* from the two arms were counted.

2.3 Statistical Analysis

Two-tailed one-sample T-tests were used to compare the statistical significance of learning scores compared to a hypothetical mean of zero, which suggests no memory expression. Data analysis was performed using R programming language version 4.3.0. All graphs and data expressed are stated as mean $\pm$ SEM. Significance is given as *$p<0.05$ and ***$p< 0.001$.

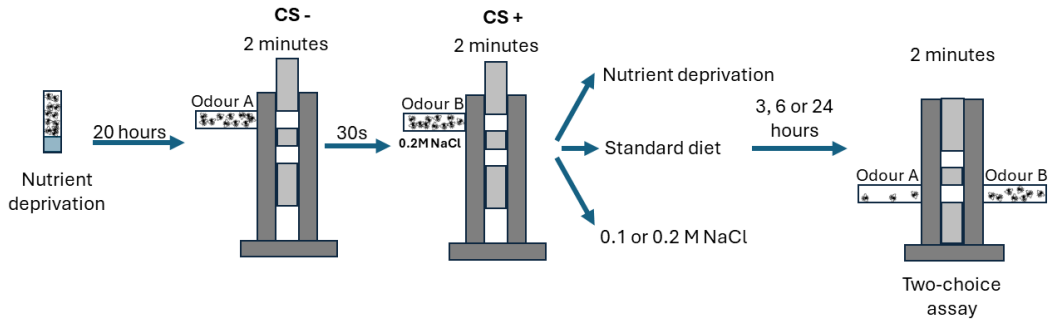


Figure 2.2: T-maze assay for testing memory persistence to 0.2 M NaCl. $\sim$ 100 mixed-sex populations of male and female *Drosophila* were nutrient-deprived 20 hours before training. During training, the populations of *Drosophila* were exposed to a novel odour (A) within one arm of the T-maze for 2 minutes (CS-). The *Drosophila* were then exposed to 30 seconds of fresh air before being transferred to a second arm of the T-maze, where a second novel odour (B) was paired with the presentation of 0.2 M NaCl (CS+). After training, the *Drosophila* were transferred to vials containing different feeding conditions. The *Drosophila* were either exposed to a standard yellow food diet, nutrient-deprivation or 0.1 or 0.2 M of NaCl in an agar and Milli-Q® water solution. After 3, 6 or 24-hour time intervals, the *Drosophila* were transferred to a two-choice assay with both odours A and B. The number of male and female *Drosophila* in each decision arm were counted.

2.4 ScRNA-seq Experiment Preparation

2.4.1 *Drosophila melanogaster*

Mated and unmated, female and male *Drosophila* were used in this experiment. To ensure the capture of virgin male and virgin female *Drosophila*, *Drosophila* were taken at pupal stage and transferred to separate feeding containers. Mated *Drosophila* were taken from a mixed population of female and male *Drosophila* to ensure copulation. 10 females and 10 males were used for each sample. All *Drosophila* were nutrient-deprived (as described previously) for 20 hours before 6 hours of continued nutrient-deprivation or 6 hours of NaCl-feeding in vials containing 1% agar, 2.5 mL of Milli-Q® water, 0.2 M NaCl and a 20 x 60 mm sized filter paper ('NaCl-fed *Drosophila*'); similar to the protocol explained in Figure 2.2. The four experimental conditions were as follows:

1. Mated male and female *Drosophila*. Nutrient-deprived for 26 hours.
2. Unmated male and female *Drosophila*. Nutrient-deprived for 26 hours.
3. Mated male and female *Drosophila*. Nutrient-deprived for 20 hours, then 6 hours of NaCl feeding.
4. Unmated male and female *Drosophila*. Nutrient-deprived for 20 hours, then 6 hours of NaCl feeding.

Two replicates of each condition were collected. The *Drosophila* brains were dissected and dissociated as described in Croset et al., (2018) and Park et al., (2022). During dissection, the optic lobes were removed, leaving the *Drosophila* midbrain.

2.4.2 *Apis mellifera*

Apis mellifera var. carnica honey bees were collected by Professor Geraldine Wright's group from the The John Krebs Field Station, Wytham, Oxford, UK. The bees were processed for sequencing by members of Professor Scott Waddell's group at the Centre for Neural Circuits and Behaviour, University of Oxford, UK. To ensure the capture of newly emerged *Apis*, a capped brood frame from a parent hive was housed within an incubator (Sanyo: at 34 °C, 60% RH). Four bees were collected from a single hive in Oxfordshire (UK) in August 2019, and eight more were collected in October 2020 from the same location but from a different colony (Queen from the same supplier). $\sim$ 200 bees were tagged with a numbered tag (Opalith, Germany), applied to the bee's shaved thorax with a drop of cyanoacrylate glue (Gorilla Super glue). The bees were placed back into their parent hive after tagging to mature. Foraging *Apis* were collected 36 days post-emergence to allow for approximately 15 days of foraging, which typically occurs 21 days post-emergence. To ensure the *Apis* were the desired class of pollen-forager, only the bees which were returning to the hive that had visible pollen baskets on their legs and a numbered tag on their thorax were captured. The captured bees were placed in a plastic feeding box with feeding tubes (2.0 ml eppendorf tubes, modified as in Stabler et al. 2021), with *ad libitum* 1.0 M sucrose in an incubator (same conditions as above) for 22 hours before dissection. The *Apis* brains were dissected and dissociated as described in Stabler et al., 2021, adapted to remove the optic lobes, trachea and the ventral portion of the brain, leaving the dorsal protocerebrum.

2.5 Dissection and Dissociation Procedure

For both *Apis* and *Drosophila* samples, brains were dissected in DPBS (ThermoFischer, 14190094) supplemented with toxin (toxin supplement: 20 mM 6,7-dinitroquinoxaline-2,3-dione, 0.1 mM tetrodotoxin and 50 mM d(-)-2- amino-5-phosphonovaleric acid) and transferred to 500 μL of Schneider's medium supplemented with toxin (tSM: ThermoFischer, 21720024, plus the toxin supplement) in a Protein LoBind tube on ice. Samples were washed in 1mL tSM and then incubated with a dissociation solution (PaCo: 25μL

papaine, Sigma P4762, and 25μL collagenase, Sigma C2674, in 450μL Schneider medium) for 20 minutes. Samples were then washed with tSM. The samples were titrated using a flame-rounded P200 tip before straining through a 10 mm CellTrix strainer (Sysmex, 04-0042-2314) into a Protein LoBind microcentrifuge tube, rinsing with Schnider's culture medium. Cell concentration was measured using a Leica DMIL LED Fluo microscope and a disposable Fuchs-Rosenthal haemocytometer (VWR, 631–1096) with a 1:10 solution of cell suspension.

2.6 Library preparation and scRNA-seq

The 10X Genomics, Chromium Single Cell 3' Reagent Kit V3 was used, as per the manufacturer's instructions, to perform mRNA barcoding. To capture a high proportion of cells in *Apis* samples, two libraries were prepared per replicate. Samples were processed at the Oxford's Wellcome Trust Centre for Human Genetics using NovaSeq 6000 (Illumina).

2.7 Cell Ranger and Read Quality Checks

Sequencing output files were checked for quality using FastQC (V 0.11.9 Andrews 2010) and MutiQC (V 1.14). Fastq output files were processed using CellRanger (v5) to align reads to Ensembl's reference genome Amel_Hav3.1 for *Apis* and Dmel_6.50 for *Drosophila*.

2.8 Quality Control in R

For both experiments, the output from Cell Ranger was read into RStudio using Read10X(), and a Seurat object was built for each sample using CreateSeuratObject(), filtering for min.cells = 3 and min.features = 200.

2.8.1 *Drosophila melanogaster*

Drosophila data was filtered similar to what was described by Park et al., (2022). Barcodes were filtered by removing those with >15% ribosomal proteins and mitochondrial

RNA, and >10% rRNA. Barcodes with <200 or >2500 features, or >8000 UMIs were removed as these were assumed to be empty droplets or multiplets. Data was scaled and normalised, and variable features were identified using SCTransform(). Replicates for each condition were merged before integrating with Seurat V3 using reduction = "cca" and dims = 1:40. For both *Apis* and *Drosophila* datasets, dimensionality reduction was performed using runPCA(). Elbow plots identified PC 1:40 as the range that captures the most variation in the data for each sample. RunUMAP(), FindNeighbors() and FindClusters() was used to find clusters. SingleR() was used to transfer labels from the Park et al., (2022) paper to the *Drosophila* data in this report.

Differential gene expression analysis was carried out using Bioconductor's DESeq2 package (Love et al. 2014) on both the Park et al., (2022) data as a reference, comparing satiated *Drosophila* with 12-hour water-deprived *Drosophila*, and the NaCl-related scRNA-seq data in this report. For this report, I analysed differential expression of genes between the nutrient-deprived and NaCl-fed conditions in all *Drosophila*, female *Drosophila* and the interaction effect of mating status on condition effects. DESeq2 was performed on count data from each cluster. Counts for each cluster were aggregated by sample to create pseudobulk data. Clusters with less than three cells per condition were excluded from the analysis. Genes were considered differentially expressed if the adjusted p-value <0.05 and the absolute log2(FC) values >1.

2.8.2 *Apis mellifera*

The data was filtered for cells with >250 features and >500 UMIs. As two libraries were prepared for each sample, the two datasets per sample were filtered by UMI and feature count before being merged together. Each sample was processed using the SCTransform() to normalise and scale the data, and to find variable features. Clustering at a resolution of 0.1 found a range of 11-17 clusters per sample. scDblFinder() and decontX() were used to identify doublets and cells with high contamination, respectively. Cells which were identified as "doublets" were removed. Samples were integrated using Harmony(), correcting for sample batch effects. FindAllMarkers() with only.pos = TRUE, a log2c.threshold of 1 and a cluster resolution of 0.1 unless otherwise stated.

2.9 Machine Learning Methods for Classifying Kenyon Cell Clusters

For classifying *Apis* Kenyon Cell clusters, Kenyon Cell cell-by-feature count matrices were filtered for neuropeptide, neurotransmitter, neuropeptide receptor and neurotransmitter receptor genes from FlyBase (accessed 06/12/2024) that have an identified homologous gene in *Apis*. For this, genes that were found using the search terms "Neurotransmitter receptor", "Neuropeptide receptor", "Neurotransmitter" and "Neuropeptide" in FlyBase were then matched with *Apis mellifera* homologues identified from EnsemblMetazoa BioMart. For the comparison of top ranked features output from classifying *Apis* and *Drosophila* Kenyon Cell clusters, cell-by-feature count matrices were filtered for all homologous genes between *Apis* and *Drosophila* from EnsemblMetazoa BioMart.

To classify Kenyon Cell clusters a Random Forest machine-learning model run in Python's scikit-learn was used based on cell-by-feature count matrices as described above. train_test_split was imported from sklearn.model_selection and run with random_state = 17 and test_size = 0.2. This was fit to the RandomForestClassifier() and the classification_report imported from sklearn.metrics and run to produce the output report for the model. Feature importances were found using feature_importances_. Confusion matrices were produced by importing confusion_matrix and ConfusionMatrixDisplay from sklearn.metrics.

Chapter 3

A State-Dependent Nutrient Memory for NaCl in Adult *Drosophila*

Contents

3.1 Introduction

Optimal intake of sodium chloride is required for normal muscle and nerve function, and osmotic balance. *Drosophila melanogaster* can discriminate low from high NaCl concentrations. *Drosophila* will choose to consume low NaCl concentrations ($\sim$ 0.05 M NaCl) in a sucrose solution compared to sucrose alone, but will avoid high NaCl concentrations ($\geq$ 0.2 M) (Zhang et al. 2013). Feeding behaviours towards high NaCl concentrations change with sodium deprivation or when the nutritional demand increases after copulation (Walker et al. 2015; Jaeger et al. 2018). Considering *Drosophila* can form state-dependent associative olfactory memories to other essential nutrients, such as nutritious sugars and water (Burke et al. 2011; Huetteroth et al. 2015; Lin et al. 2014; Senapati et al. 2019), we wanted to understand if *Drosophila* can also form memories of NaCl, if memory valence is affected by NaCl concentration, and if the animal's internal need for NaCl (NaCl deprivation or post-copulation nutritional changes) influences these memories.

As mentioned in Chapter 1, a memory for NaCl has been identified in rats, crickets and *Drosophila* larvae. In crickets, a high concentration of NaCl (2 M NaCl) was used as the US during aversive olfactory conditioning (Mizunami et al. 2010). In these experiments, one of two odours (CS+) was present to the cricket's antenna before the US was presented to the mouth. During a preference test between the CS+ and CS- odours, the crickets showed a greater preference for the CS- odour, compared to the CS+ odour, suggesting the cricket formed an aversive associative memory to the high NaCl solution. In experiments using *Drosophila* larvae, larvae were placed onto a petri dish with either agarose or agarose mixed with different concentrations of NaCl (US) and each petri dish was associated with a different odour (CS- and CS+, respectively). During the odour preference test, more larvae chose the CS+ odour when the US was 0.375 or 0.750 M NaCl. No memory performance could be detected at higher concentrations of NaCl unless the preference test was presented with the US (US presented with both CS- and CS+ odours); under these conditions, more larvae approached the CS- odours

when the US was 1.5 M NaCl (Niewalda et al. 2008). In rats, behaviours towards a previously aversive cue associated with a solution of high NaCl became appetitive when in a sodium-deficient state (Robinson et al. 2013).

Drosophila is a desirable model organism for studying nutrient memories for the well-established appetitive olfactory learning paradigm (Tempel et al. 1983; Krashes et al. 2008), that has been adapted to study state-dependent learning and memory dynamics of essential nutrients such as nutritious sugar and water (Burke et al. 2011; Huetteroth et al. 2015; Lin et al. 2014; Senapati et al. 2019). This assay, combined with the use of powerful genetic tools such as the GAL4/UAS system, has allowed for the dissection of the neuronal circuits involved in these nutrient memories (see section 1.3). Experiments such as these also identified neuropeptide signalling on DANs that project to the MBs which coordinate state-dependent memories for water and nutritious sugar (Krashes et al. 2009; Senapati et al. 2019).

Considering a state-dependent memory for NaCl has been identified in rats, and appetitive or aversive memories have been found in *Drosophila* larvae depending on the US NaCl concentration, we wanted to explore the existence of a state-dependent memory for NaCl in adult *Drosophila*. Because adult *Drosophila* have different preferences for NaCl depending on the concentration and their nutritional status, we also wanted to understand if the nutrient-status and / or the concentration of the NaCl solution presented as the US influence approach or avoidance behaviours to the CS+ odours. Therefore, this chapter aims to utilise state-of-the-art behavioural assays to identify novel NaCl memories in *Drosophila*. Further, we aimed to understand if the formation or expression of these memories was influenced by the animal's internal state, such as mating or nutrient status.

The data in this chapter was collected by me and Dr. Senapati from Prof. Waddell's group, CNCB, unless stated otherwise.

3.2 Results

3.2.1 Appetitive Immediate Memory to Cues Predictive of 0.05 M NaCl

The essential micronutrients in NaCl elicit an appetitive or an aversive behavioural response in *Drosophila* depending on the concentration (see introduction section 3.1). Given that taste memories for nutrients, such as nutritious sugars, can be measured immediately after training, and the expression of these memories is affected by the animal's internal state (Lin et al. 2014; Senapati et al. 2019; Krashes et al. 2008; Krashes et al. 2009), we wanted to understand if this was also the case for NaCl. Additionally, we hypothesised that the valence of the memories for different concentrations of NaCl would correspond to the results observed in feeding assays (Zhang et al. 2013). We utilised the olfactory associative learning assay using a T-maze to condition nutrient-deprived or fed *Drosophila* to associate the second of two odours with different concentrations of NaCl in a 1% agar solid medium (see Methods). The performance index (PI) was calculated by subtracting the number of flies trapped in the CS- arm from the number of flies in the CS+ arm, divided by the total number of flies in the experiment. The PI score from one experiment was averaged with the PI score from another experiment where the odours for CS- and CS+ were reversed. A positive PI score suggests more flies were choosing the conditioned odour and a negative score suggests more flies were choosing the alternative odour. In experiments where hungry *Drosophila* were trained with sucrose, the PI score after testing immediate memory was $\sim$0.3 (Krashes et al. 2008; Krashes et al. 2009), whereas, *Drosophila* trained with the aversive tastant, DEET mixed with a sugar solution (0.4% DEET, 3 M Xylose and 0.1 M sucrose), had an immediate memory PI score of $\sim$-0.4 (Das et al. 2014). *Drosophila* trained with 0 M NaCl showed no significant behavioural response. Nutrient-deprived *Drosophila* preferred the odour previously paired with 0.05 M NaCl, with an average performance index (PI) score of 0.13 (see Figure 3.1). This appetitive score was lower than the appetitive scores obtained from sucrose learning (Krashes et al. 2008; Krashes et al. 2009), however, this score was significantly different from zero (p-value = .04, t(7) = 2.517, mean= 0.134, SD=0.053, two-tailed one-sample

t-test, Figure 3.1), therefore evidencing the detection of an appetitive memory. This behaviour was not seen in *Drosophila* that were fed a standard diet. This suggests that the presentation of 0.05 M NaCl was sufficient for learning, and nutrient-deprivation increases motivation to seek out the odour predictive of the low NaCl concentration. Considering the appetitive concentration of NaCl provided an appetitive PI score, we expected that the aversive concentration of NaCl would provide an aversive PI score. However, both nutrient-deprived and fed *Drosophila* showed no preference for the conditioned odour previously paired with 0.2 M NaCl or the alternative odour.

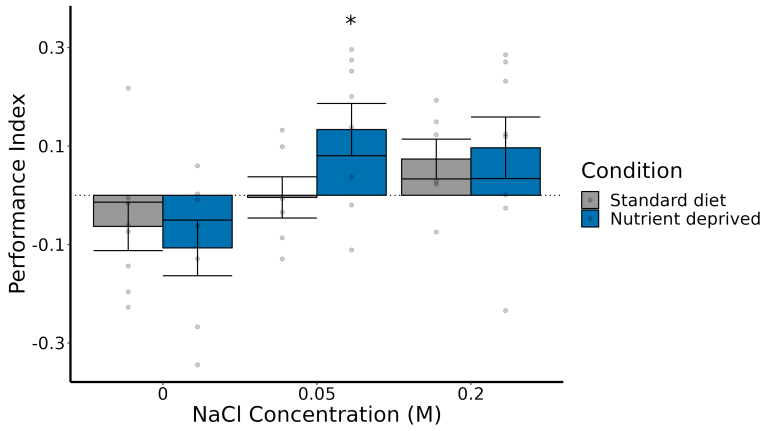


Figure 3.1: An appetitive immediate memory performance detected in nutrient-deprived *Drosophila* to low concentrations of NaCl (0.05 M NaCl). Performance index scores from 20-hour fed (grey bars) or nutrient-deprived (blue bars) mixed-sex populations of *Drosophila* trained with 0 - 0.2 M NaCl and testing immediate memory. Nutrient-deprived *Drosophila* preferred the conditioned odour associated with 0.05 M NaCl compared to the unconditioned odour (p-value = .04, t(7) = 2.517, mean= 0.134, SD=0.053, two-tailed one-sample t-test). Fed-*Drosophila* showed no preference for either the conditioned or unconditioned odour (p-value = .922, t(5) = -0.103, mean = -0.004, SD = 0.103, two-tailed, one-sample t-test). N = 6-8, all data is shown as mean ± SEM.

3.2.2 The Valence of long-lasting Memory to NaCl is Dependent on Nutrient Status

Memories of sweet-tasting, non-nutritious sugars are short-lived in *Drosophila*. However, memories of sugars with a nutrient value are remembered long-term (Huetteroth et al. 2015; Burke et al. 2011), with approach behaviours towards the odour previously paired with nutritious sugar detected at 36 hours in hungry *Drosophila* (Krashes et al. 2008). Feeding the *Drosophila* reduces memory performance (Krashes et al. 2009, Figure 1a), suggesting the *Drosophila*'s internal state affects the animal's motivation to seek out the odour paired previously with the nutritious sugar also in the long term. We, therefore, aimed to understand if the presentation of 0.05 M NaCl or 0.2 M NaCl could provide value to the animal to reinforce the immediate memory into a detectable long-lasting memory performance, and if the animal's internal state affects the performance outcomes. Here, we nutrient-deprived *Drosophila* before training with 0.05 M or 0.2 M NaCl and then put *Drosophila* into different feeding conditions before testing 0, 3, 6, and 24-hour memory. *Drosophila* were either nutrient-deprived or fed a standard diet before testing.

As evidenced in section 3.2.1, an immediate memory for 0.05 M NaCl can be detected in nutrient-deprived *Drosophila*. However, the PI scores for both nutrient-deprived and fed *Drosophila* decreased with time, tending towards zero (Figure 3.2). This suggests that the low concentration of NaCl was sufficient for short-term memory, like sweet-tasting sugars, but was not sufficient to provide the additional rewarding values to reinforce the short-term memory into long-lasting memory. Interestingly, when testing 24-hour memory, *Drosophila* trained with 0.2 M NaCl and then nutrient-deprived preferred the previously NaCl-paired odour compared to the alternative odour, with a mean average positive PI score of 0.126. The opposite was true in *Drosophila* trained with 0.2 M NaCl and fed before testing 6-hour memory, with a mean average negative PI score of -0.194 (Figure 3.2).

The results from 0.2 M NaCl suggest that, although this concentration of NaCl was not sufficient to provide detectable short-term memory (section 3.2.1) it did provide value to

the animal to create long-lasting memories. The memory of 0.2 M NaCl has an opposing valence that depends on the animal's internal state. This result is similar to those reported by Das et al., (2014) where the aversive stimulus, DEET, combined with an appetitive nutritious sugar, formed parallel and opposing memories. The *Drosophila* exhibited an aversive short-term memory which later became a robust appetitive long-term memory 30 minutes after training, suggesting two memory components were guiding their behaviours (Das et al. 2014). In contrast, the memory for 0.2 M NaCl appears to have opposing values, but to the same stimuli, the memory output appears to be gated by the animal's internal state.

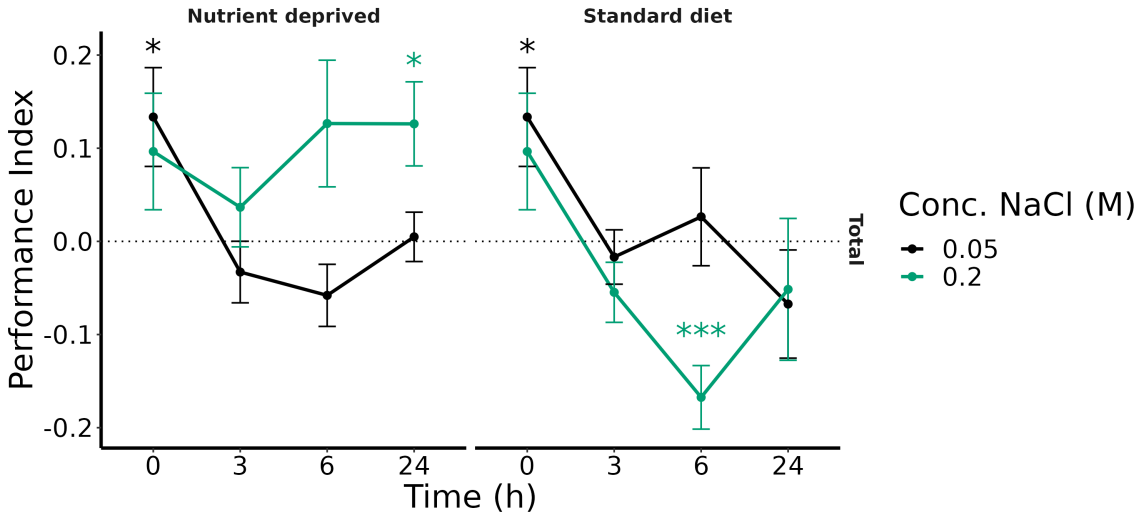


Figure 3.2: Behaviour towards odours predictive of high NaCl (0.2 M NaCl) depends on the feeding status of the *Drosophila* before testing. Performance index scores from populations of mixed-sex nutrient-deprived *Drosophila* trained with low or high concentrations of NaCl (0.05 M and 0.2 M, respectively), which were placed into different feeding conditions before testing 3, 6 or 24-hour memory recall. *Drosophila* which were nutrient-deprived before testing 24-hour memory approached the odour predictive of 0.2 M NaCl (p-value = 0.027, t(7) = 2.795, mean = 0.126, two-sided, one-sample t-test). However, *Drosophila* fed a standard diet before testing 0.2 M NaCl memory approached the alternative odour (p-value <0.001, t(6) = -7.69, mean = -0.194, two-sided, one-sample t-test). In both nutrient-deprived and fed *Drosophila*, the learning scores associated with 0.05 M NaCl tended towards zero over time. N = 8 for all.

3.2.3 Aversive 6-hour Memory Performance in Fed *Drosophila* is Significant in Females

Feeding behaviours towards NaCl increase in female *Drosophila* after copulation (Walker et al. 2015). We, therefore, evaluated the PI scores from females and males by splitting the data from Figure 3.2 into male and female counts. This revealed that the aversive PI score when testing 6-hour memory in *Drosophila* fed a standard diet was only significantly different from zero in female *Drosophila*, although the PI score for male *Drosophila* fed a standard diet was negative.

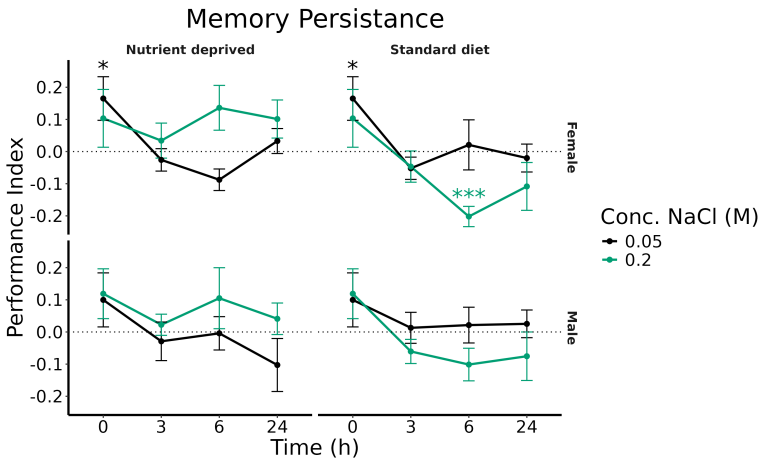


Figure 3.3: Female *Drosophila* fed a standard diet before testing, expressed an aversive high NaCl memory at 6 hours. The same data is shown in Figure 3.2, but male and female counts were plotted separately. PI scores in female *Drosophila* fed a standard diet before testing 6-hour memory were significantly different from zero (p-value <0.001, $t(7) = -6.41$, mean = -0.2, two-sided, one-sample t-test). The 6-hour PI scores for fed male *Drosophila*, however, was not significantly different from zero (p-value = 0.086, $t(7) = -2.00$, mean = -0.10, two-sided, one-sample t-test). N = 7-8 for all. Due to some uncertainty with the count data collected for one trail of 24-hour 0.2 M NaCl memory in female *Drosophila* that were nutrient-deprived before testing, the corresponding PI score was removed from analysis. It is likely, as result of removing this PI, the sex-specific average PI scores no longer align with the average total PI scores in Figure 3.2.

3.2.4 High NaCl 6-Hour Memory Performance is State-Specific

Given that providing the *Drosophila* with *ad libitum* access to standard fly food before testing 0.2 M NaCl memory performance affected the behavioural output, we wanted to understand if providing NaCl before testing would affect 0.2 M NaCl memory scores. We, therefore, tested 6-hour 0.2 M NaCl memory performance in nutrient-deprived *Drosophila*, fed *Drosophila* and *Drosophila* with *ad libitum* access to 0.1 M NaCl before testing. Interestingly, the PI scores for *Drosophila* that were fed 0.1 M NaCl before testing were close to zero (Figure 3.4a), with no observable difference between males and females (Figure 3.4b). This suggests that providing NaCl reduced the *Drosophila*'s motivation to approach or avoid the previously NaCl-paired odour.

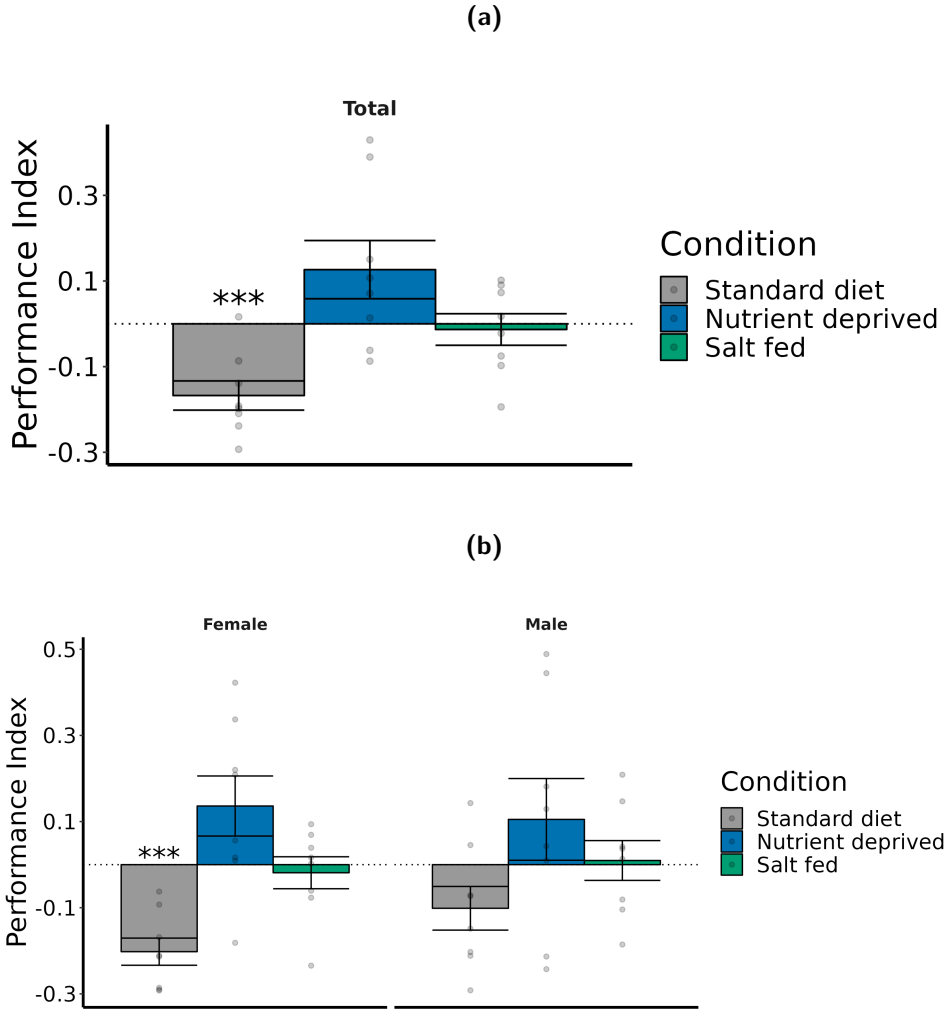


Figure 3.4: Feeding *Drosophila* 0.1 M NaCl before testing 6-hour memory recall to odours predictive of 0.2 M NaCl resulted in a PI score close to zero.
Populations of mixed males and females were nutrient-deprived for 20 hours before olfactory associative learning conditioning with 0.2 M NaCl. After training, the *Drosophila* were fed a standard diet, nutrient-deprived or fed 0.1 M NaCl before testing memory performance. As seen in Figures 3.2 and 3.3, *Drosophila* fed a standard diet before testing high-NaCl memory had an aversive response to the conditioned odour. When fed NaCl before testing memory, all *Drosophila* (males and females) showed no preference towards the conditioned or alternative odour (p-value = 0.73, t(7) = -0.35951, mean = -0.013, N = 8).

Unfortunately, due to the development of severe allergies to the *Drosophila*, I had to abandon this project. Further behavioural experiments were collected by Dr. Senapati from Prof. Waddell's group, CNCB. The preparation of the single-cell RNA sequencing experiment (outlined in Chapter 4) were carried out by Dr. Christoph Treiber and colleagues from the Waddell group, CNCB.

3.2.5 High NaCl-Fed Virgin Female *Drosophila* Tend to Avoid the Odour Predictive of High NaCl

Given that NaCl feeding behaviours of mated females are different from those of males and virgin females (Walker et al. 2015), we wanted to understand if there is a difference in PI scores when testing 6-hour 0.2 M NaCl memory in nutrient-deprived *Drosophila* compared to *Drosophila* fed 0.2 M NaCl, in mated and unmated female and male *Drosophila*. As before, the scores for NaCl-fed mated males and females were close to zero, however, the PI score for virgin female *Drosophila* tended towards an aversive PI score, but this was not significantly different from zero ($p = 0.06$, Figure 3.5). Data collection for behavioural experiments were usually collected over two or more days. On this occasion, the data for the experiment shown in Figure 3.5 was collected in one day. There may have been external variability that may have contributed to the behavioural scores from this experiment. It would be valuable to collect more data to improve confidence in the results from this experiment.

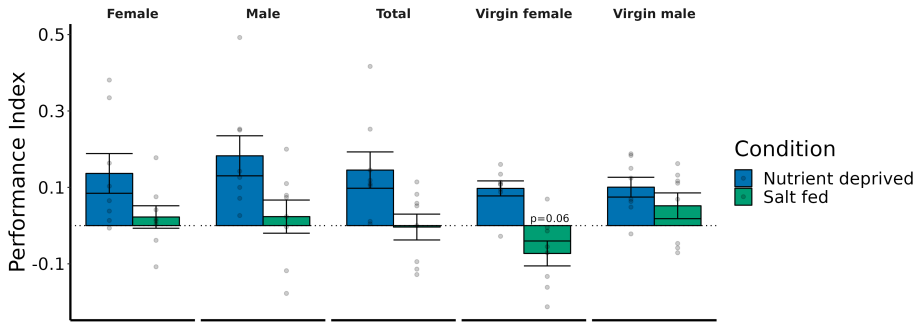


Figure 3.5: NaCl fed (salt fed) virgin female *Drosophila* tend to avoid the odour predictive of high NaCl 6 hours after conditioning.
Populations of mixed males and females, virgin females and virgin males were nutrient-deprived for 20 hours before training. *Drosophila* were trained as described in Figure 3.4. Salt-fed virgin-female *Drosophila* tend towards conditioned odour aversion (p-value = 0.06, t(7) = -2.234, mean = -0.07, two-tailed, one-sample t-test). N = 8.

3.3 Discussion

3.3.1 Results Summary

The results from this chapter provide evidence for the existence of NaCl olfactory memories in *Drosophila*. Appetitive olfactory memories are formed to 0.05 M NaCl, these are short-lived and decay over time. Short-term memory performance was not detected in *Drosophila* trained with 0.2 M NaCl, however, the valence of the memory detected at 6 hours and 24 hours depended on the internal state of the *Drosophila*. *Drosophila* that were nutrient-deprived before testing 24-hour 0.2 M NaCl memory performance approached the previously NaCl-paired odour, whereas, *Drosophila* fed a standard diet before testing 6-hour memory avoided the paired odour. The detected aversive behaviour was more pronounced in female *Drosophila*. Feeding the *Drosophila* NaCl before testing 6-hour 0.2 M NaCl memory seemed to abolish memory expression as no performance was detected. Virgin female *Drosophila* fed NaCl before testing 6-hour 0.2 M NaCl memory tend to avoid the CS+ odour.

3.3.2 Dissecting the Components of NaCl Memory

Drosophila can remember different components of the US. For example, there are two memory components to the memory of nutritious sugars, like sucrose, the sweet-taste short-term component and the long-term nutrient component (Huetteroth et al. 2015; Yamagata et al. 2015). These elements can be of opposing valence. For example, the memory of the US combining DEET (bitter taste) with sucrose was stored as a short-term aversive memory component and a long-term appetitive component (Das et al. 2014). These different components are controlled by different subsets of DANs (Das et al. 2014).

Short-term NaCl Memory

The results from this report detected an appetitive short-term memory to a low NaCl concentration (0.05M NaCl) in nutrient-deprived *Drosophila*. This short-term memory might be taste-related. Preliminary feeding experiments carried out by Dr. Bhagyashree Senapati showed that nutrient-deprived male and female *Drosophila* will consume NaCl

at different concentrations. These experiments detected staining on the proboscis and the abdomen of *Drosophila* after presenting solutions containing 0, 0.05 or 0.2 M NaCl dyed blue (Appendix Figure A.1). Different taste modalities are detected by subsets of Gustatory Receptor Neurones (GRNs) located in the proboscis (and also the tarsae and wing margin, Hiroi et al. 2004). Different GRNs respond to low and high salt concentrations (Jaeger et al. 2018; Zhang et al. 2013). GRNs project their axons to the brain region that processes gustatory information, the subesophageal zone (SEZ, Kendroud et al. 2018; Talay et al. 2017). Work carried out by Otto et al., (2020) identified neurones which project from the SEZ to the PAM-DANs which signal information on taste (bitter and sweet components) to reinforce memories. Given this information, it is possible that taste information detected by the GRNs could be carried to the DANs via the SEZ. To dissect the circuits involved in low-NaCl appetitive memory, target neurones identified from the connectomics data downstream of the NaCl-associated GRNs could be artificially inactivated using the optogenetically activated anion channel rhodopsin (GtACR1, Mohammad et al. 2017) during NaCl olfactory learning acquisition. Inactivation of neurones that are involved in low NaCl memory acquisition would impair learning, the *Drosophila* would show no preference for the conditioned or the alternative odour when testing memory recall.

Long-lasting NaCl Memory

24-hour and 6-hour high NaCl memory in nutrient-deprived and in *Drosophila* fed a standard diet, respectively, were of opposing valence. These memories of opposing valence could be created in parallel but independently. These NaCl memories may be coordinated by different DAN subtypes, such as those that are involved in long-term memory for sucrose (like PAM-α1, Yamagata et al. 2015) and those involved in aversive long-term memory for shock (like PPL-DANs, see section 1.3 and Plaçais et al. 2012). To test if PAM or PPL-DAN clusters are required for coordinating the opposing memories, their activity could be blocked during training, before testing 6 or 24 hour high-NaCl memory recall in *Drosophila* that are either nutrient-deprived or fed a standard diet. I would anticipate that blocking the activity of PAM-DANs during

training, and nutrient-depriving the *Drosophila*, would result in no approach behaviours during testing at 24 hours. I would expect that blocking PPL-DANs during training and then feeding the *Drosophila* before testing 6-hour memory would result in a reduced aversive response to the conditioned odour.

Detection of a long-lasting memory for high-NaCl suggests a nutritional element to this memory. Kim et al., (2024) identified a set of anterior enteric neurones that they refer to as 'internal sodium-sensing' (INSO) neurones. Inactivating these neurones in salt-deprived *Drosophila* reduced the preference for NaCl in sucrose compared to sucrose alone and ex vivo calcium imaging of the INSO neurones recorded neuronal activity when applying sodium-containing salts. This suggests that INSO neurones are sodium sensors and provide sodium-associated state-dependent signals. The INSO neurones project their axons to the SEZ of the *Drosophila* brain. Trans-synaptic tracing analysis tools (such as *trans-Tango*, Talay et al. 2017) could be employed to identify downstream connections between INSO neurones and higher *Drosophila* brain regions. It is possible that INSO neurones do not make direct connections to the DANs to provide the teaching signals to reinforce the nutrient memory. For example, neurones carrying the fructose sensor gustatory receptor 43a (Gr43a, Yamagata et al. 2015; Miyamoto et al. 2012; Miyamoto et al. 2013) are required for long-term sucrose memory, but they do not form direct connections with the PAM-DANs involved in signalling for the long-term memory for sucrose.

As INSO neurones are suggested to provide a sodium-associated state-dependent signal, it is possible that signalling by these neurones could affect 6 and 24-hour memory recall. For example, inactivating these neurones before testing 24-hour memory recall in nutrient-deprived *Drosophila* might reduce the approach behaviours. Other possible mechanisms for internal state signalling are discussed below.

3.3.3 Internal State

Nutrient-Status

In *Drosophila*, neuropeptides such as dNPF and Lk gate sucrose and water memories by providing internal state signals (such as thirst and hunger, see section 1.3). Neuropeptides may be responsible for signalling sodium-status. Given that Lk neurones respond to solutions of different osmolality (Senapati et al. 2019), they may respond to changes in sodium content of the fly's hemolymph. To understand if Lk signalling plays a role in state-dependent NaCl memory, RNAi knockdown of the Lk receptor (*Lkr*) in subsets of DANs (such as PPL1-DANs that gate water memory, Senapati et al. 2019) would help us understand its role in NaCl-memory. Once the DANs involved in NaCl-memory have been identified, single-cell RNA sequencing data could be used to identify other neuropeptide receptors that might be involved in signalling NaCl nutrient-status.

Mating-Status

Drosophila feeding behaviours towards NaCl change after copulation as nutritional demand shifts to support reproduction (Walker et al. 2015). This behaviour change is attributed to the changes in the post-mating circuit. The female *Drosophila* post-mating circuit involves the sex peptide receptor (SPR), the sex peptide sensory neurones (SPSNs) and the sex peptide abdominal ganglion (SAG). The SAG projects to the brain. SPR activation depresses the post-mating circuit (Feng et al. 2014). Walker et al., (2015) found that depressing the circuit (such as inactivating SAG and SPSN using the genetic tool *UAS-Kir2.1*) increased PER to a NaCl solution in virgin female *Drosophila*, similar to levels measured with mated-female *Drosophila*. As there was a difference in memory performance to high-NaCl 6-hour memory in mated female *Drosophila* compared to mated male *Drosophila*, it is possible that these differences are due to the post-mating circuits in female *Drosophila*. Repeating this experiment using virgin female *Drosophila* that were fed a standard diet before testing 6-hour high-NaCl memory would be useful to understand if the post-mating circuit is involved. I would expect virgin female *Drosophila* to perform similarly to male *Drosophila*. To probe the post-mating circuit, I would use RNAi against the SPR in SPSNs of mated female *Drosophila*. If these receptors are

involved in coordinating the differences in behaviour between mated males and females I would expect there to be no difference between memory performance between mated and virgin female *Drosophila*.

Interestingly, when testing 6-hour high-NaCl memory in virgin female *Drosophila* that were fed NaCl before testing, there was a trend towards aversion, whereas mated males and females showed no preference for the conditioned or the alternative odour. This difference in behaviour in virgin female *Drosophila* could also be attributed to the post-mating circuit. I expect that blocking the activity of SPSN and SAG neurones using GtACR1 in virgin female *Drosophila* during 6-hour high-NaCl memory tasks would create behaviours similar to mated females. As the SAG projects to the brain, it would be important to understand if the post-mating circuit connects with learning and memory circuits. To probe this, *trans-Tango* and connectomics would help identify potential connections between these circuits.

3.3.4 Conclusions

The results from this report identify the existence of a new NaCl olfactory associative memory in *Drosophila* adults. Memory dynamics appear to be influenced by the concentration of NaCl and the expressed valence is influenced by nutrient status and sex. Further research is required to understand the circuits involved in NaCl memories and how these circuits are influenced by signals associated with nutrient or mating status. The next chapter in my thesis (Chapter 4) outlines the research performed to understand how NaCl-related nutrient status is represented in the *Drosophila* brain, and how this might be influenced by mating status.

Chapter 4

The Transcriptomic Representation of Sodium Status in the Adult *Drosophila* Midbrain

Contents

4.1 Introduction

The results outlined in Chapter 3 provided evidence for the existence of a NaCl long-lasting memory in adult *Drosophila*. The valence of this NaCl long-lasting memory depended on nutrient status. *Drosophila* that were nutrient-deprived before testing 0.2 M NaCl long-lasting memory recall approached the conditioned odour. Whereas *Drosophila* fed a standard diet avoided the conditioned odour, and *Drosophila* fed NaCl before testing memory recall showed no preference for the conditioned or alternative odour (see Figures 3.2 and 3.4a). The conditioned aversion was more pronounced in fed female *Drosophila* (Figure 3.3), and the 6-hour memory performance of NaCl-fed virgin female *Drosophila* tended towards aversion (Figure 3.5). This chapter, therefore, aimed to use scRNA-seq data to understand how internal sodium status is represented in the *Drosophila* brain and how this is influenced by mating status.

The experiment outlined in this chapter was modelled on the scRNA-seq experiment published by Park et al., (2022). Park et al., (2022) compared gene expression between *Drosophila* that were satiated (no water deprivation) with *Drosophila* that were water deprived for twelve hours. As outlined in the Introduction section 1.4, their analysis identified biologically relevant differential gene expression (DGE) events in glial cell clusters. The scRNA-seq data for this chapter consisted of eight samples, each containing twenty *Drosophila* midbrains, ten of which were male and ten of which were female. Four samples contained midbrains of *Drosophila* that were nutrient-deprived, and the other were fed 0.2 M NaCl (see Methods). Of the four samples for each feeding condition, two samples contained midbrains from virgin *Drosophila*, and the other two samples contained midbrains from mated *Drosophila* (see Methods section 2.4). The cells from each sample could be classed as male or female based on the expression or absence of the male-specific gene *lncRNA:roX1*.

This chapter aimed to identify DGE events between all *Drosophila* that were nutrient-deprived and those fed 0.2 M NaCl. In addition, as the long-lasting NaCl memory behaviours were

more pronounced in fed female *Drosophila* than males, I aimed to compare DGE events from the two feeding conditions in female-only cells. Finally, I wanted to understand if there was an interaction effect of mating status on the DGE events between the two feeding conditions.

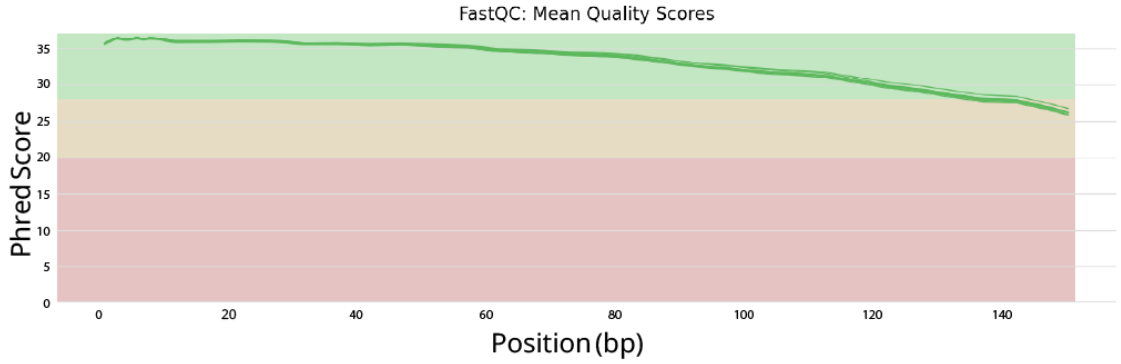
4.2 Results

4.2.1 Quality Control of Raw Sequencing Reads

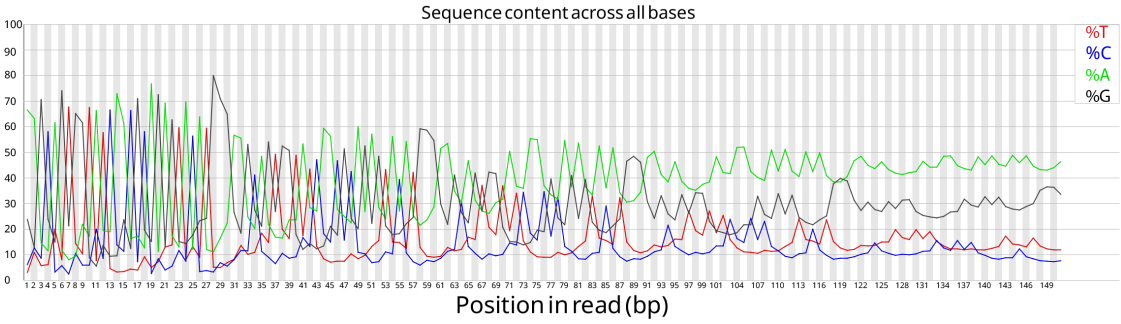
The quality of the FASTQ files output from scRNA-seq was assessed using FastQC and visualised using MultiQC. The confidence score for base calling was high for each of the reads (measured using a Phred score, Figure 4.1a). However, reviewing the percentage sequence content across the bases for each read found a repeating pattern in the base sequence (see example Figure 4.1b from sample 1, evidenced by the over representation of certain bases along the sequence). The percentage GC content for good-quality reads should be approximately normally distributed. For the reads in this report, the GC content was not normally distributed, suggesting there were overrepresented sequences or contamination (Figure 4.1c). A high proportion of the overrepresented sequences were from the "Clontech SMARTer II A Oligonucleotide" (Figure 4.1d), with more than 40% of reads in sample 6 containing this sequence.

To address the possibility of contamination caused by sample preparation, FastQ Screen was used to compare reads to multiple reference genomes, including *Drosophila*, Yeast, Mouse and Human. $\sim$1.9% of reads for each sample mapped to the *Drosophila genome*, $\sim$2.6% of reads mapped to the yeast genome and $\sim$22.7% of reads mapped exclusively to the "contamination" reference, which was the Clontech SMARTer II A Oligonucleotide sequence (Figure 4.2). This suggests that a large portion of the contamination is due to the Clontech adapter sequence, with a minor contribution from yeast.

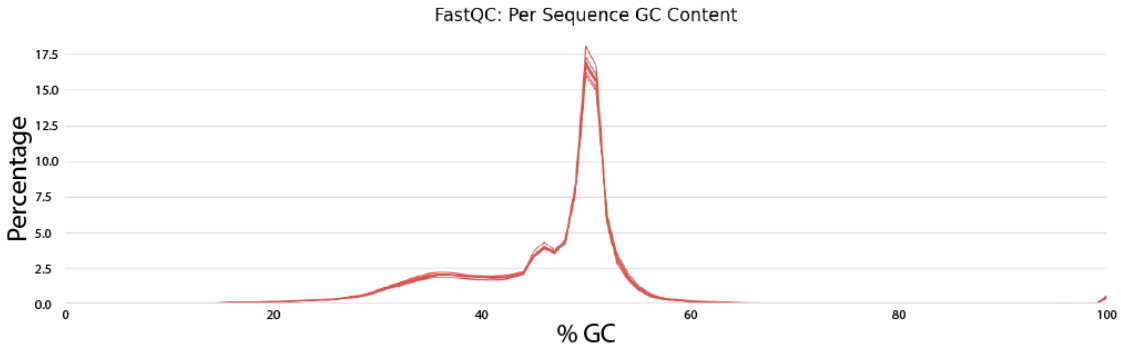
(a)


FastQC: Mean Quality Scores
Phred Score
Position (bp)

(b)


Sequence content across all bases
%T
%C
%A
%G
Position in read (bp)

(c)


FastQC: Per Sequence GC Content
Percentage
% GC

(d)

Overrepresented sequences

Sequence	Count	Percentage	Possible Source
AAGCAGTGGTATCAACGCAGAGTACATGGGAAGCAGTGGTATCAACGCAG	133110634	34.31777695537423	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
GCAGTGGTATCAACGCAGAGTACATGGGAAGCAGTGGTATCAACGCAGAG	16614963	4.283569067485589	Clontech SMARTer II A Oligonucleotide (100% over 23bp)
GTGGTATCAACGCAGAGTACATGGGAAGCAGTGGTATCAACGCAGAGTAC	10274903	2.649013221528984	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
AGCAGTGGTATCAACGCAGAGTACATGGGAAGCAGTGGTATCAACGCAGA	7891906	2.0346433768828684	Clontech SMARTer II A Oligonucleotide (100% over 24bp)
AAGCAGTGGTATCAACGCAGAGTACATGGGAAAGCAGTGGTATCAACGCA	6824090	1.759345527145485	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
CAGTGGTATCAACGCAGAGTACATGGGAAGCAGTGGTATCAACGCAGAGT	5766298	1.4866320043387404	Clontech Universal Primer Mix Long (100% over 23bp)
AAGCAGTGGTATCAACGCAGAGTACATGGGGAAGCAGTGGTATCAACGCA	5316889	1.3707680995530582	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	3992839	1.0294095528139358	No Hit
AAGCAGTGGTATCAACGCAGAGTACATGGGAGCAGTGGTATCAACGCAGA	3941775	1.0162445418017485	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
AGTGGTATCAACGCAGAGTACATGGGAAGCAGTGGTATCAACGCAGAGTA	3280388	0.8457297537256577	Clontech SMARTer II A Oligonucleotide (100% over 24bp)
AAGCAGTGGTATCAACGCAGAGTACATGGGAAAAGCAGTGGTATCAACGC	2751059	0.7092613588864348	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
GGTATCAACGCAGAGTACATGGGAAGCAGTGGTATCAACGCAGAGTACAT	2722544	0.7019097943984879	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
ATCAACGCAGAGTACATGGGAAGCAGTGGTATCAACGCAGAGTACATGGG	2571647	0.6630064443533285	Clontech Universal Primer Mix Long (96% over 26bp)
AAGCAGTGGTATCAACGCAGAGTACATGGGAAAAAAAAAAAAAAAAAAAA	2046346	0.52757652406285	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
AAGCAGTGGTATCAACGCAGAGTACATGGGGTGGTATCAACGCAGAGTAC	1684172	0.43420301829894764	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
AAGCAGTGGTATCAACGCAGAGTACATGGGGTCAGATGTGTATAAGAGAC	1596381	0.4115692747267454	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
AAGCAGTGGTATCAACGCAGAGTACATGGGAGTGGTATCAACGCAGAGTA	1463644	0.3773478258248831	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
AAGCAGTGGTATCAACGCAGAGTACATGGGGCAGTGGTATCAACGCAGAG	1453150	0.37464232634262756	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
AAGCAGTGGTATCAACGCAGAGTACATGGAAGCAGTGGTATCAACGCAGA	1334951	0.3441689764947989	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
AAGCAGTGGTATCAACGCAGAGTACATGGGTGGTATCAACGCAGAGTACA	1242332	0.32029050722216507	Clontech SMARTer II A Oligonucleotide (100% over 25bp)

Figure 4.1: FastQC and MultiQC output reports.
a. Mean quality scores (Phred scores) for each sample across read sequences. The quality for each read and sample were overall good (within the green zone). b. Base content across read sequences. There is a high representation of specific bases along the read sequences. c. Percentage sequence GC content. The GC content for each sample is not normally distributed, suggesting sequence contamination or over representation of sequences. Each line represents a sample. d. Overrepresented sequences. A high level of overrepresented sequences contained the "Clontech SMARTer II A Oligonucleotide" adapter sequence.

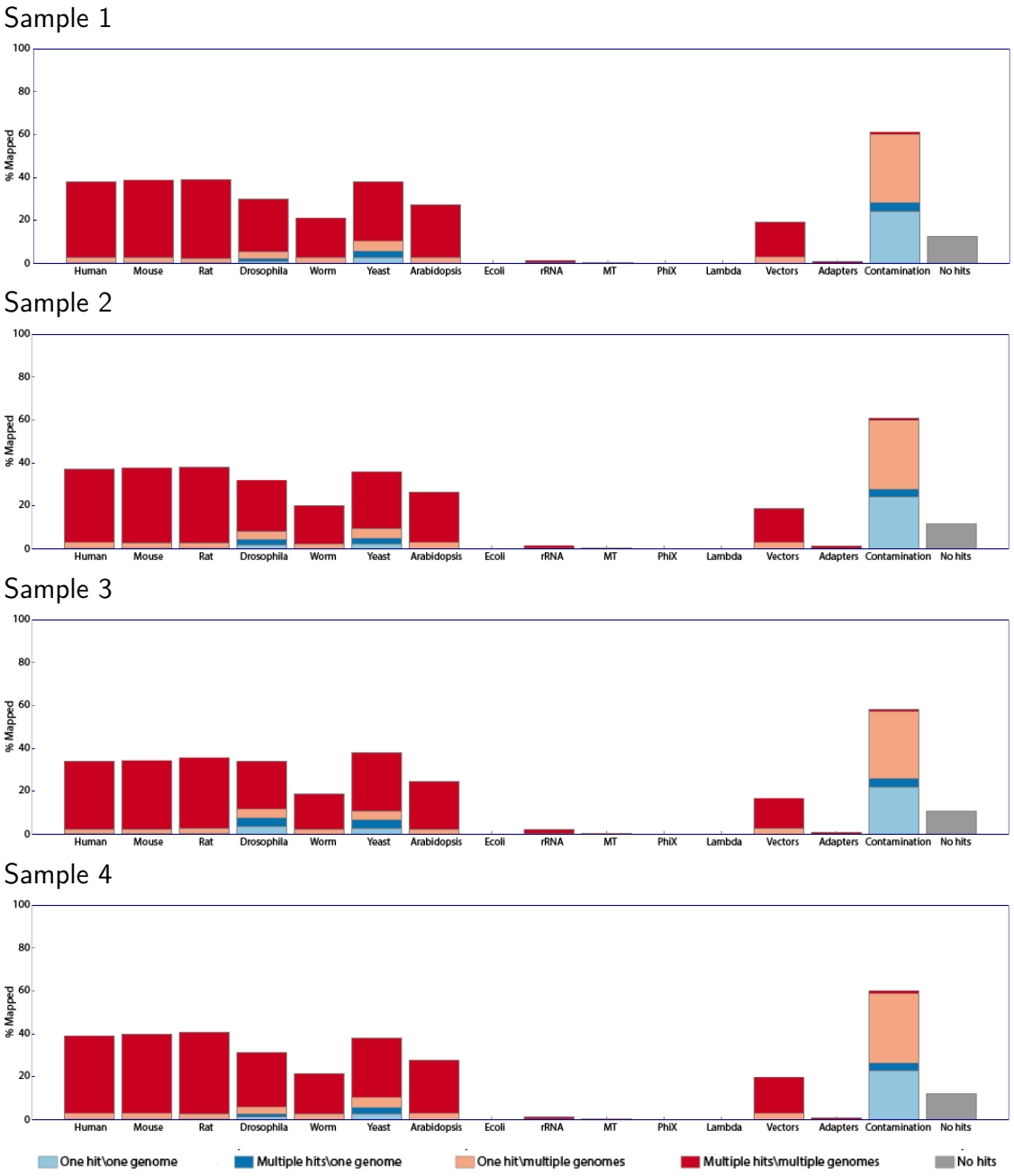


Figure 4.2: High level of sequence contamination in reads, from samples 1-4. Fastq Screen analysis comparing reads to different reference genomes. The "Contamination" column is the Clontech SMARTer II A Oligonucleotide sequence. For all samples, ∼22.7% of reads are contaminated with the Clontech adapter. ∼1.9% of reads mapped to the *Drosophila* reference genome and ∼2.6% mapped to the yeast genome. See the appendix Figures A.2 for FastQ Screen plots for the other samples.

4.2.2 Few Reads Mapped to the *Drosophila melanogaster* Reference Genome

Cell Ranger was used to align sample reads to the *Drosophila* reference genome (dmel650). Between 1.8% and 5.5% reads mapped to the reference genome and between 1.5% and 4.9% of reads mapped to the reference transcriptome (see Table 4.1). The low percentage of reads mapping to the reference transcriptome is likely a reflection of reads that were contaminated with the "Clontech SMARTer II A Oligonucleotide" adapter sequence.

Barcode rank plots from Cell Ranger identify barcodes associated with likely cells and barcodes associated with background noise (empty droplets containing ambient RNA). Figure 4.3 is a ranked barcode by UMI counts plot. These plots will typically have a "knee" or "cliff" when the number of UMIs drops, separating cells from empty droplets. The number of cells for Cell Ranger to expect was set to 10,000, this captured most samples just after the "knee"/"cliff" in the data (see Figure 4.3). For some samples, the cut-off point includes barcodes with few UMI. Barcodes associated with low UMI counts were filtered out in section 4.2.3.

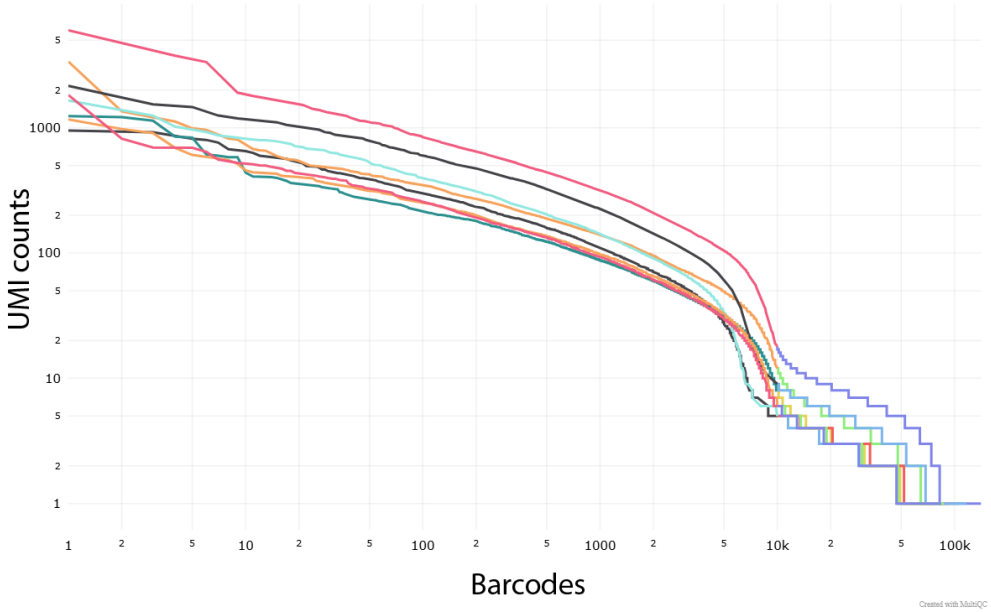


Figure 4.3: Barcode by UMI counts. The cut off for number of barcodes retained in each sample was 10,000. This captured most samples towards the end of the "knee" or "cliff" which is characteristic of the separation between barcodes associated with cells and those associated with empty droplets. Each line represents a sample.

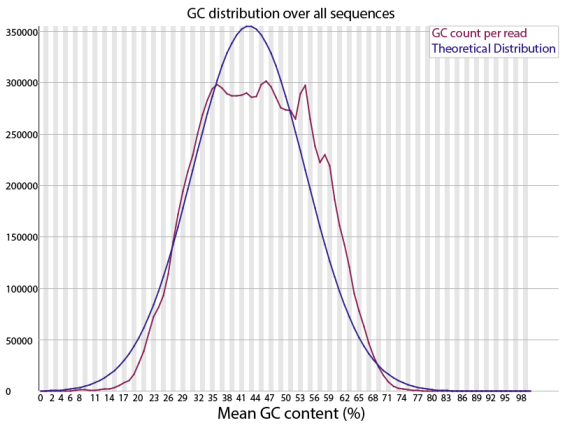
Table 4.1

Sample	**Percentage of reads mapping to the reference genome.**	**Percentage of reads mapping to the reference transcriptome.**
Sample 1	1.8%	1.5%
Sample 2	3.0%	2.6%
Sample 3	5.5%	4.9%
Sample 4	2.0%	1.7%
Sample 5	2.3%	1.9%
Sample 6	3.9%	3.5%
Sample 7	1.9%	1.6%
Sample 8	1.9%	1.6%

Percentage of reads mapping to the reference genome or transcriptome per sample.

Cell Ranger BAM output files contain additional bitwise flags which provide information about which reads are included in the UMI count. Cell Ranger was effective at identifying good quality reads as the population of reads identified as "confidently mapped to the transcriptome" had an improved GC content (Figure 4.4a) and a greater representation of reads mapping exclusively to the *Drosophila* genome (Figure 4.4d). Some of the mapped reads contained the Clontech SMARTer II A Oligonucleotide sequence (Figure 4.4b), however, these were mostly at the start of the reads (Figure 4.4c), and Cell Ranger 'soft clips' these sequences before alignment.

(a)

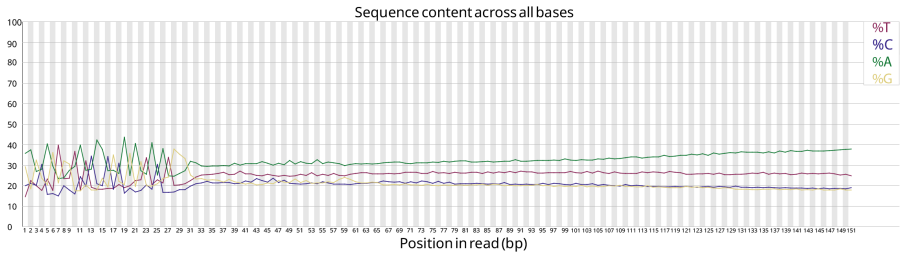


(b)

Overrepresented sequences

Sequence	Count	Percentage	Possible Source
AAGCAGTGGTATCAACGCAGAGTACATGGGAAGCAGTGGTATCAACGCAG	104194	1.0194210900626257	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
GCAGTGGTATCAACGCAGAGTACATGGGAAGCAGTGGTATCAACGCAGAG	47663	0.46632884250201473	Clontech SMARTer II A Oligonucleotide (100% over 23bp)
GTGGTATCAACGCAGAGTACATGGGAAGCAGTGGTATCAACGCAGAGTAC	30469	0.2981048927300818	Clontech SMARTer II A Oligonucleotide (100% over 25bp)
AGCAGTGGTATCAACGCAGAGTACATGGGAAGCAGTGGTATCAACGCAGA	20787	0.20337741327842102	Clontech SMARTer II A Oligonucleotide (100% over 24bp)
CAGTGGTATCAACGCAGAGTACATGGGAAGCAGTGGTATCAACGCAGAGT	16521	0.16163940177864983	Clontech Universal Primer Mix Long (100% over 23bp)

(c)



(d)

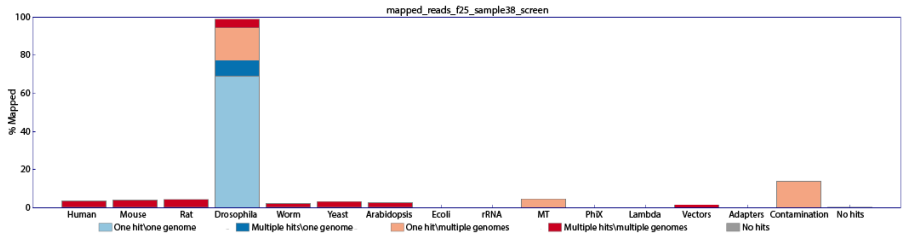


Figure 4.4: Cell Ranger improved the representation of good quality reads. FastQC and FastQ Screen output plots from reads tagged with the bitwise flag xf25 in the Cell Ranger output BAM files. a. GC content of mapped reads. The GC content is more normally distributed. b. The Clontech SMARTer II A Oligonucleotide sequences are found in less than 3% of reads. c. Sequence content across all bases is biased to the Clontech SMARTer II A Oligonucleotide sequence at the start of reads. d. $\sim$ 70% of reads were mapped exclusively to the *Drosophila melanogaster* genome.

4.2.3 Cell Quality Analysis

To further investigate cell quality, read counts from Cell Ranger for each sample were loaded into R as Seurat objects (see Methods). Park et al., (2022) filtered out cells that had <300 or >4500 features or >25,000 UMIs. Feature and UMI counts for the dataset used in this report fall well below these thresholds (Figures 4.5a and 4.5b). I, therefore, adjusted the thresholds manually to filter cells based on our data. The dotted lines in Figures 4.5a and 4.5b show the thresholds for filtering. The cells were filtered for cells that had >200 and <2500 features and <8000 UMIs. This filtered out 28535 cells. 33 cells had greater than 2500 features, and 76 cells had more than 8000 UMIs. Like Park et al., (2022), I also filtered out cells with >15% mitochondrial RNA and ribosomal proteins, and cells with >10% rRNA (see Methods, and Appendix Figures A.3a and A.3b). Reviewing the distribution of the percentage of reads that are mitochondrial by UMI count shows that cells with greater than >15% of mitochondrial RNA are unlikely to be healthy cells, as many have low UMI counts but high mitochondrial counts (Figure 4.5d). After filtering, between 2672 and 6618 cells per sample remain (Table 4.2).

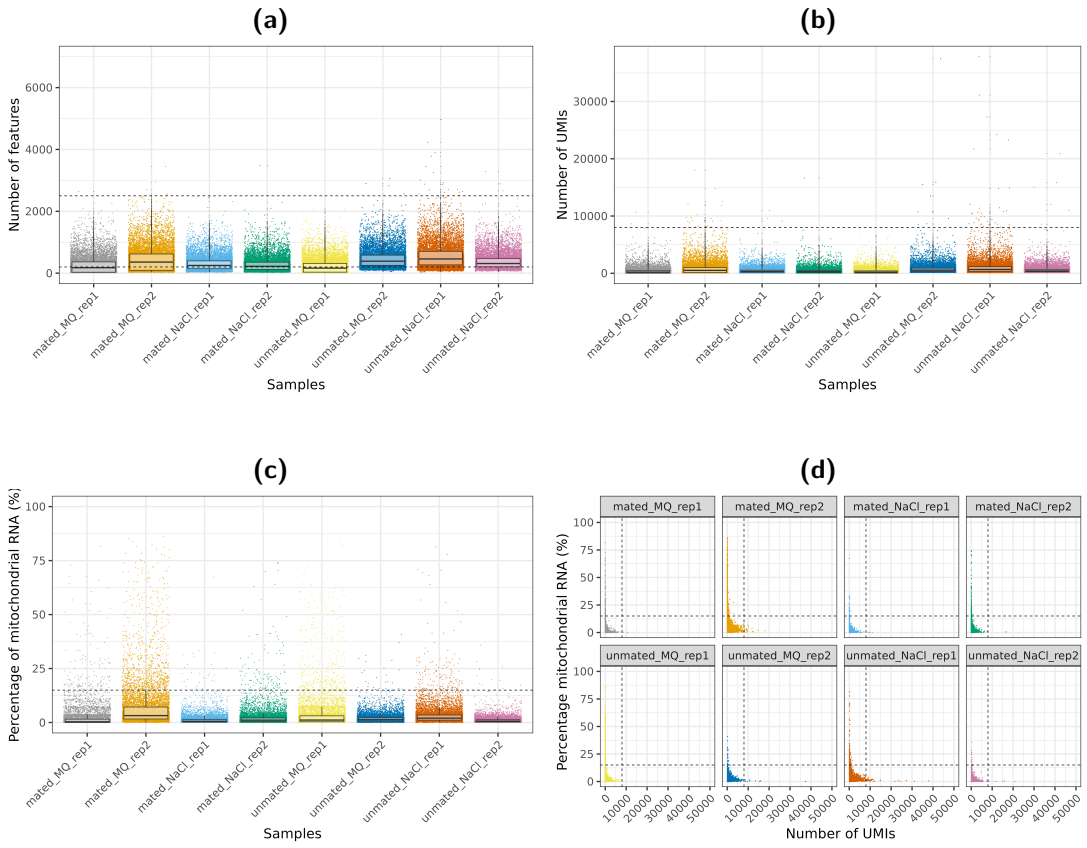


Figure 4.5: Reviewing quality of cells for each sample.
a. Number of features per cell for each sample. Dotted lines show the lower (200) and upper (2500) threshold cut-off values. b. Number of UMIs per cell for each sample. The dotted line shows the cut-off threshold of 8000. c. Percentage of mitochondrial RNA (%) per cell for each sample. The dotted lines show the cut-off threshold of 15%. d. Percentage mitochondrial RNA (%) by UMI counts. The cut-off threshold for cells with >15% mitochondrial reads appears to remove many cells with low UMI counts but a high representation of mitochondrial reads.

Table 4.2

Sample	Number of cells	Median UMI	Median feature count
20230110 S1 unmated MQ rep1	2672	459	341.5
20230110 S1 unmated MQ rep2	6618	621	432.5
20230110 S3 unmated NaCl rep1	7055	777	516
20230110 S4 unmated NaCl rep2	5213	505	371
20230110 S5 mated MQ rep1	3236	515	377
20230110 S6 mated MQ rep2	5458	824	536
20230110 S7 mated NaCl rep1	3887	471	347
20230110 S8 mated NaCl rep2	3281	450.5	334

Total number of cells, median number of UMIs and feature counts per sample.

4.2.4 Sample Integration and Clustering Analysis

Replicates from each condition were merged before the samples were integrated together (see Methods section 2.8.1). Clustering using the first 40 principle components (PCs) and a clustering resolution of 0.1 identified seventeen clusters (Figure 4.6a). PCs were selected using an Elbow plot (see Methods); 1:40 PCs seemed to explain most of the data's variance. Selecting a greater number of PCs would likely increase noise in the data. Integrating the samples together appeared to achieve a good representation of each sample within each cluster (Figure 4.6b).

4.2.5 Identification of Male and Female Cells

As I was interested in the genes that might be differentially expressed in NaCl-fed compared to nutrient-deprived female *Drosophila*, I separated males and female *Drosophila* cells by the expression of the male sex-specific gene *lncRNA:roX1* (Kelley et al. 2003). Plotting the distribution of module score values for *lncRNA:roX1* identified two populations of cells (Figure 4.7a). Cells with a module score >0 were classified as male, and the other cells were classified as female. Plotting the distribution of cells classified as male and female in UMAP space shows good integration of these cells for most clusters (Figure 4.7b). Cluster 6 showed some male and female separation (Figure 4.7b).

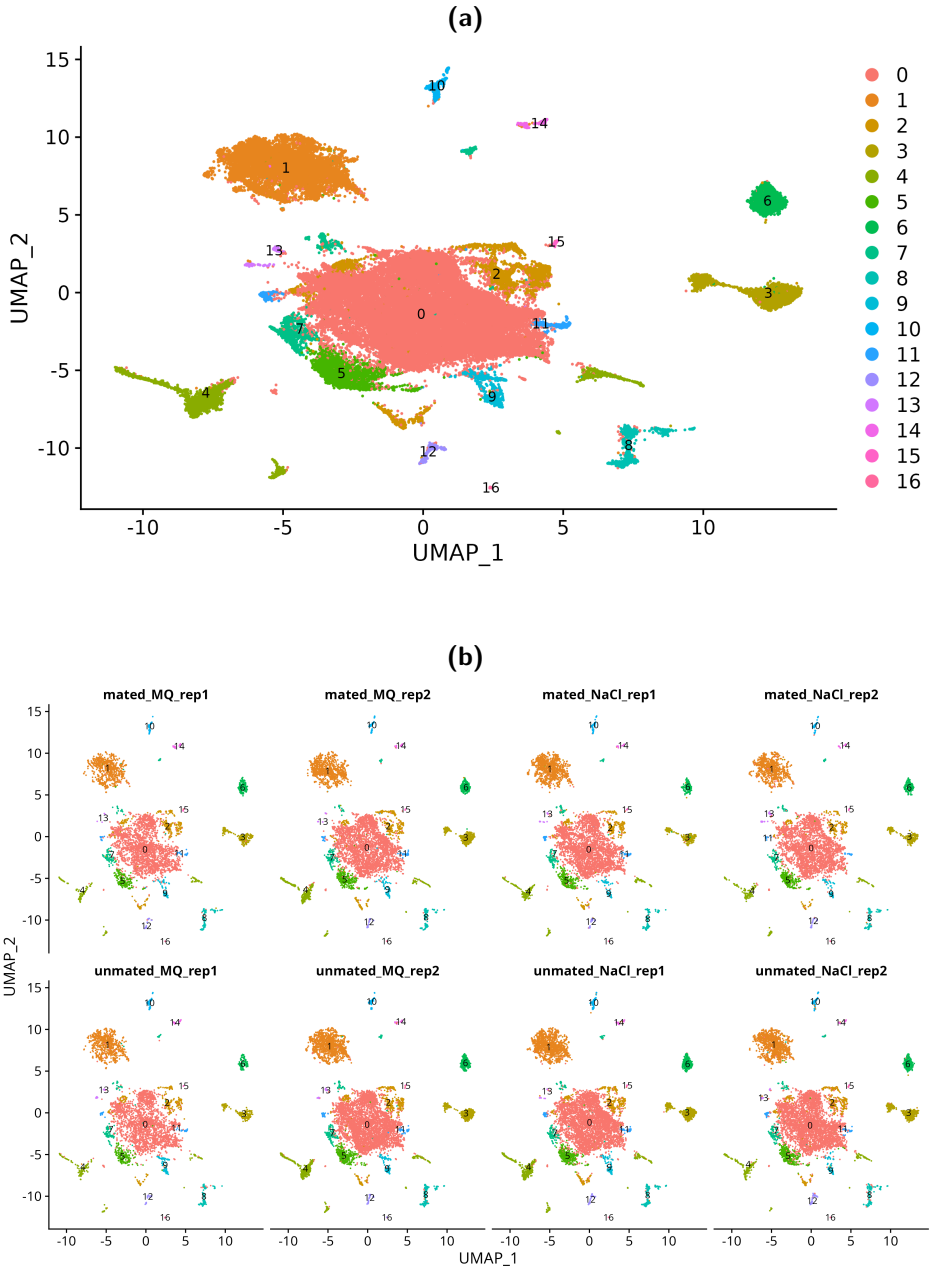


Figure 4.6: UMAP plot of the integrated samples.
a. Integrating the samples and clustering at a resolution of 0.1 revealed 17 clusters. b. Separating the integrated data in UMAP space by sample identity shows a good representation of each sample within each cluster.

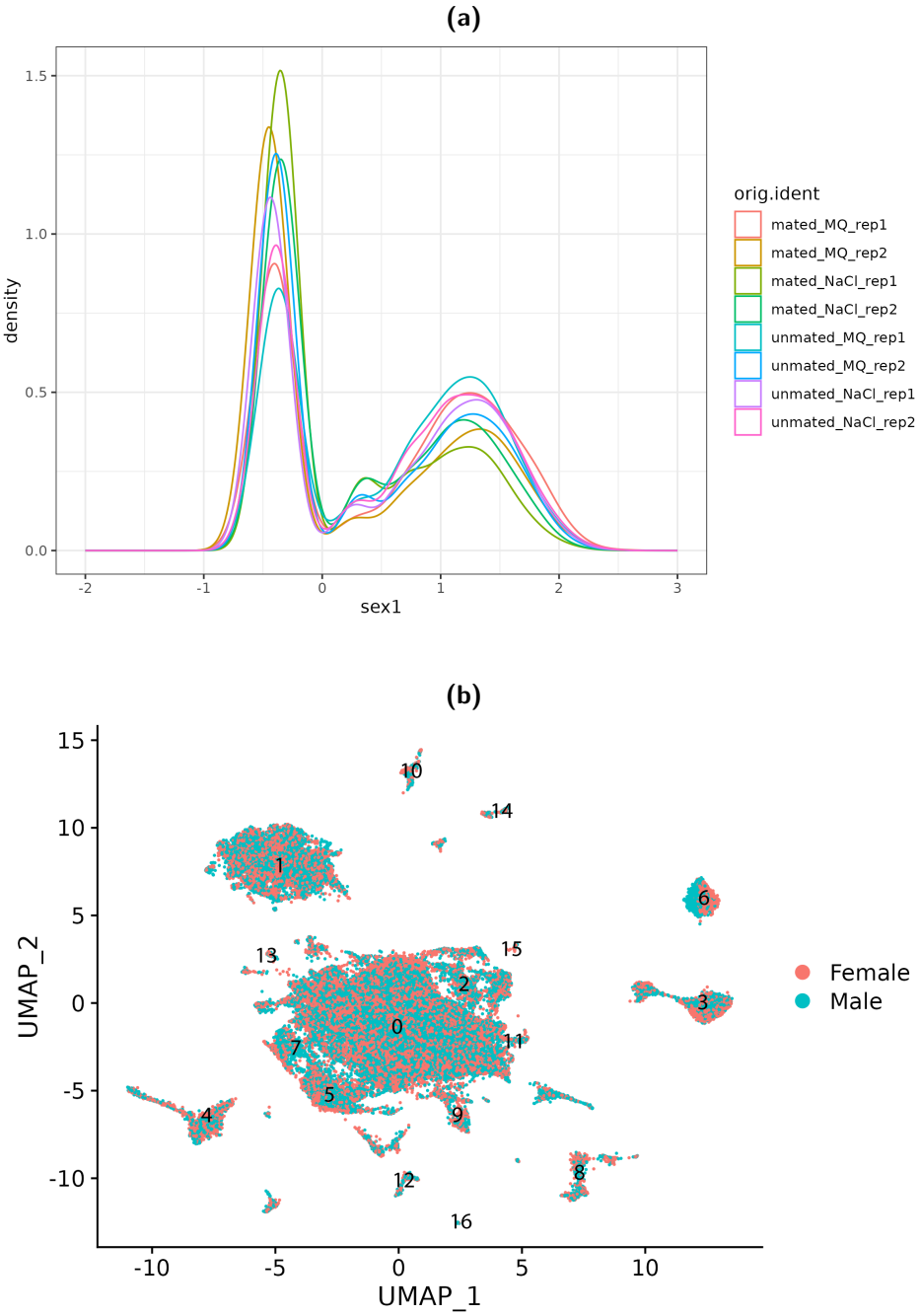


Figure 4.7: Identifying male and female cells for each sample.
a. Density plot of *lncRNA:roX1* module scores for each sample. Cells with a *lncRNA:roX1* module score (labelled 'sex1') greater than zero were classified as male. b. The distribution of cells classified as male and female in UMAP space. Male and female cells seem to be distributed well for most clusters. Cluster 6 appears to be split by male and female cells in UMAP space.

4.2.6 Transfer Labels Using SingleR

Due to the low feature and UMI numbers per cell for each sample, there was likely some dropout of key marker genes for annotating clusters. To overcome this challenge, I employed the label transfer tool, SingleR, to transfer the cell identity labels from the Park et al., (2022) dataset onto the data outlined in this report (referred to as the "salt dataset"). This tool transfers labels by considering and comparing the expression profiles of each cell type between the datasets and is not reliant on single marker genes (Aran et al. 2019). Out of the 184 possible cluster labels from the Park et al., (2022) dataset, 171 were transferred over, with 156 labels associated with three or more cells. Grouping cells based on major cell type classes, such as cholinergic, GABAergic, glial, monoaminergic and Kenyon Cell types found that the ratios of cell type numbers were similar between the salt dataset and the Park et al., (2022) dataset (Figures 4.8a and 4.8b). However, there appears to be a greater representation of cholinergic cells in the salt dataset, and a lower representation of glial cell types. Appendix Figure A.4 created from the Park et al., (2022) dataset shows that glial cells have a lower UMI count compared to other cell types. It is possible that the glial cell types were filtered out when filtering for cells with greater than 200 UMI counts. It is also possible that there was an over-classification of cholinergic cells. This could be due to ambient RNA contamination. Further analysis using the contamination evaluation tools, such as DecontX (Yang et al. 2020), could be valuable to understand if ambient RNA has influenced cluster labelling. The male and female cell contributions were consistent across the different cell types (Appendix Figure A.5).

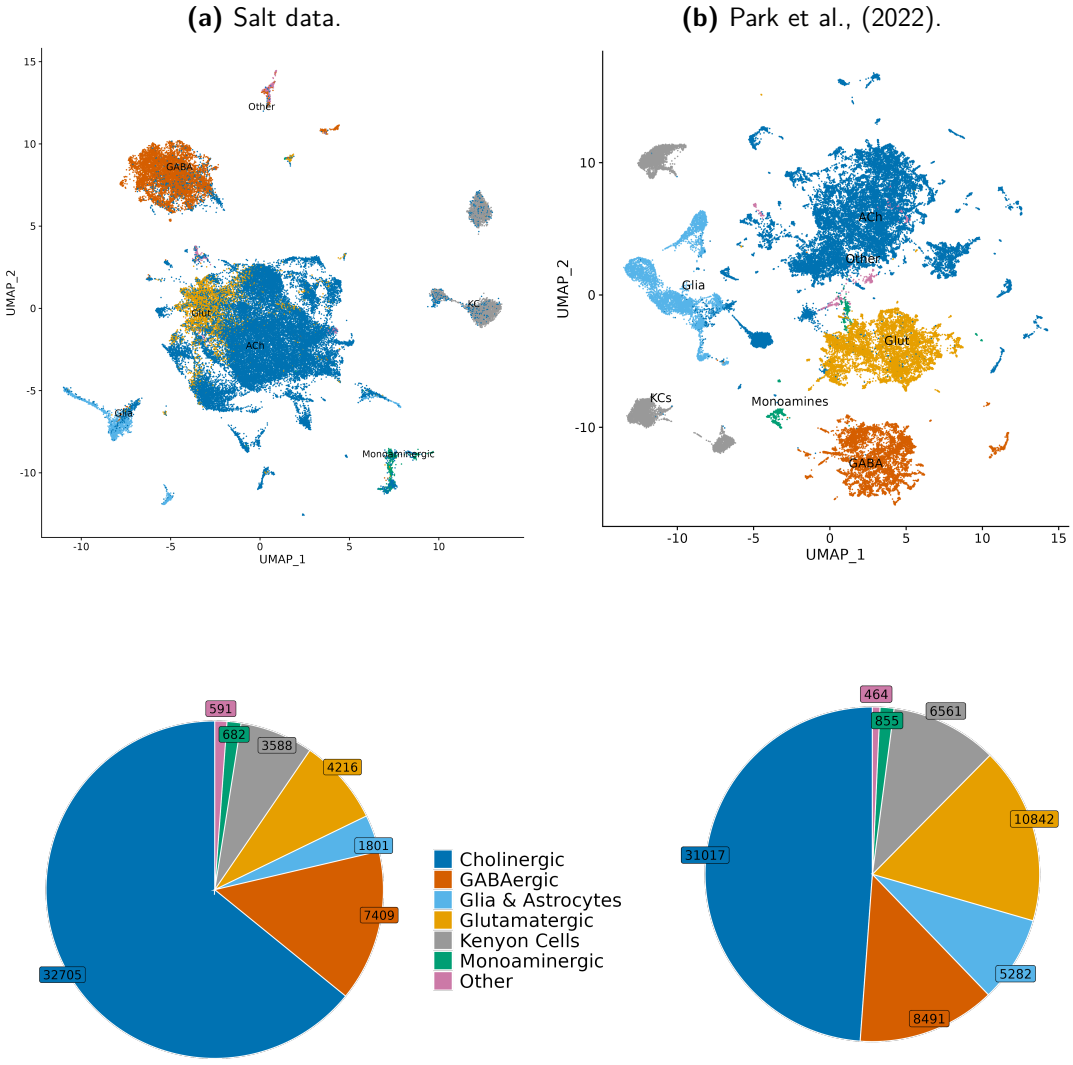


Figure 4.8: Major cell types labelled in UMAP space showing labels transferred from Park et al., (2022, b) to the salt data (a) and cell type proportions (bottom, right and left, respectively).
a-b. UMAP plot and pie chart representing the major cell types within the salt dataset, (a) and the Park et al., (2022) dataset, (b).
Cholinergic cells: blue, GABAergic cells: dark orange, Glia and Astrocytes: light blue, Glutamateric cells: light orange, Kenyon Cells: grey, Monoaminergic cells: green and other: pink. The proportion of the different major cell types were similar between the two datasets.

4.2.7 The Proposed DGE Setup was Sufficient to Identify DGE Events From Published Work

Due to the low UMI and feature counts in the salt dataset (Table 4.2), it is possible that important gene expression was missed in many cells. Therefore, the data for each cluster was aggregated by sample to create a 'pseudobulk' dataset. Pseudobulk data preserves cell-type information whilst allowing the use of bulk differential expression methods like DESeq2 (Love et al. 2014). This approach reduces the risk of false positives by treating individual samples as replicates, unlike single-cell methods that compare individual cells (Squair et al. 2021). To understand the pseudobulk workflow, I performed differential gene expression (DGE) analysis on a good quality scRNA-seq dataset published by Park et al. 2022 which had a similar experimental design. For this analysis, I aggregated the single-cell count data by sample replicate to create a pseudobulk count matrix for each cell-type sub-cluster (see Methods). I used DESeq2 to test DGE between the experimental conditions - twelve-hour dehydrated *Drosophila* compared to satiated *Drosophila*.

The output from DESeq2 DGE analysis from subclusters was grouped into volcano plots for each major cell type (as was done Park et al. 2022). Similar to the results from Park et al., (2022), the greatest number of DGE events were found in glial cells (Figure 4.9a). I also found that *aay* and *Drat* were upregulated in dehydrated *Drosophila* compared to satiated *Drosophila* in some glial cell clusters (Figure 4.9a).

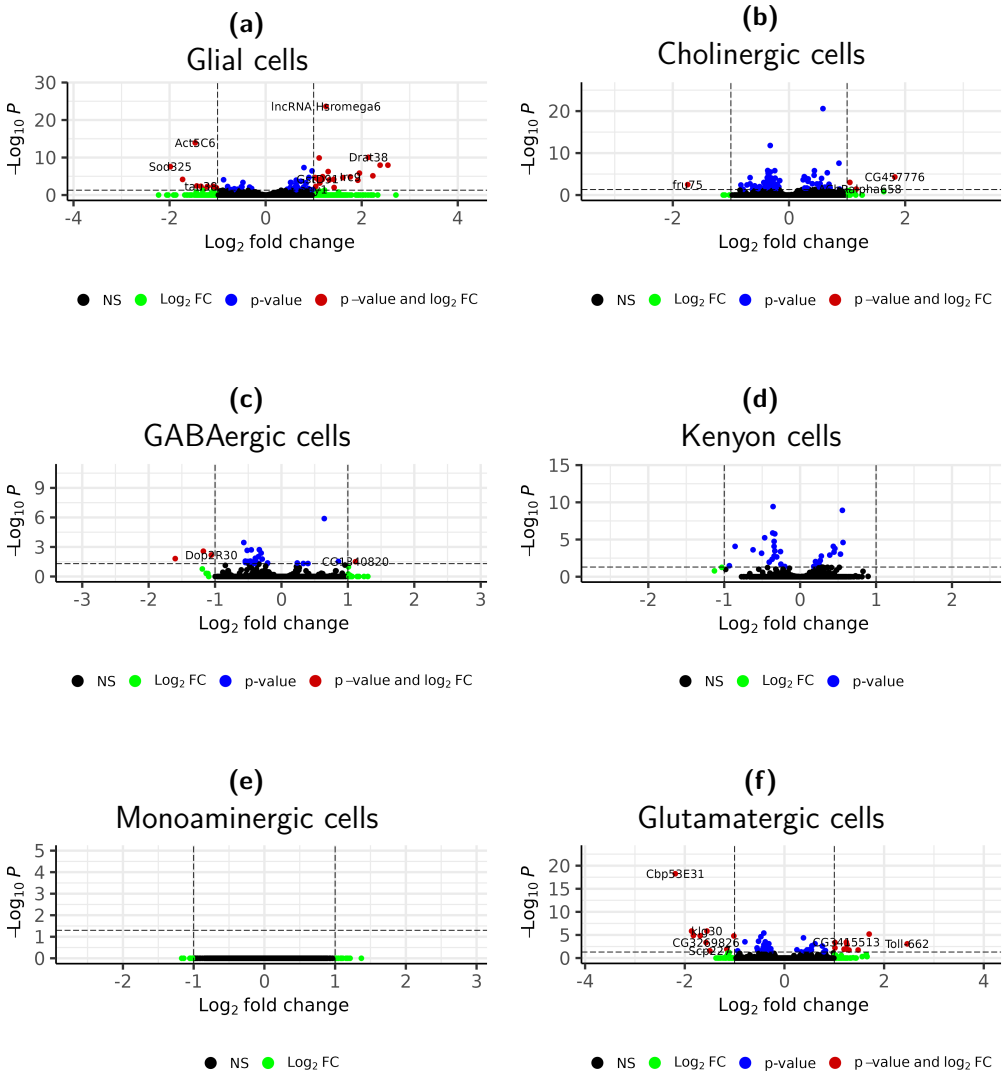


Figure 4.9: Differential gene expression analysis of Pseudobulk data from Park et al., (2022). Performing differential expression analysis using a modified workflow published by Park et al., (2022) reproduced similar results. When comparing gene expression between satiated *Drosophila* (no water deprivation, represented on the left-hand side of the volcano plots) compared to 12-hour dehydrated *Drosophila* (right side of the volcano plots), there were greater DGE events in the glial cell clusters compared to the alternative cell types. -$\text{Log}_{10}P$ values were adjusted -$\text{Log}_{10}P$ values, this is true for the following volcano plots.
a. There were 31 DGE events in the glial cell clusters. b. There were 4 DGE events in cholinergic cell clusters. c. There were 4 DGE events in GABAergic cell clusters. d. There were no DGE in Kenyon Cells. e. There were no DGE events in monoaminergic cell clusters. f. There were 19 DGE events in Glutamatergic cell clusters.

4.2.8 No Detectable Transcriptomic Changes were Found Between Nutrient-Deprived and NaCl-Fed *Drosophila*

Performing the same workflow as above, but controlling for sex and mating status, I looked for differentially expressed genes within each cell sub-cluster, between *Drosophila* fed 0 M NaCl and *Drosophila* fed 0.2 M NaCl. This analysis did not find differentially expressed genes with a |log2FC| >1 and an adjusted p-value <0.05 (Figures 4.10a to 4.10f, 0 M NaCl feeding condition on the left and the 0.2 M NaCl condition on the right).

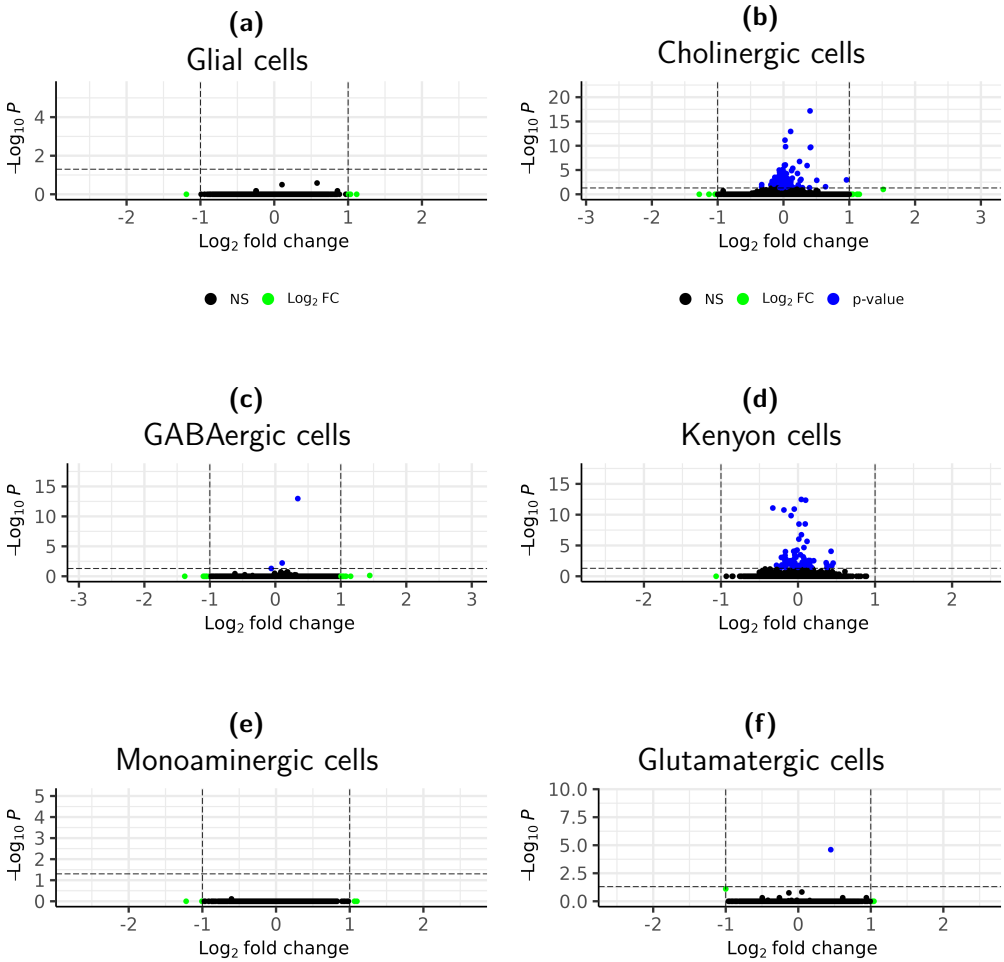


Figure 4.10: Volcano plots of DESeq2 output results for pseudobulk data for each cell type, comparing *Drosophila* fed 0.2 M NaCl to 0 M NaCl. Output from DESeq2 analysis identified no genes with an adjusted p-value <0.05 and |log2FC| >1 from any cell cluster types.

4.2.9 *Myosuppressin* was Differentially Expressed in a Subset of Cholinergic Cells from Female *Drosophila*

Next, I tested DGE between the different feeding conditions in female *Drosophila*. The overall results from this analysis found lower adjusted p-values across all cell types despite the lower number of cells (Figure 4.11a to 4.11f). One feature was classed as statistically differentially expressed. *Myosuppressin* (*Ms*), of the cluster labelled 22-ACh, was upregulated in the 0 M NaCl feeding condition compared to the 0.2 M NaCl condition (Figure 4.11b). *Ms* is expressed by medial neurosecretory cells (m-NSCs) in the adult brain (McKim et al. 2024). Exploring the difference in pseudobulk counts between the two feeding conditions shows that cells from the 0 M NaCl feeding condition have higher pseudobulk count values for *Ms* compared to the 0.2 M NaCl feeding condition (Figure 4.12). There is a difference in counts between the mating status of female *Drosophila* fed 0 M NaCl (see Figure 4.12), however, this difference could be due to dropout of this feature within the mated samples. This difference was also seen when plotting a SCTransformed version of the original count data, Figure A.6a shows a greater expression of *Ms* in nutrient-deprived female *Drosophila* compared to NaCl-fed, with greater counts in the un-mated condition. Many cells appear to have zero count values for *Ms*. RT-qPCR could be used to confirm the increase in *Ms* expression in nutrient-deprived *Drosophila* in future experiments.

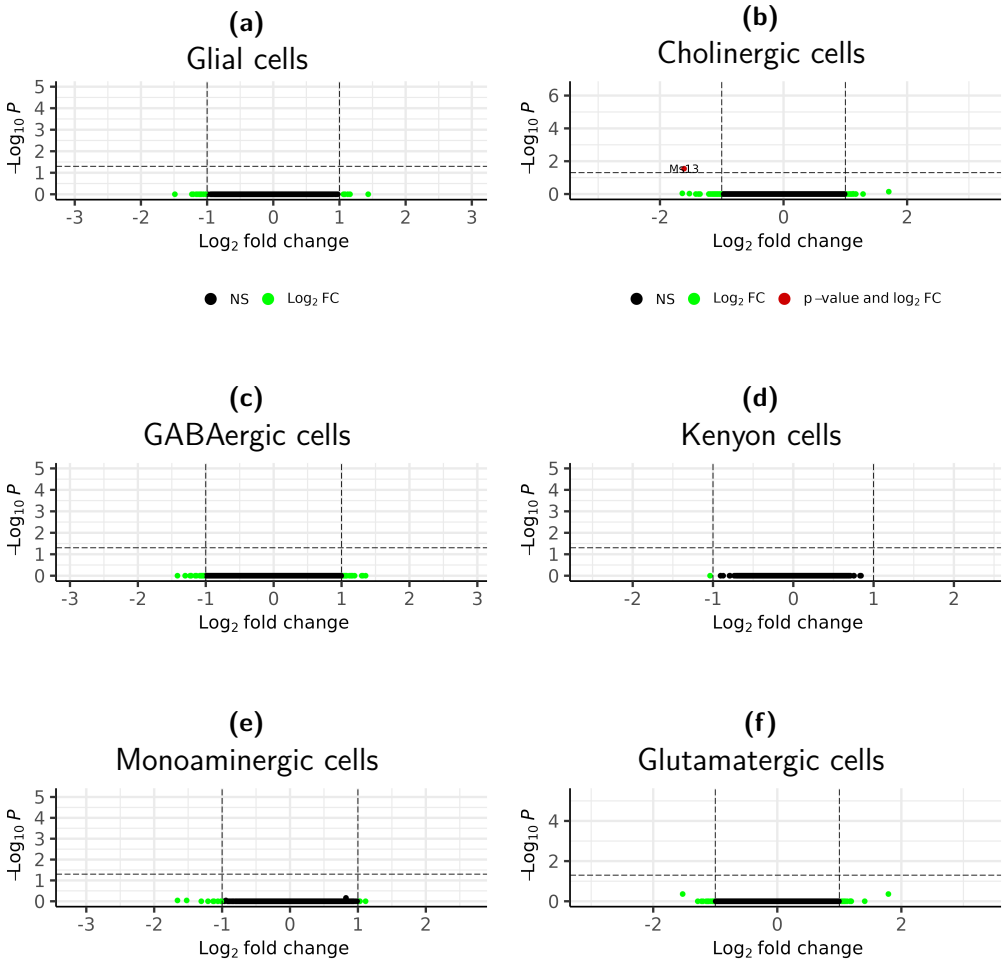


Figure 4.11: Volcano plots of DESeq2 output results for the pseudobulk data for each cell type, comparing female *Drosophila* fed 0.2 M NaCl to female *Drosophila* fed 0 M NaCl. Compared to Figures 4.10a to 4.10f, overall adjusted p-values in plots a-g were reduced. *Ms* expression in cluster 22-ACh was greater in female *Drosophila* fed 0 M NaCl with a log2FC of 1.61 and adjusted p-value of 0.029.

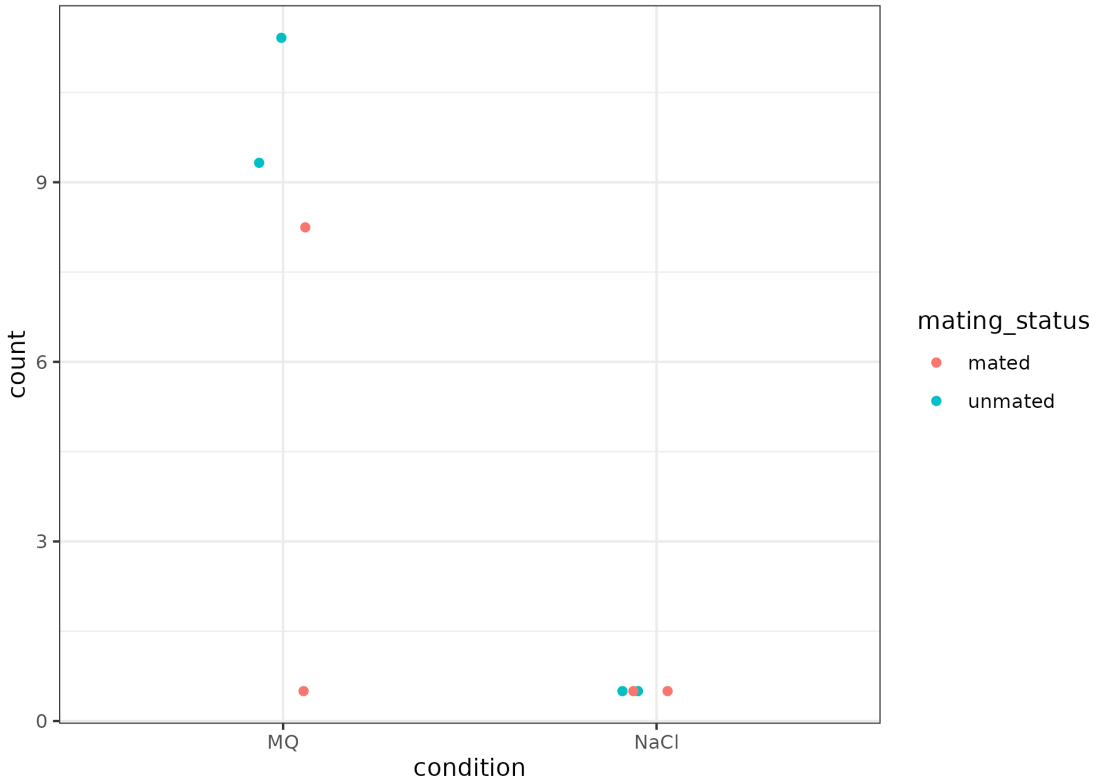


Figure 4.12: Pseudobulk *Ms* counts from cluster 22-ACh. There were greater pseudobulk counts in cholinergic cells of cluster 22-ACh in *Drosophila* that were nutrient-deprived (0 M NaCl, MQ) compared to those that were fed NaCl (NaCl). Pseudobulk counts for mated *Drosophila*: pink, pseudobulk counts for unmated *Drosophila*: blue. One replicate for nutrient-deprived mated *Drosophila* is close to zero, this could be due to dropout.

4.3 Discussion

4.3.1 Results Summary

ScRNA-seq data from eight samples, of mated and unmated adult *Drosophila* that were either nutrient-deprived or fed 0.2 M NaCl, were of poor quality. Many sequenced reads did not map to the *Drosophila* reference genome, with 1.5-4.9% of reads from each sample being confidently mapped to the transcriptome. The poor read quality was attributed to the high level of Clonetech SMARTer II A Oligonucleotide adapter sequence contamination. Despite the low cell recovery per sample, the representation of the different cell types was good. *Ms* was identified as a differentially expressed gene when comparing nutrient-deprived to NaCl-fed female *Drosophila*.

4.3.2 Neuropeptide Expression and Sodium Status

Ms influences gut motility, food intake and defecation in *Gryllus bimaculatus* crickets (Zhou et al., 2019), as application of the *Ms* peptide in the cricket reduced foregut muscle contractions, but increased food intake and defecation. As mentioned in section 4.2.9, *Ms* is expressed by m-NSCs. Another group of m-NSCs express the *diuretic hormone 44* (*Dh44*, McKim et al. 2024), a neuropeptide that affects osmotic balance. Dh44 reduces fluid secretion by interacting with Dh44 receptors (Dh44-R2) expressed by Malpighian tube cells (as reviewed by Lee et al. 2023). A global RNAi knockdown of *Dh44-R2* reduced the survival rate of *Drosophila* exposed to a high NaCl concentration (0.6 M NaCl, Hector et al. 2009). When ordering the output from DESeq2 analysis, comparing 0 M NaCl fed to 0.2 M NaCl fed female *Drosophila*, by ascending adjusted p-values, *Dh44* was second after *Ms*. *Dh44* expression in the cluster labelled 20-Glut was greater in 0.2 M NaCl-fed female *Drosophila* compared to 0 M NaCl-fed female *Drosophila*. The log2FC value was 1.79, the p-value was <0.001, but the adjusted p-value was 0.43. It is possible that in NaCl-fed female *Drosophila*, there is a greater expression of *Dh44* to inhibit water secretion and balance osmotic homeostasis. This evidence could suggest that m-NSCs are involved in signalling sodium status. When plotting both pseudobulk and SCTransformed counts for *Dh44* in cluster 20-Glut, the

increase in *Dh44* expression in NaCl-fed female *Drosophila* seemed to be carried by the unmated NaCl-fed samples (Appendix Figure A.6b and A.6c). It is possible that *Dh44* signalling is different in mated and unmated female *Drosophila* in response to NaCl feeding. It would be important to validate the results in this chapter with further experiments, such as RT-qPCR to detect differences in neuropeptide expression in *Drosophila* brains after experiencing the feeding conditions used in this chapter.

4.3.3 Conclusion

Considering the level of contamination in this dataset and the reduction of power to detect DGE events between the two feeding conditions, it is challenging to draw conclusions from this analysis alone. The limited evidence points towards neuropeptide signalling from the m-NSCs of the *Drosophila* brain. Further research is required to understand the function of these neurones in sodium-sensing and how this might influence NaCl-associated learned behaviours.

Chapter 5

Transcriptomic Characterisation of Learning and Memory-Relevant Cells of *Apis mellifera* and Comparison to *Drosophila melanogaster*

Contents

5.1 Introduction

As outlined in Chapter 1 (section 1.2), *Apis mellifera* and *Drosophila melanogaster* are model organisms used for studying learning and memory in the lab. Although *Apis* and *Drosophila* diverged ∼300 million years ago (Maier et al. 2008), they still share common cell types that are involved in learning and memory. Both insects use the intrinsic neurones of the Mushroom bodies (MBs), the Kenyon Cells (KCs), and monoaminergic neurones, such as dopaminergic neurones (DANs), to coordinate associative learning and memory. However, there are some differences between the two species, such as the differing relative size and morphologies of the MBs (see section 1.2.1 and Figures 1.1a and 1.1b). In addition, the cellular circuits involved in learning and memory are better understood in *Drosophila* thanks to neuronal connectomic information and powerful genetic tools combined with behavioural experiments (Owald et al. 2015b; Scheffer et al. 2020; Li et al. 2020a). Therefore, the analysis reported in this chapter aimed to provide a detailed understanding of the transcriptomic profiles of KC and DANs in *Apis*, and compare these cells to those of *Drosophila*, with the objective of identifying any conserved or differing cell-type profiles between the species.

Apis have two classes of KCs, Class I KCs, which are subdivided into sKC, mKCs, lKCs and FoxP-KCs, and Class II KCs. *Drosophila* have three major KC types; γ KCs which are further divided into γd and γm KCs, $\alpha\beta$ KCs are divided into $\alpha\beta$p, $\alpha\beta$s and $\alpha\beta$c KCs, and $\alpha'\beta'$ KCs are divided into $\alpha'\beta'$ap and $\alpha'\beta'$m KCs (Aso et al. 2014; Shih et al. 2018). In *Drosophila* the $\alpha\beta$ and $\alpha'\beta'$, and in *Apis* the Class I KC axons bifurcate along the horizontal and vertical MB lobes in *Drosophila* and vertical and medial MB lobes in *Apis* (Strausfeld 2002; Lee et al. 1999). *Drosophila* γ KCs and *Apis* Class II KCs do not bifurcate, they project their axons to the γ lobe. Marker genes for individual subtypes of *Apis* Class I KCs have been described by Suenami et al., (2018), but there are currently no discrete marker genes for Class II KCs.

Dopaminergic neurones (DAN) provide value information to reinforce olfactory associative memories. In *Drosophila*, DAN clusters PPL1 and PAM reinforce aversive and appetitive memories, respectively (see section 1.2.3). In *Drosophila* DANs tile the MBs, innervating discrete, non-overlapping, compartments of the horizontal lobe (PAM-DANs) and vertical lobes (PPL1), see Figure 1.1a (Tanaka et al. 2008; Aso et al. 2014; Bornstein et al. 2021). In *Apis* there are four main DAN clusters, C1 - C4 (see Figure 1.1b). C1 and C2 are suggested to correspond to *Drosophila* PAM-DANs, and C3 is suggested to correspond with *Drosophila* PPL1, PPM3 and PPL2ab-DANs based on anatomy (Tedjakumala et al. 2017). In *Drosophila*, there are 12 PPL1, 6∼8 PPM3, 6 PPL2ab and ∼100 PAM-DANs per hemisphere (Mao & Davis, 2009). In *Apis*, there are about 75 cells in C1 and C2, and 140 cells in C3 (Tedjakumala et al., 2017). *Apis* C3 might be made of subsets of DANs (Tedjakumala et al. 2017).

Single-cell RNA sequencing (scRNA-seq) experiments of insect brains have been indispensable for an unbiased approach to cell type characterisation. In the analysis of scRNA-seq data from the midbrains of *Drosophila*, distinct clusters are found for the three major KC types, suggesting that they have a distinct transcriptomic profile (Croset et al. 2018; Park et al. 2022). ScRNA-seq data from whole adult *Apis mellifera* brains have been published by Zhang et al., (2022). Zhang et al., found clusters corresponding to Class I lKCs, sKCs and FoxP-KCs (Zhang et al., 2022, Figure 3A). However, they did not find Class I mKCs and had KC clusters left unidentified. Zhang et al., (2022), Figure 3A shows some additional granularity within the Class I KC subtypes, which suggests there might be additional heterogeneity to the *Apis* KC clusters than what was previously reviewed by Suenami et al., (2018). Zhang et al., (2022) also did not find discrete clusters dedicated to monoaminergic cells, cluster types that are commonly seen in scRNA-seq data from *Drosophila*. It is possible that there were too few monoaminergic cells to form a distinct cluster, and as Zhang et al., did not perform an analysis to review ambient RNA contamination; this data may have suffered from additional noise, thus preventing clustering of cell types with lower representation.

As discussed in this introduction section, a detailed understanding of the cell types involved in *Apis* associative learning is incomplete. This chapter outlines the scRNA-seq analysis of twelve pollen-forager *Apis mellifera* honey bees, using only the dorsal protocerebrum to increase the representation of KC and monoaminergic cell types. This chapter aims to characterise the transcriptomic profiles of the cell types involved in associative learning and to compare these clusters to KC and dopaminergic cell types in *Drosophila* using the scRNA-seq data published by Park et al., (2022).

5.2 Results

5.2.1 High GC Content in Raw Reads Was Likely Due to Sequence Duplication

As mentioned in the Methods section, there were two FASTQ files for each individual bee collected as the dorsal protocerebrum was too large to be processed at once. The two samples A and B per individual bee were processed in parallel and the data was recombined after QC (see section 5.2.3). FastQC and MultiQC were used to assess the quality of the FASTQ files output from scRNA-seq. Phred scores, a score quantifying base calling confidence at each position, were high across all reads for each sample (Figure 5.1a). A normal distribution of raw read GC content was observed in most samples. However, one of the FASTQ files from the initial round of data collection (October 2019) exhibited an additional peak at around 50% GC content (see Figure 1b). The GC content is assessed based on a reference distribution calculated from the modal GC content for each sample (Ewels et al. 2016). Distributions that deviate from a normal distribution may indicate overrepresented sequences or contamination. The increased GC content noted in this report is likely a consequence of a higher percentage of sequence duplication, with duplication rates ranging from 74.2% to 76.9% in these samples, compared to 50.1% to 70.4% in the other samples (data not shown). This observation is further supported by reviewing sequence duplication levels per sample. Figure 5.1c shows the sequence duplication levels for sample 12a, which has a normal distribution of GC content. These levels are much lower than sample 4a, which has a higher GC content 5.1d.

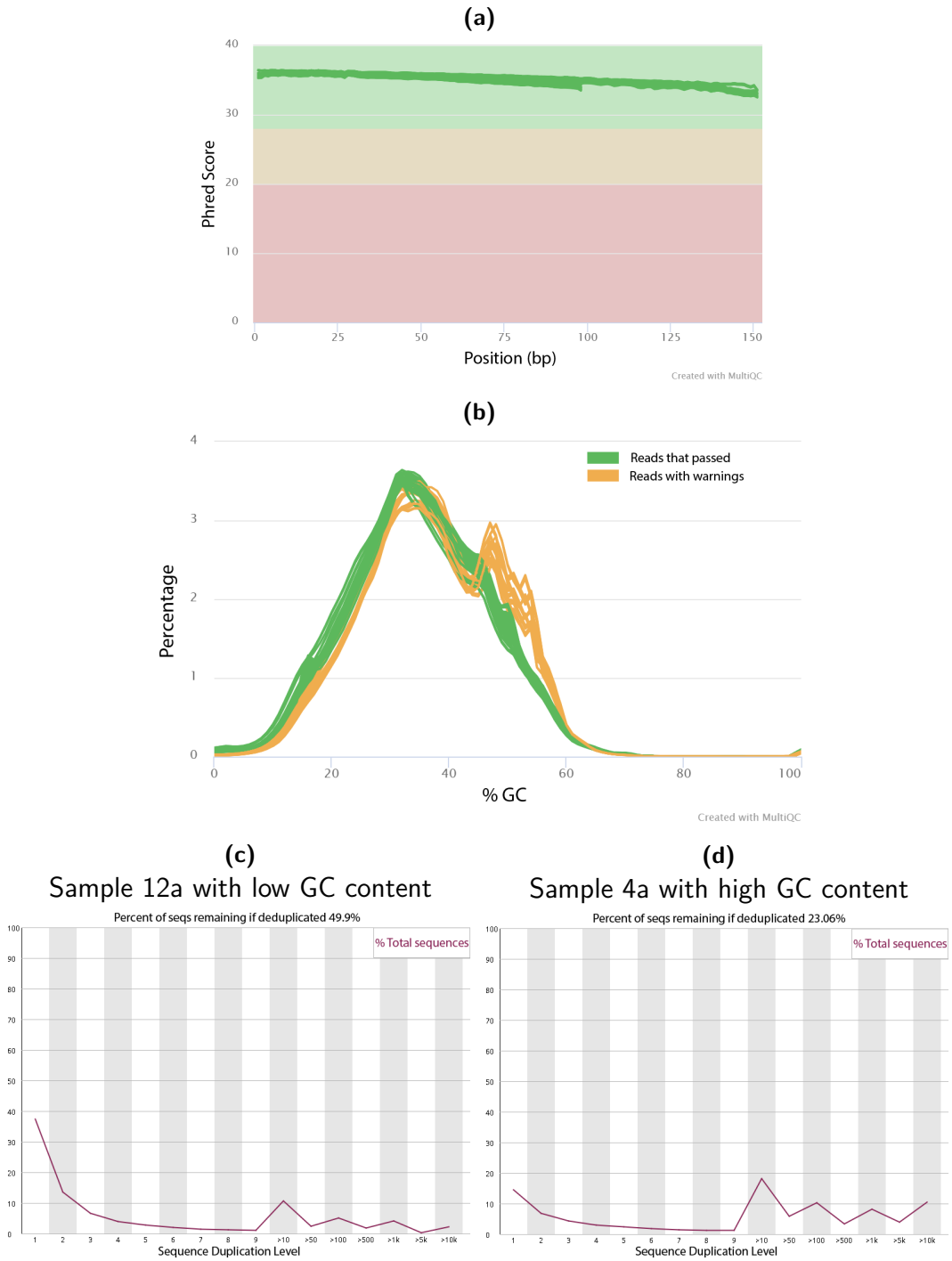


Figure 5.1: FastQC and MultiQC output reports.
a. Phred scores across each sequence for each sample. The base sequence for each sample is of good quality. b. Per sequence GC content. Eight samples from the first round of data collection (October 2019) had read sequences with a higher GC content, these samples were flagged with warnings (orange). The other samples collected in August 2020 were not flagged with warnings (green). c. Sequence duplication levels of a representative sample (12a) with a GC content of a normal distribution. This sample had 49.9% of reads, that are not duplicates. c. Sequence duplication levels of a representative sample (4a) with a higher GC content. This sample had 23.06% of reads that are not duplicates.

5.2.2 Evaluating Read Alignment and Cell Barcode Calling with Cell Ranger

Reads from the FASTQ files for each sample were aligned to Ensembl's reference genome, 'Amel_Hav3.1' using Cell Ranger. Across all samples, 68.2 - 82.5 % of reads were confidently mapped to the genome, and 49.7 - 57.0 % of reads were confidently mapped to the transcriptome. Cell Ranger detected 11538 - 11919 genes across the different samples. As there are 12,398 annotated genes in the *Apis mellifera* genome (GenBank GCA_003254395.2), a large portion of these annotated genes were captured.

As explained in section 4.2.2, Cell Ranger identifies barcodes likely associated with cells and removes barcodes linked with empty droplets by evaluating and comparing the RNA content associated with each barcode in the sample. For most of the samples, except sample 10a (black line indicated by the arrow, Figure 5.2) there is a clear drop in UMI counts around $\sim$ 1000 UMI (Figure 5.2). The estimated region splitting cell barcodes and empty droplets for sample 10a was also around 1000 UMIs. However, there was no distinct "knee"/"cliff" before this cut-off. This could suggest a high level of background noise associated with these barcodes. Of the remaining cells, the average of reads per cell was between 47294 and 98861, except for sample 10a, which had 26128 average reads per cell. Cell ranger estimated 3403 - 9125 cells per sample, except sample 10a, which had an estimated cell number of 22508. The high cell number and low average read counts per cell in sample 10a suggest a high noise level in this dataset; it is possible that Cell Ranger could not accurately identify cells from empty droplets.

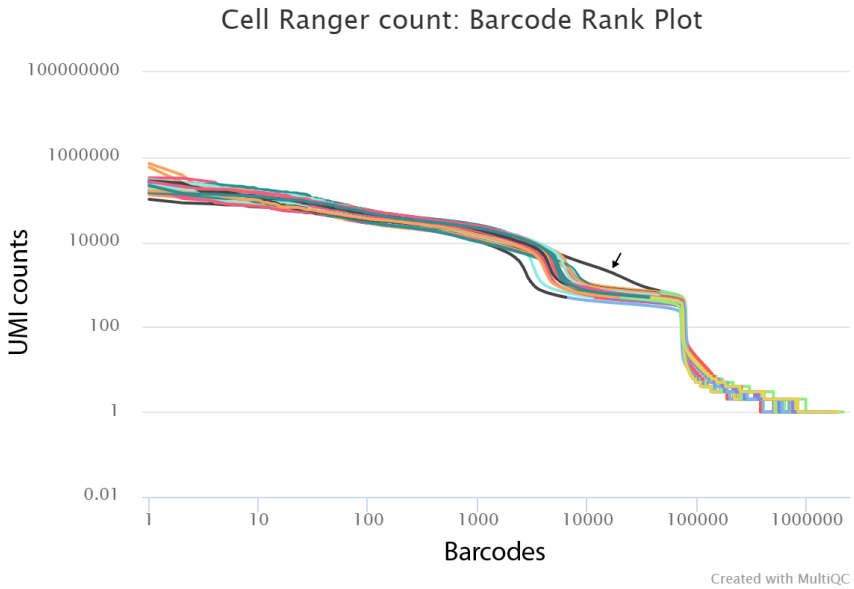


Figure 5.2: Evaluating read alignment and cell barcode calling with Cell Ranger. Barcode by UMI counts. For all samples, the threshold for barcodes associated with cells and those associated with empty droplets was about 1000 UMI. The arrow indicates the sample which did not have the characteristic "knee" or "cliff", sample 10a.

5.2.3 Assessing Cell Quality Per Sample

Feature-barcode matrices output from Cell Ranger for each sample were converted into Seurat objects for downstream analysis. Cells from each sample were evaluated for UMI count, feature count and the percentage of UMIs that were mitochondrial genes (Figure 5.3a - 5.3c). Samples 1a - 4b were collected in the first round of data collection (October 2019). The range of features per cell was consistent between these samples. The samples 5a - 12b collected in the second round of data collection (August 2020) had overall higher feature counts per cell and sample variation compared to samples 1a - 4b. However, sample 10a has the lowest median feature count and low variation between cells within this sample. This supports what was seen with Cell Ranger's output metrics that this sample contains mostly cells with low feature counts (Figure 5.3a). A similar observation was drawn from the number of UMIs per cell (Figure 5.3b, bottom figure); sample 10a appears to consist of cells with low overall features and UMI counts. A few cells had a high UMI count (Figure 5.3b). A cell from sample 10b has a UMI value of 705320, almost two-fold greater than the next highest value in this sample (Figure 5.3b, top figure). Sample 6a also had a cell with a high UMI count of 582597. These cells with high UMI counts had a low percentage of mitochondrial genes ($<$0.01% of UMIs for both cells were from mitochondrial genes). They were also not identified as doublets in downstream analysis and were not filtered out at this step. The percentage of UMIs derived from mitochondrial genes was evaluated, as a high percentage of these genes could indicate poor cell quality. As the percentage of UMIs derived from mitochondrial genes varied between each sample (Figure 5.3c), a hard cut-off value did not seem appropriate. I did not filter cells based on mitochondrial genes at this stage due to time restrictions. I aim to revisit this analysis and remove cells that have a percentage of UMIs derived from mitochondrial genes greater than three median absolute deviations (MADs).

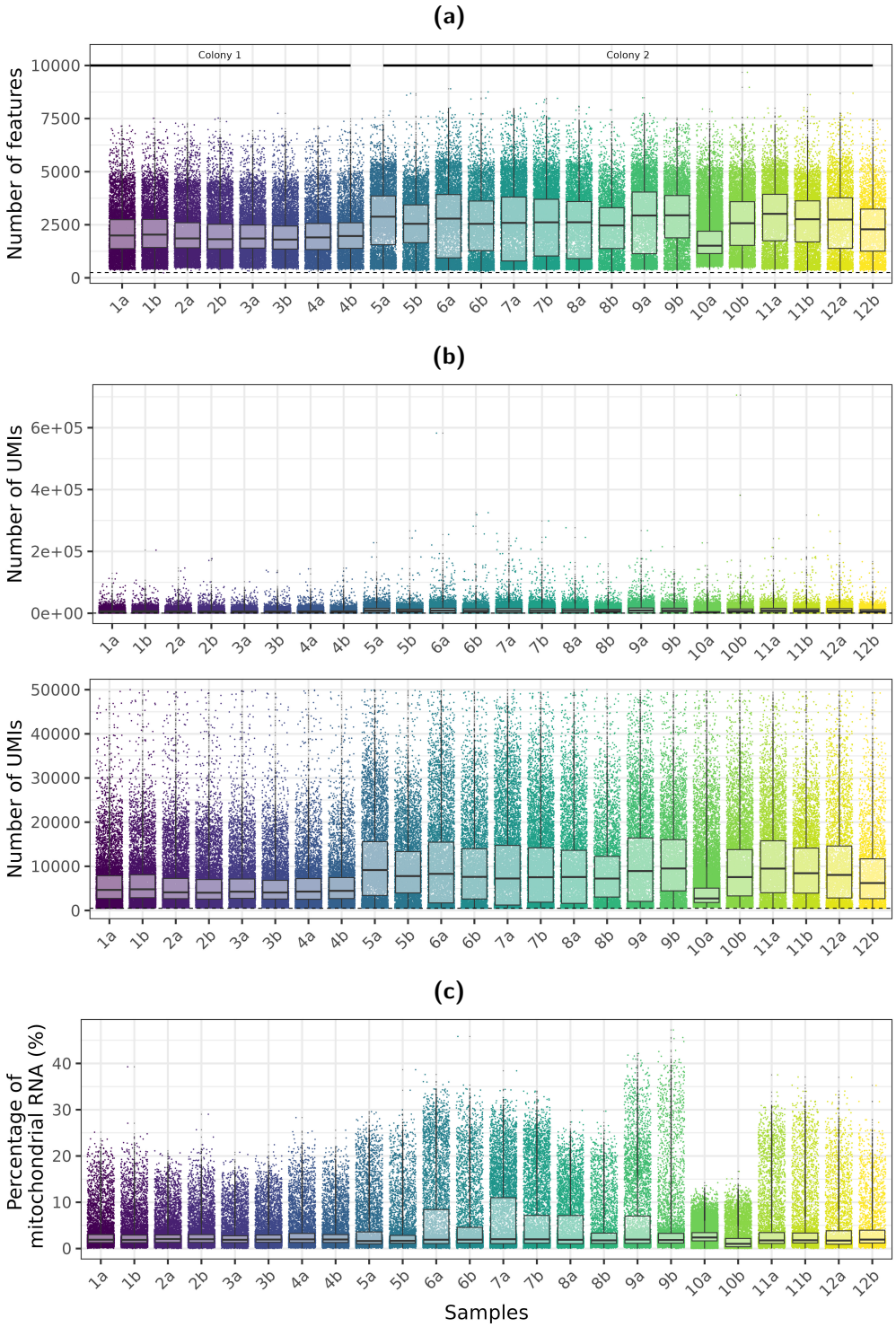
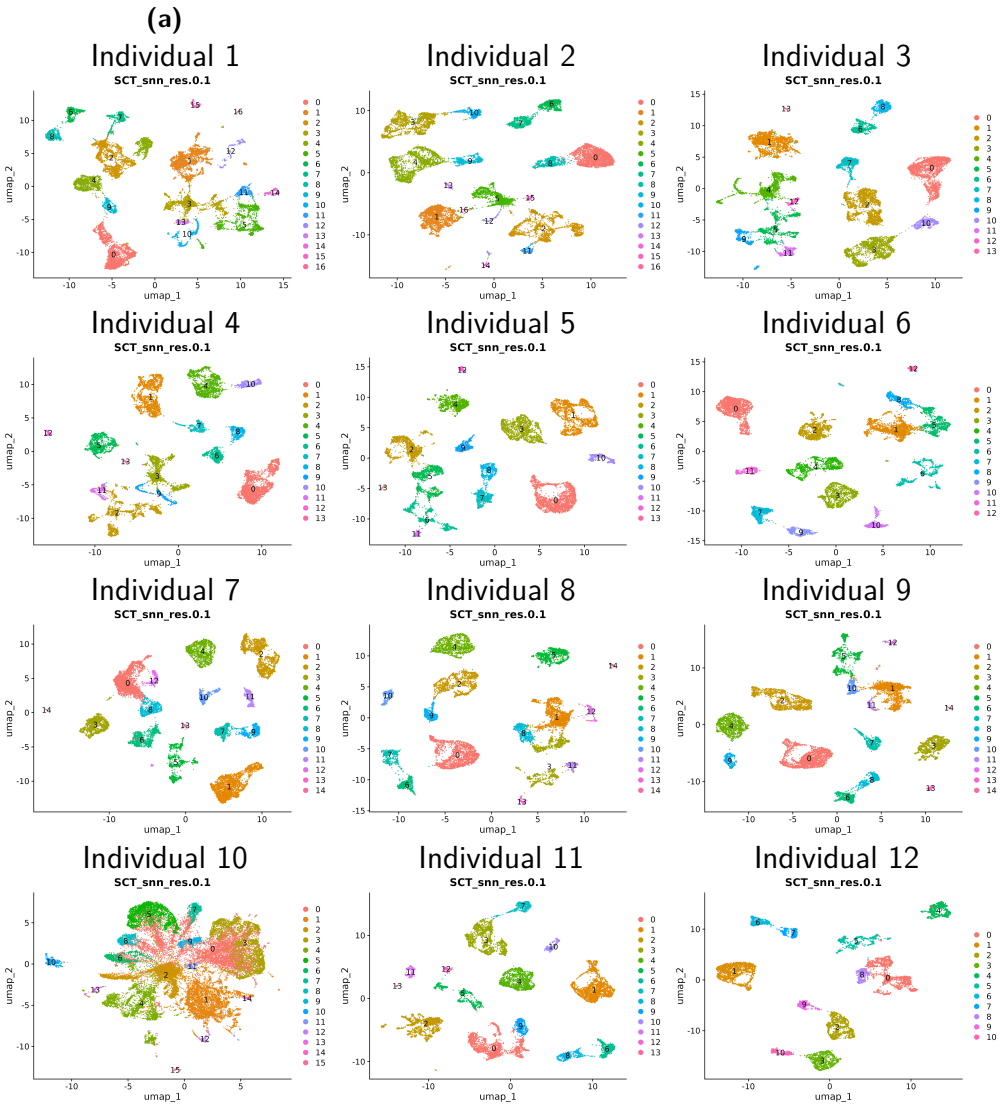


Figure 5.3: Evaluating UMI, feature and mitochondrial gene counts per sample. a. Feature counts. Samples from Colony 1 were collected in October 2019 and samples from Colony 2 were collected in August 2020. b. Number of UMIs. Top: Figure y-axis has been stretched due to two cells in samples 6a and 10b with high UMI counts. Bottom: Y-axis cropped at 50,000 UMIs. Sample 10a has the lowest median UMI score compared to the other samples. c. Percentage of UMIs that were from mitochondrial RNA for each cell per sample.

After filtering, the two samples per individual were merged. Dimensionality reduction and clustering, using a resolution of 0.1, identified 11-17 clusters per individual (Figures 5.4a). Reviewing clustering for each individual dataset can be useful for identifying any datasets with a comparatively high level of noise. For most individual datasets, the cells group into distinguishable clusters, however, the cells from individual 10 do not cluster into distinct groups suggesting there was additional noise in this dataset (Figures 5.4a). For most individuals, samples A and B merged together well, however, this was not the case for individual 10 (Figure 5.4b). Clustering 10a and 10b individually shows poor cluster quality in sample 10a compared to sample 10b (Figures 5.5a - 5.5b). Sample 10a was therefore removed from the dataset and downstream analysis.



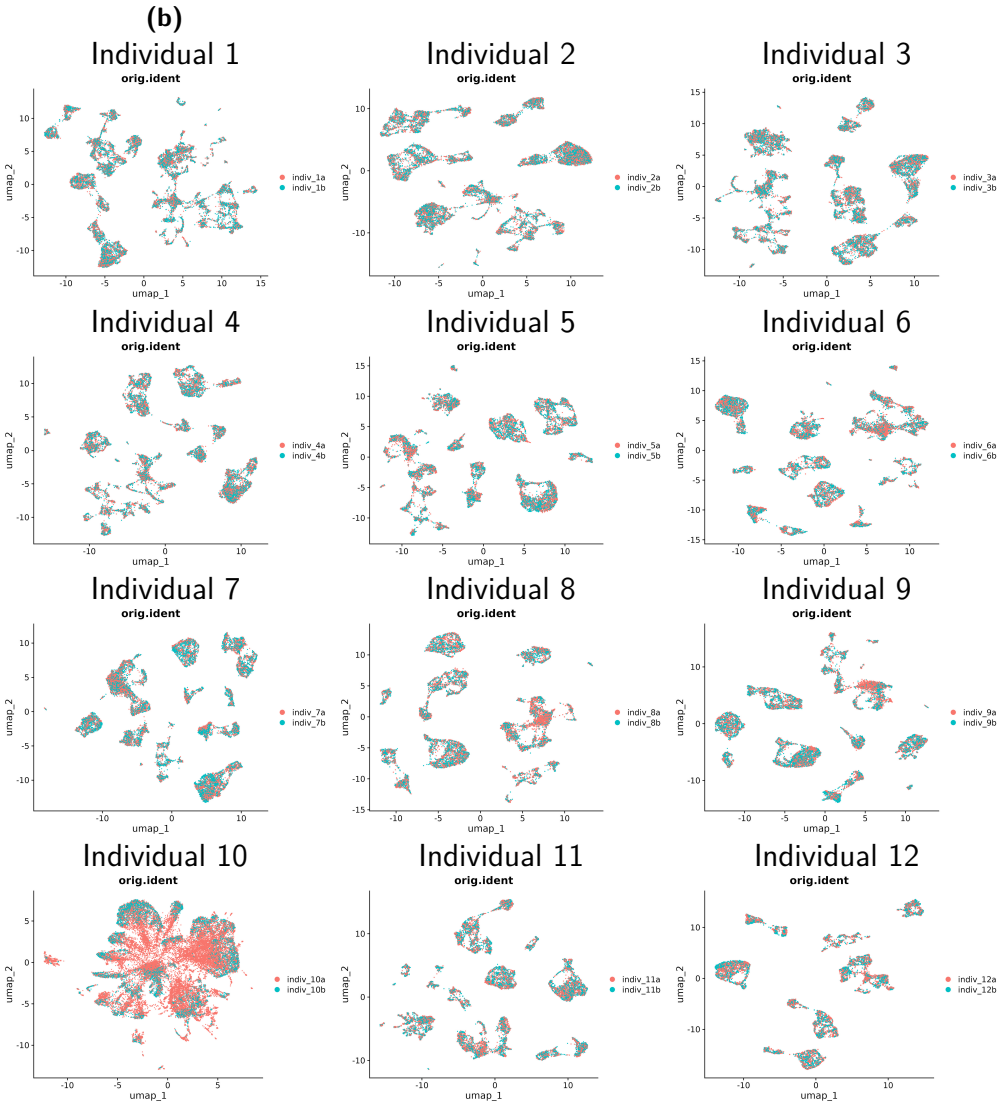


Figure 5.4: Per sample clustering and batch effect evaluation.
a. Clustering after merging A and B samples for each individual gave 11-17 clusters. b. Evaluating the integration of A and B samples after merging. There were no batch effects between most samples except samples from individual 10.

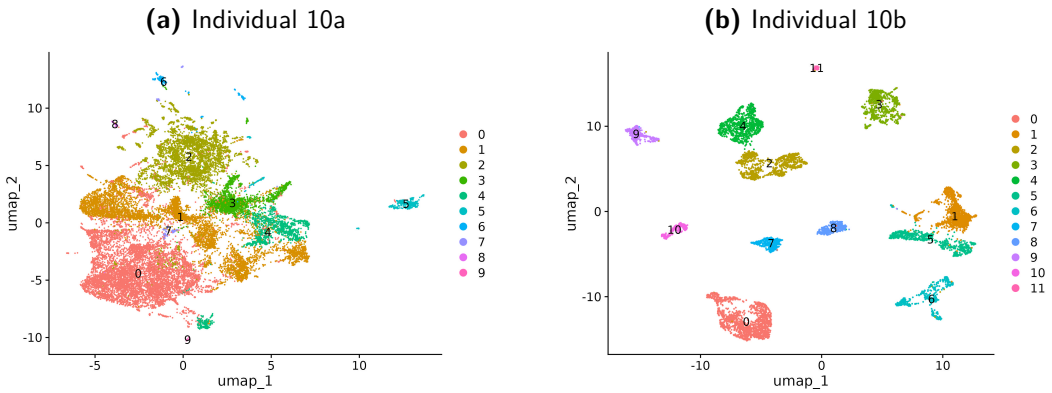


Figure 5.5: Dimension reduction plots showing clusters for samples 10a and 10b. a. Clustering of sample 10a suggests cells contain much background noise. b. Clustering of sample 10b shows 12 clusters.

5.2.4 Identification and Removal of Doublets

"Doublets" occur in scRNA-seq datasets when two (or more) cells are captured in one droplet and sequenced under one cell ID barcode. To evaluate the presence of doublets in the scRNA-seq *Apis* dorsal protocerebra dataset, the doublet detection package "DoubletFinder" was used. DoubletFinder identifies doublets based on a cell's proportion of nearest neighbours made from artificial doublets (McGinnis et al. 2019). This analysis found between 1.87% and 7.83% doublets for each sample (Table 5.1), and doublets were often positioned between the clusters (doublets shown in blue, Figure 5.6). The cells identified as doublets were removed from the dataset.

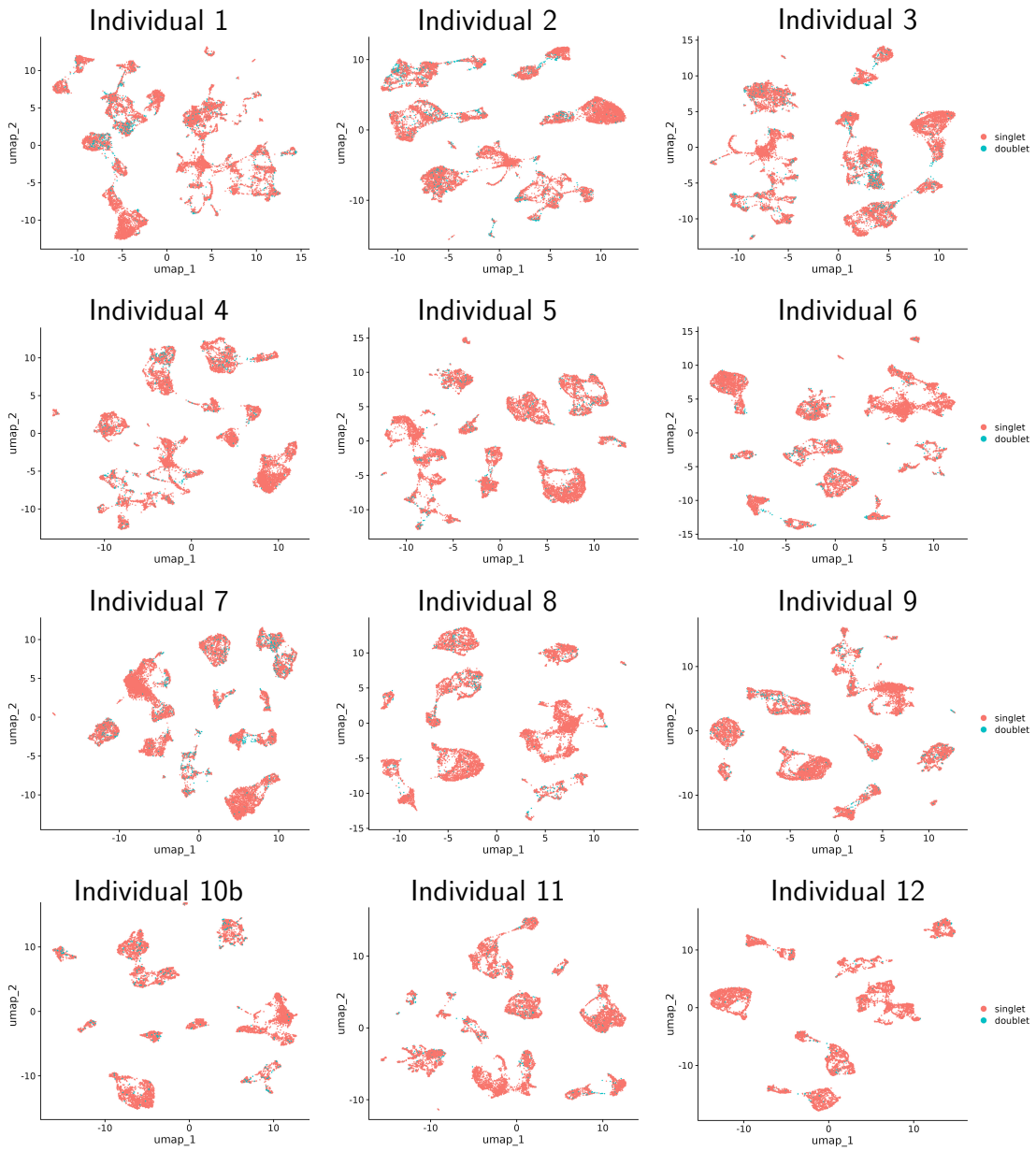


Figure 5.6: Identification of doublets. Cells identified as doublets (blue) were often positioned between clusters. Clusters were comprised of mostly singlets (pink).

5.2.5 Evaluation of Ambient RNA using DecontX

While creating the cell suspension for scRNA-seq, it is common for cells to rupture, releasing RNA into solution and creating 'ambient RNA'. During the capture of single cells in droplets before sequencing, there is a high chance that the droplet will also capture some of this ambient RNA. DecontX is a Bioconductor tool that utilises the theory of Bayesian statistics to approximate relative ambient RNA contamination within each cell. Reviewing contamination scores across the different samples shows regions of cells with high ambient RNA (Figure 5.7). For many samples, cells with the highest contamination scores are positioned within one cluster (see individuals 1, 2, 4 and 5 Figure 5.7), and between clusters. This suggests that the cells with the highest contamination are clustered together based on a high level of ambient RNA. Before filtering, the positions of cells with different severity of contamination scores were evaluated. Figure 5.8a shows the spread of contamination scores and the MAD thresholds for each sample. Plotting cells with contamination scores above four and five MADs does not appear to be stringent enough to capture the main areas of contamination (Figure 5.8b, cells with contamination scores above the MAD threshold coloured blue). Therefore, I filtered out cells with a score greater than three MADs for each sample, removing between 163 and 1547 cells per sample (see Table 5.1).

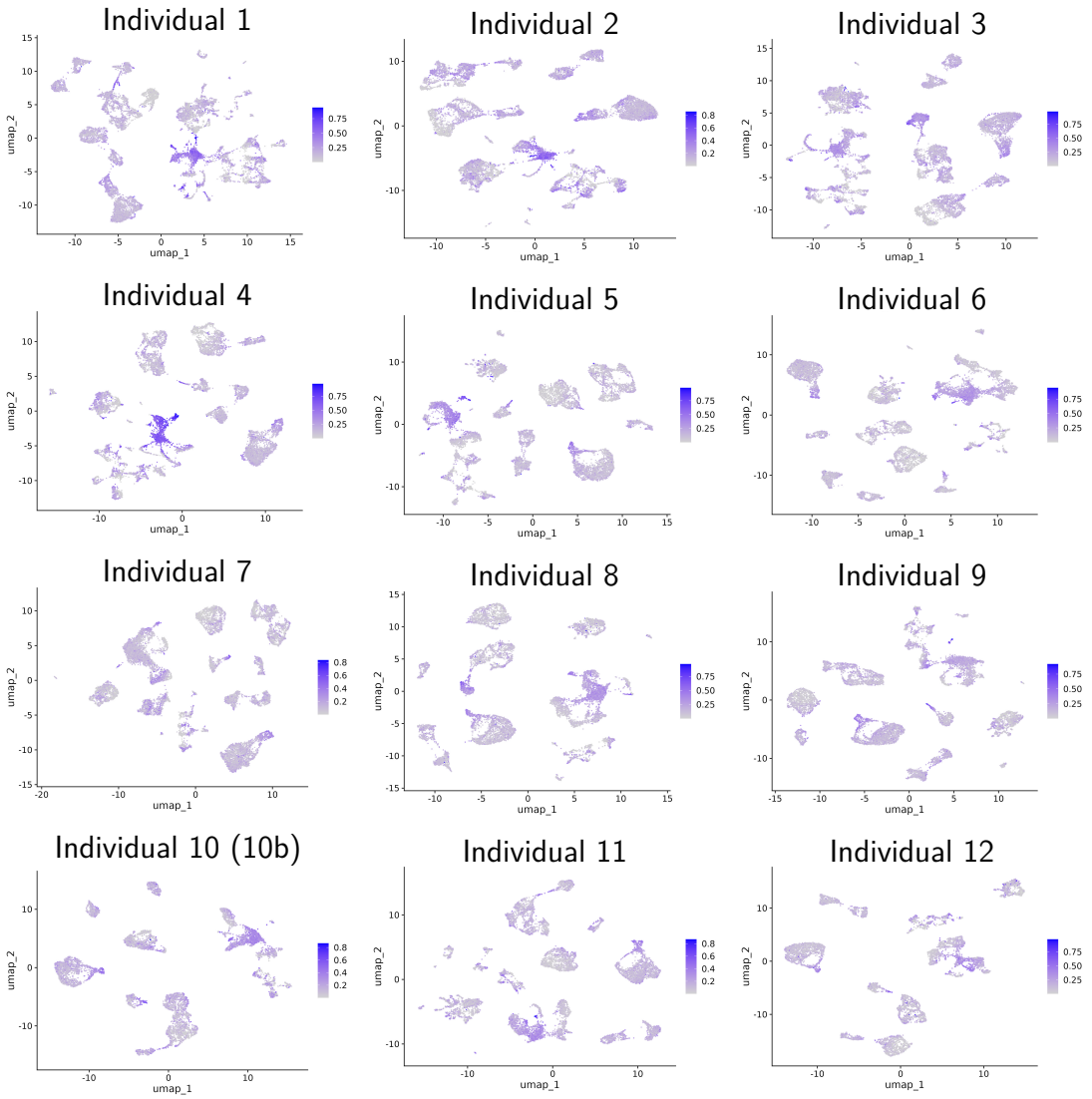


Figure 5.7: Evaluation of ambient RNA contamination levels per sample. Cells with a high contamination score appear to cluster together, as seen in the figures for individuals 1, 2 and 4, and between clusters.

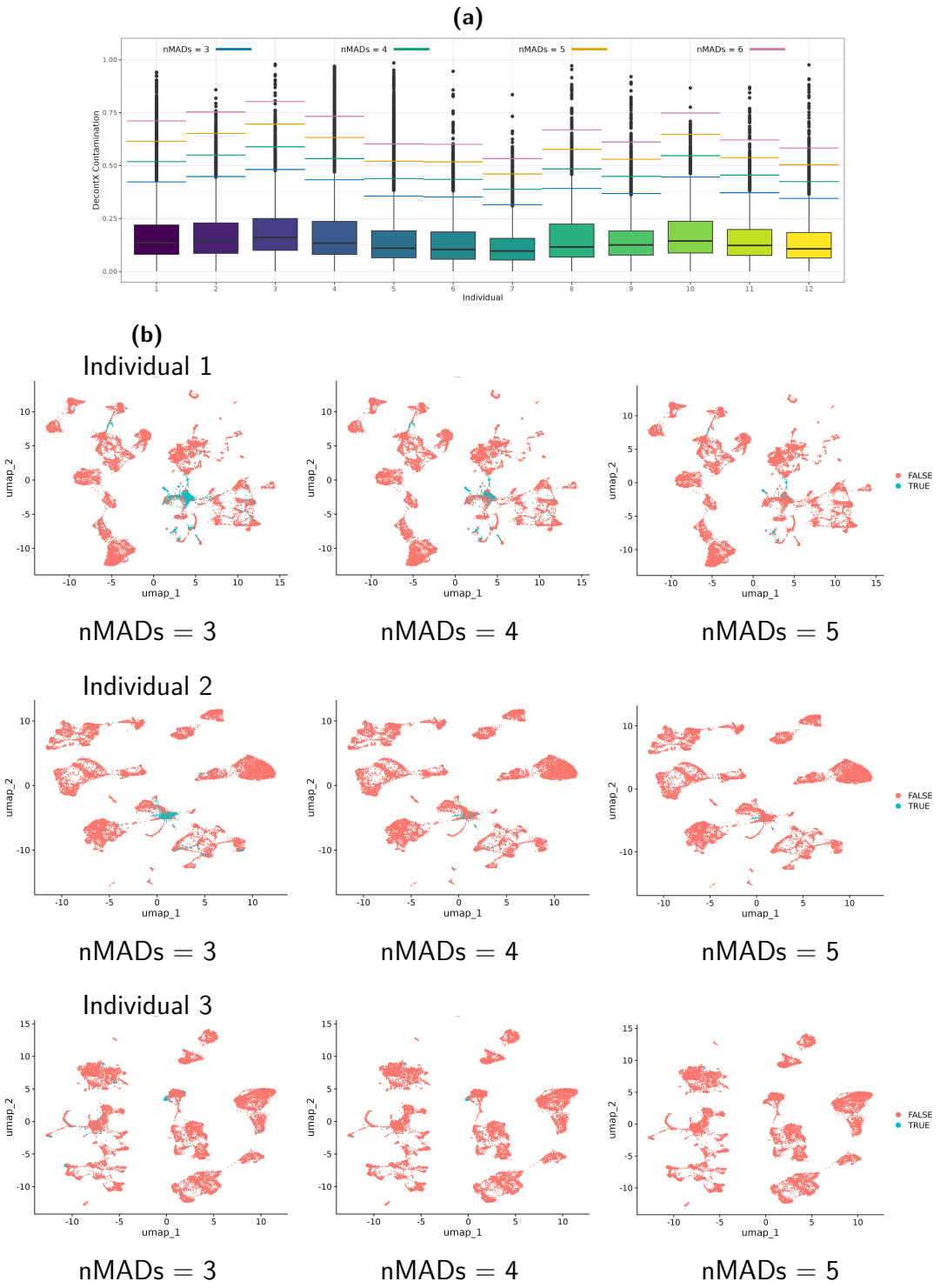


Figure 5.8: Evaluation of different thresholds for filtering cells based on median absolute deviation (MAD).
a. Box plots showing the distribution of DecontX contamination scores per sample. Coloured lines show the threshold values for different MADs values. Three MADs in blue, four MADs in green, five MADs in orange, and six MADs in pink. b. Cells annotated based on MAD thresholds. Cells with contamination scores greater than three, four or five MADs were annotated in blue (TRUE), and those below the threshold were in pink (FALSE).

Table 5.1

Individual	Total cell number	Number of Doublets	Percentage of Doublets	Number of cells removed after DecontX analysis	Percentage of cells removed after DecontX analysis
1	16030	1099	6.86%	1284	8.60%
2	17872	1400	7.83%	799	4.85%
3	16757	1210	7.22%	308	1.99%
4	13663	743	5.44%	1547	11.97%
5	12787	641	5.01%	1126	9.27%
6	12129	570	4.70%	480	4.15%
7	16806	1210	7.20%	614	3.94%
8	11569	501	4.33%	423	3.82%
9	13010	664	5.10%	339	2.75%
10 (b)	8459	189	2.23%	163	1.97%
11	12450	587	4.71%	397	3.35%
12	7595	142	1.87%	438	5.88%

The percentage and number of cells removed after identifying doublets and evaluating DecontX scores.

5.2.6 Integration using Harmony Removes Batch Effects

Before analysis, the individual samples must be merged or integrated. Merging datasets without integration can reveal batch effects that might exist between samples in the dataset. Merging the data and finding clusters at a resolution of 0.1 identifies 20 clusters (Figure 5.9a). However, Figure 5.9b shows a batch separation between the datasets from different colonies (colony 1, collected in October 2019 and colony 2, collected in August 2020). Harmony, the integration tool, was employed to integrate the data and account for batch effects. Initially, Harmony finds clusters using a "soft k-means" method on the cell embeddings (PCAs). A "penalty term" ensures the representation of different cells within batch types in each cluster. The algorithm iteratively adjusts clustering by using correction factors based on sample-specific centroids for each cluster (Korsunsky et al. 2019). Harmony integrated the 12 datasets, creating 19 clusters at a cluster resolution of 0.1 (Figure 5.9c) and corrected for batch effects between the two colonies (Figure 5.9d).

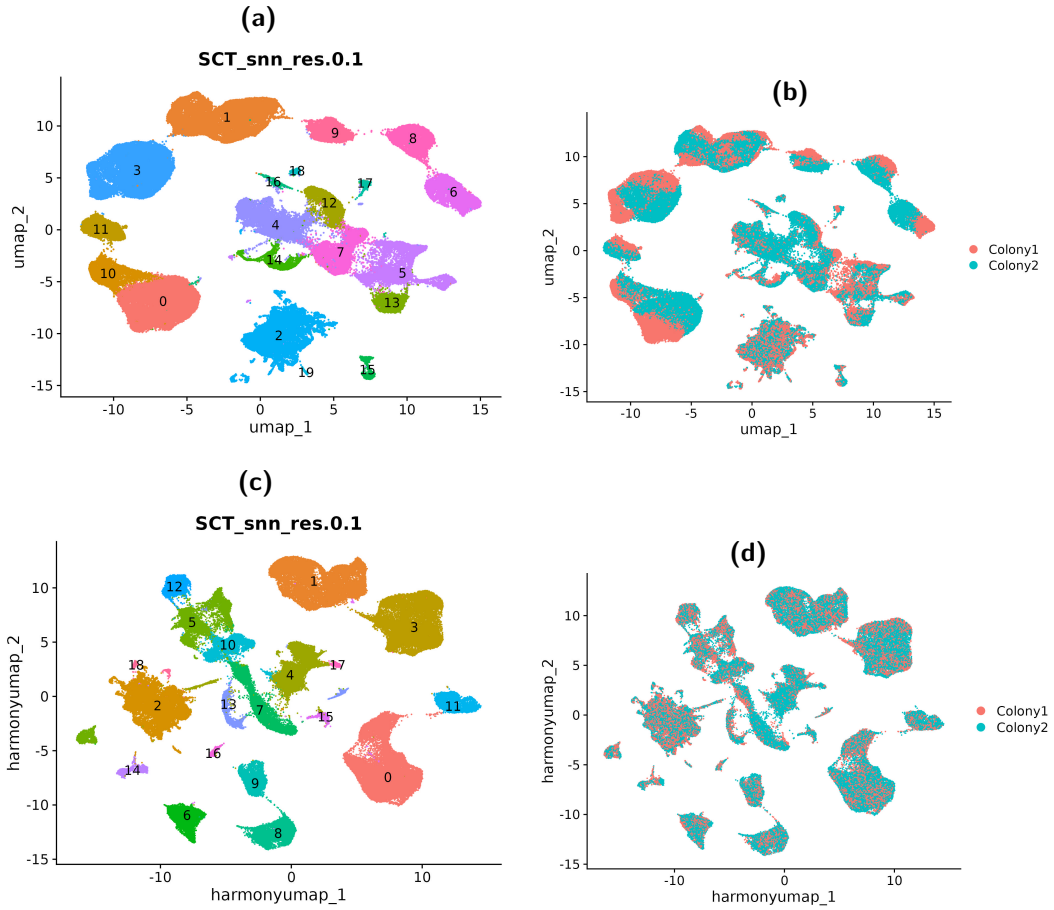


Figure 5.9: Review of methods used to combine filtered scRNA-seq data of the dorsal protocerebum from twelve *Apis mellifera* pollen-forager honey bees.
a-b. Merging the scRNA-seq data from the twelve *Apis* dorsal protocerebra. a. Merging the data produced 20 clusters. b. The merged object shows some batch effect as the cells from each cluster are split by colony. c-d. Integrating the twelve dorsal protocerebra datasets using the integration tool Harmony. c. Integrating the data using harmony created 19 clusters. d. The Harmony integrated object shows that the cells from the two colonies were well integrated.

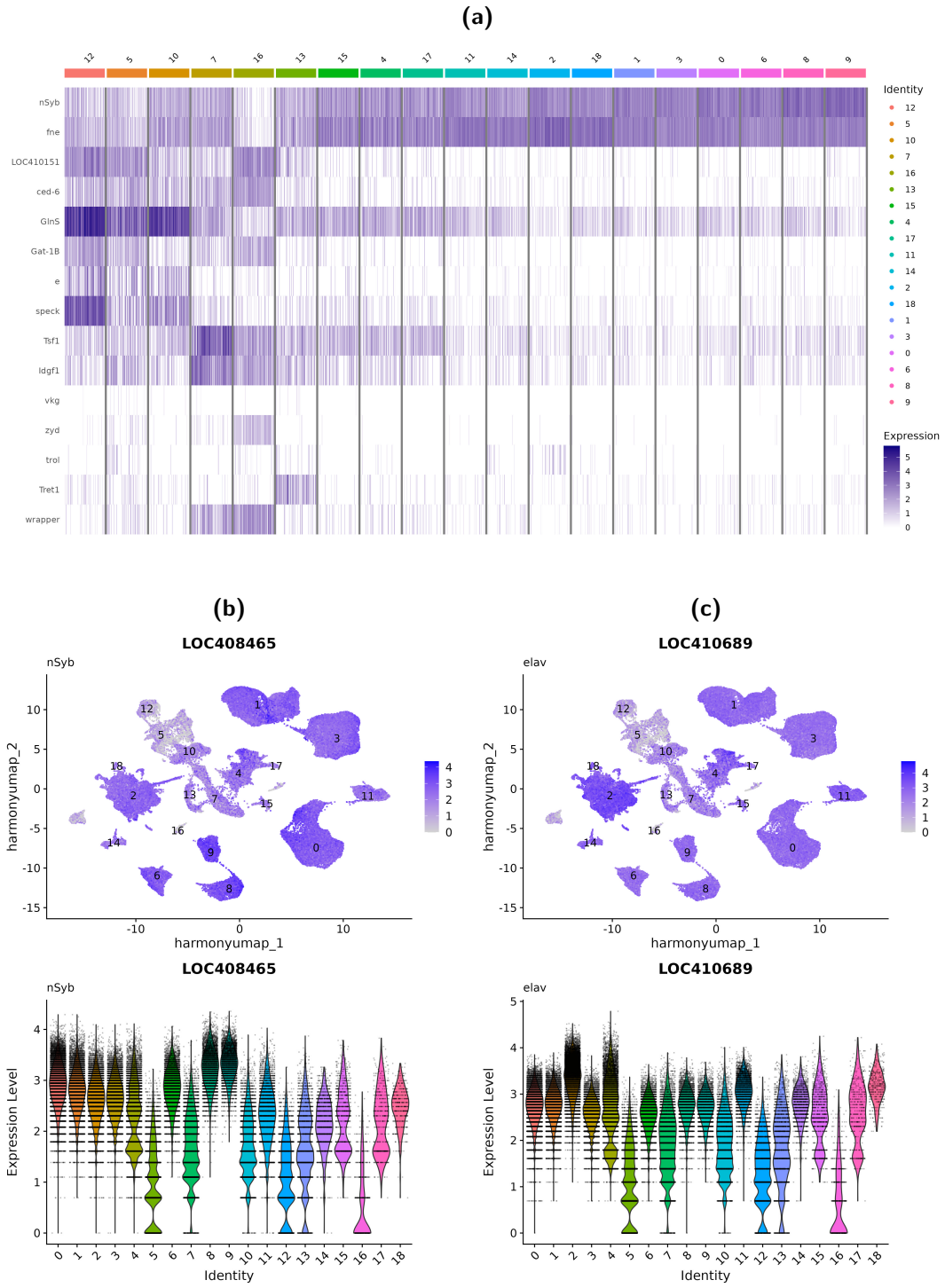
5.2.7 Annotation

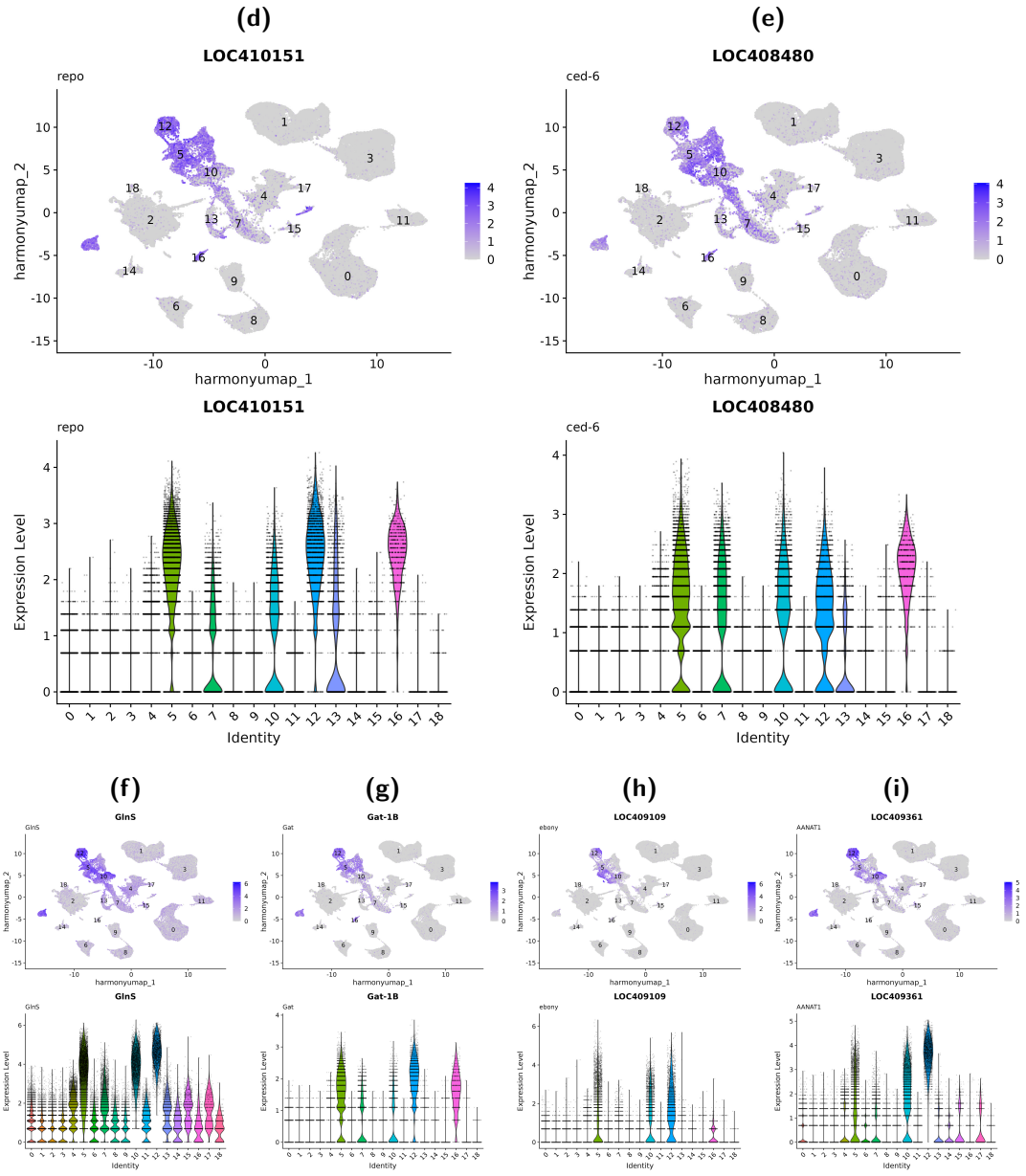
5.2.8 Neurones and Glia

To better identify Kenyon Cell and monoaminergic cells types for downstream analysis, I first separated neuronal from non-neuronal cell types using both *Apis* and *Drosophila* neuronal and glial marker genes. Expression of the neuronal cell-specific markers *elav* (Zhang et al. 2022) and *nSyb* (Croset et al. 2018) was high in all clusters, except for clusters 5, 7, 10, 12 and 16 (Figures 5.10b and 5.10c), these clusters had a high expression of the glial cell marker genes *rx2* and *ced-6*. Cluster 13 had some expression of both neuronal and glial cell marker genes. I, therefore, plotted the expression of marker genes for glial cell subtypes. Cluster 13 had a high expression of *Tret1* (Figure 5.10p), a marker for surface glia (Zhang et al. 2022; Park et al. 2022), therefore, cluster 13 was classified as glia. Overall, six clusters were annotated as glial and thirteen were classed as neuronal (Figure 5.10a), with 26547 cells being glial and 115705 labelled as neuronal cells.

Glial cells can be divided further by the expression of marker genes for glial subtypes. The marker genes for astrocytes were expressed mostly in clusters 5, 10, 12 (*GlnS, Gat-1B, ebony* and *ANNAT1*, Figures 5.10f to 5.10i), with some expression of *Gat-1B* in cluster 16 (Figure 5.10g). Marker genes for ensheathing glia (*Tsf1, Igdf4, trol* and *zyd*) were expressed in clusters 5, 7, 13 and 16, with *trol* being lowly expressed. Marker genes for cortex glia (*zyd* and *wrapper*) were expressed in clusters 7 and 16 (Figures 5.10m and 5.10n). As mentioned above, the marker gene for surface glia *Tret-1* was expressed in cluster 13.

Reclustering and reintegrating the glial cells gave eleven clusters (Appendix Figure A.7).


(a)
12 5 10 7 16 13 15 4 17 11 14 2 18 1 3 0 6 8 9
nSyb
fne
LOC410151
ced-6
GlnS
Gat-1B
e
speck
Tsf1
Idgf1
vkg
zyd
trol
Tret1
wrapper
Identity
12 5 10 7 16 13 15 4 17 11 14 2 18 1 3 0 6 8 9
Expression
5 4 3 2 1 0
(b)
LOC408465
nSyb
harmonyumap_2
harmonyumap_1
(c)
LOC410689
elav
harmonyumap_2
harmonyumap_1
LOC408465
nSyb
Expression Level
Identity
LOC410689
elav
Expression Level
Identity


(d)
LOC410151
repo
harmonyumap_2
harmonyumap_1
LOC410151
repo
Expression Level
Identity
(e)
LOC408480
ced-6
harmonyumap_2
harmonyumap_1
LOC408480
ced-6
Expression Level
Identity
(f)
GlnS
(g)
Gat-1B
(h)
LOC409109
(i)
LOC409361

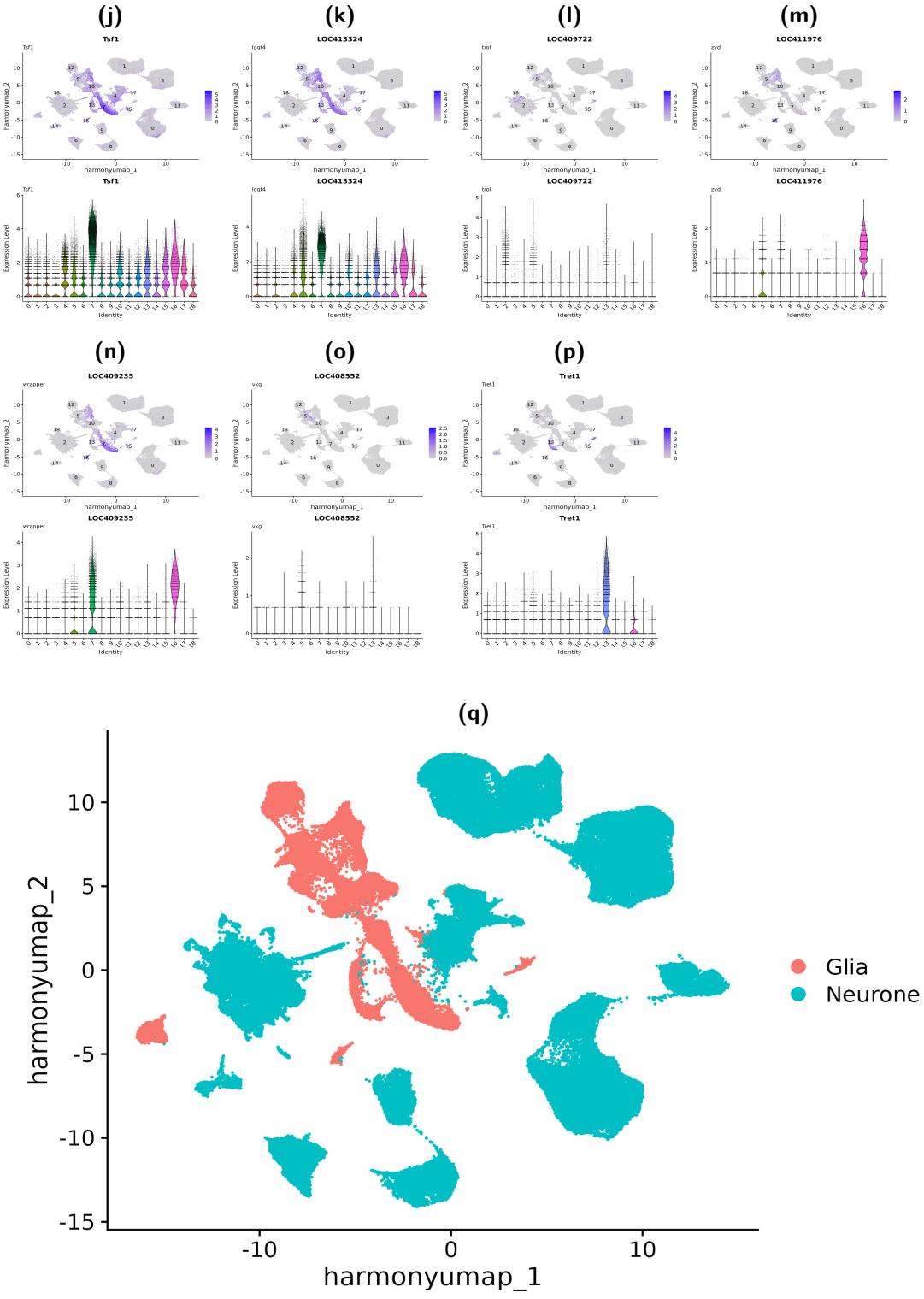
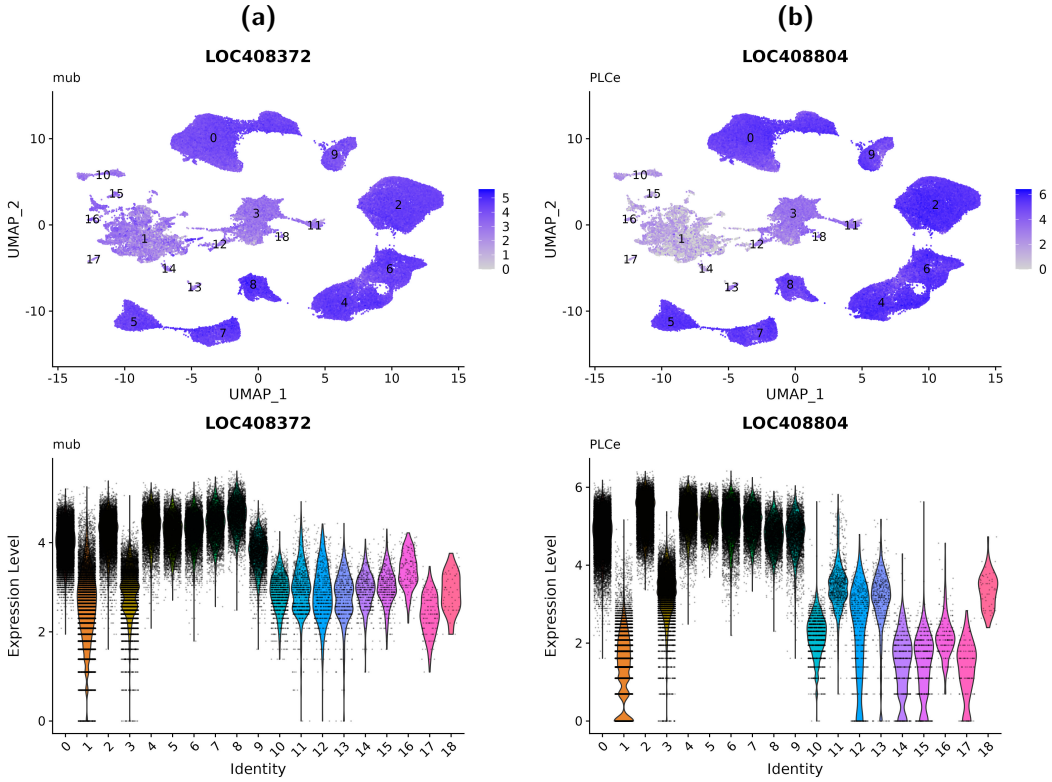


Figure 5.10: Identifying neuronal and glial cell clusters. a. Heatmap plotting neuronal and glial marker genes at a cluster resolution of 0.1. b - c. Feature and violin plots for neuronal markers *nSyb* and *elav*. d - p. Feature and violin plots for Glial markers *GlnS*, *Gat-1B*, *ebony*, *ANNAT1*, *Tsf1*, *Idgf4*, *vkg*, *zyd*, *trol*, *Tret1* and *wrapper*. q. Glial and neuronal labels on the full dataset.

5.2.9 Identifying Kenyon Cell Clusters

Reintegrating and clustering the neuronal cell types at a resolution of 0.1 gave nineteen clusters. Kenyon Cell markers *mub* and *PLCe* revealed eight Kenyon Cell clusters and thirteen non-Kenyon Cell neuronal cells. Isolating and reprocessing the Kenyon Cell clusters from the other neuronal clusters resulted in nine Kenyon Cell clusters (clustering resolution of 0.1) made of 83203 cells, and twenty-five non-Kenyon Cell neuronal clusters (resolution 0.1) made of 32502 cells.



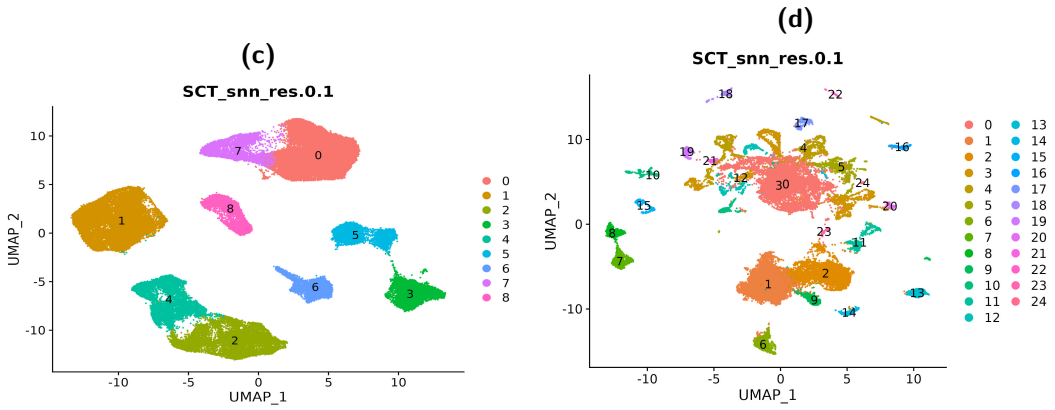


Figure 5.11: Identifying Kenyon Cell clusters from non-Kenyon Cell neuronal clusters. a - b. Feature and violin plots of the expression of Kenyon Cell marker genes *mub* and *PLCe*. Expression of *mub* and *PLCe* was expressed the highest in clusters 0, 2, 4, 5, 6, 7, 8 and 9. c. Reintegration of cells identified as Kenyon Cells identified nine clusters at a resolution of 0.1 d. Reintegration of non-Kenyon Cell neuronal cells revealed twenty-five clusters at a cluster resolution of 0.1.

5.2.10 Kenyon Cell Annotation

As mentioned in section 1.2.2, Zhang et al. (2022) identified Class - I lKCs, sKCs and FoxP-KCs clusters in *Apis* nurse, forager and queen scRNA-seq data using the marker genes *Mblk-1*, *E74* and *FoxP*, respectively. In this report, three clusters were labelled as Class - I lKCs, three clusters labelled Class-I sKCs and one cluster labelled as FoxP-KCs also using marker genes *Mblk-1*, *E74* and *FoxP*, respectively (Figures 5.12a - 5.12c). In addition, two clusters were identified as Class - I mKCs expressing *mKast*, a marker gene for mKCs (Suenami et al., 2018, Figure 5.12d). Cluster 8 also expressed mKast on the portion that was not *FoxP* positive. As the expression of *FoxP* in cluster 8 was greater than *mKast*, the annotation for cluster 8 remained FoxP-KCs. Cluster 7 expressed both *mKast* and *E74* and was therefore classified as both mKC and sKC. Cluster 6 did not have a high expression of any KC marker genes and was left unannotated (Figure 5.13).

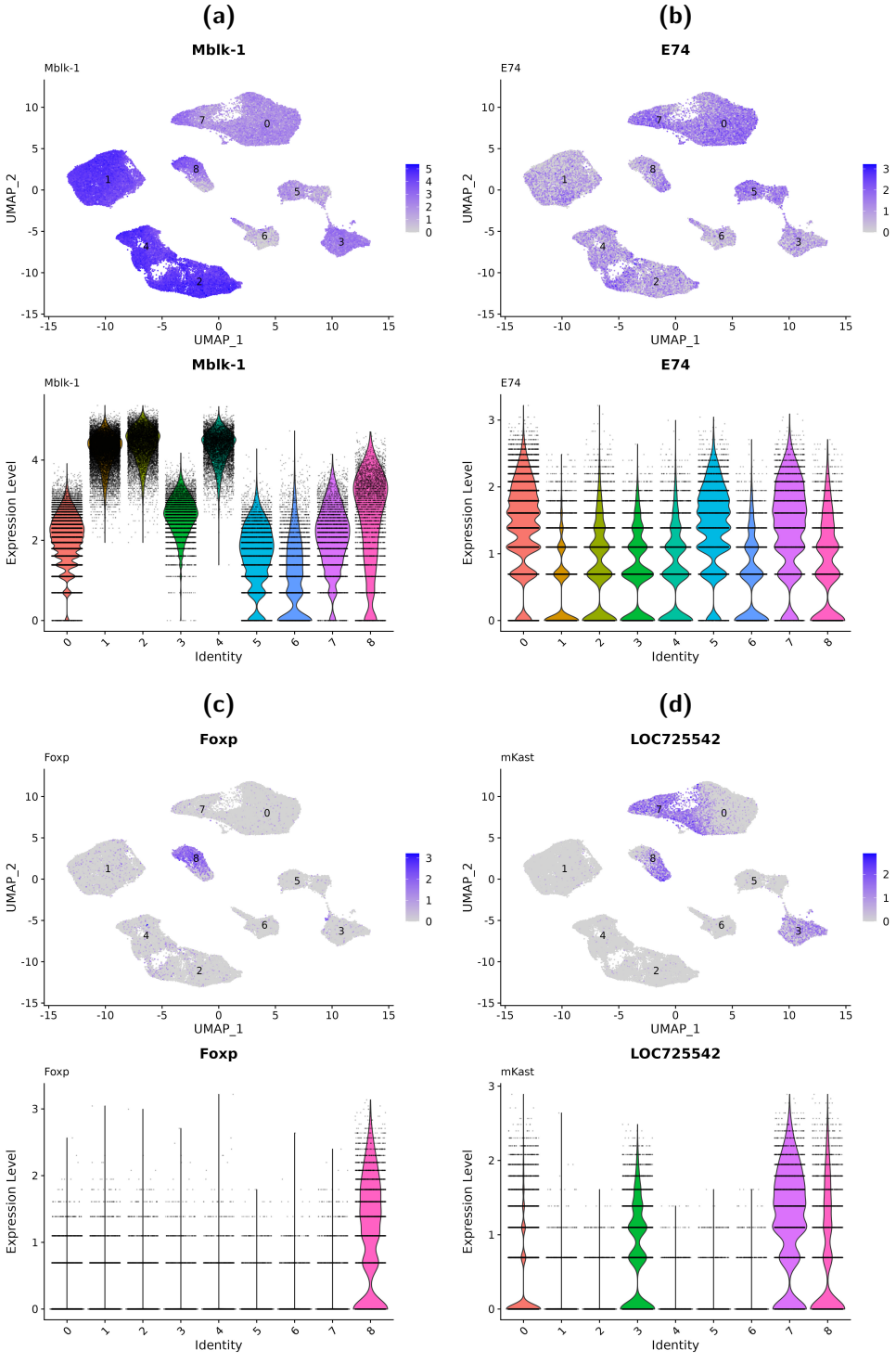


Figure 5.12: Identification of Kenyon Cell sub-clusters.
a - d. Feature and violin plots showing expression profiles of Kenyon Cell marker genes *Mblk-1*, *E74*, *FoxP* and *mKast*. a. The Class-I lKC marker *Mblk-1* was expressed in clusters 1, 2 and 4. b. The Class-I sKC marker genes *E74* has a broad expression in most clusters but was the most expressed in clusters 0, 7 and 5. c. *FoxP* the marker gene for FoxP-KCs was expressed in cluster 8. d. *mKast* the marker gene for Class-I mKCs annotated clusters 7, cluster 3 and some of cluster 8.

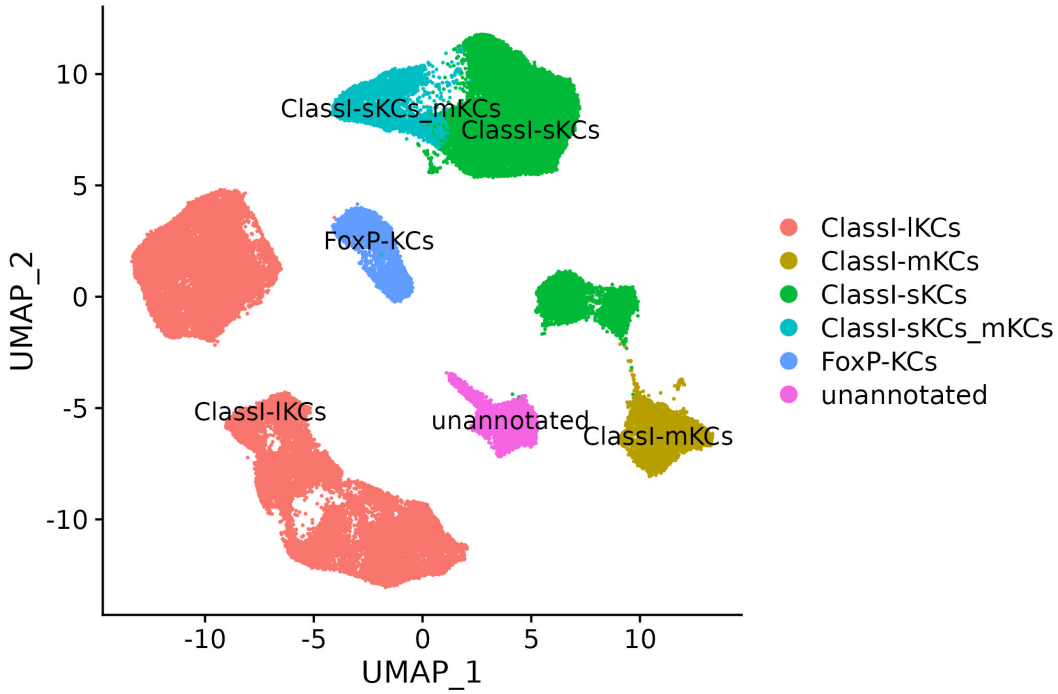


Figure 5.13: Kenyon Cell clusters annotated with Kenyon Cell subtypes from the literature. Three clusters were identified as Class-I lKCs by the expression of *Mblk-1*, one cluster was labelled Class-I mKCs based on the expression of *mKast*, two clusters were annotated Class-I sKCs based on the expression of *E74*, one cluster shared the expression of *mKast* and *E74* and was annotated Class-I sKC and mKC, and one cluster was annotated FoxP-KC based on the expression of *FoxP*. One cluster did not have a high expression of any marker genes and was left unannotated.

5.2.11 *Apis mellifera* Kenyon Cells have Greater Granularity than the Five Proposed Cell Subtypes Identified in the Literature

As evidenced in Figure 5.13, there is additional granularity to the *Apis* KCs compared to what has been previously described in the literature (Suenami et al. 2018). Therefore, I looked to find marker genes that are unique for each cluster, and marker genes that discriminate clusters within each published KC subtype. Figure 5.14 shows that the nine KC clusters differ by the expression of six to seventeen unique marker genes each. Figures 5.15b and 5.15c show that there are marker genes in common between clusters of Class-I type lKCs and clusters of Class-I type sKCs; however, Figures 5.15c and 5.15d show marker genes that discriminate clusters within the lKC, and the sKC subtypes, respectively (see below).

Clusters Belonging to Class-I lKCs

The three clusters in the lKC group were identified by the expression of *Mblk-1* (Figure 5.12a). These clusters also have a high expression of other marker genes such as the *nicotinic acetylcholine receptor alpha1 subunit, nAChRa1* and *Sulphated1* (Figures 5.15b and 5.15c). Figure 5.15c shows differentially expressed genes between the clusters within the lKC category. For example, cluster 4 has a greater expression of the serotonin receptor gene *5-ht7*, cluster 2 has a greater expression of *mamo* (a gene exclusively expressed in $\alpha'\beta'$ *Drosophila* KCs, Liu et al., 2019) and cluster 1 has a greater expression of the metabotropic glutamate receptor B gene, *Glurb*, compared to the other lKC clusters. This suggests that the clusters within the lKC category are functionally distinct.

Clusters Belonging to Class-I sKCs

Clusters 0 and 5 that were categorised as sKCs based on their expression of *E74* (Figure 5.12b) also have a high expression of *Frq2* and *chinmo* (a transcription factor regulating KC fate in *Drosophila*, reviewed by Lin 2023, gene ID LOC551086) compared to the other clusters (Figure 5.15d). Clusters 0 and 5 differ by the expression of several genes. For example, there was a greater expression of the Allatostatin-C receptor gene

AstC-R, *Fas1* and *Ptp36E* in cluster 0 compared to the other sKC clusters and a greater expression of *5-ht7*, short neuropeptide F (*sNPF*, gene ID LOC100576163) and *Imp* in cluster 5 compared to the other sKC clusters (Figure 5.15d).

Clusters Belonging to Class-I mKCs and sKC/mKCs

Two clusters, clusters 3 and 7, had a high expression of *mKast*, which annotates mKCs. The cell bodies of mKCs are located between sKC and lKC cell bodies (Kaneko et al. 2013). Cluster 7 also expressed *E74*, a marker for sKCs. Cluster 7 could represent a group of cells with cell bodies positioned at the border between the sKC and mKC cell bodies. However, cluster 7 differs from the other sKC and mKC clusters by several genes (Figures 5.14, 5.15b & 5.15d); cluster 7 is likely a discrete subcluster that contains similarities with clusters belonging to sKCs and mKCs.

The Cluster Belonging to Fox-P Expressing KCs

Cluster 8 was labelled FoxP-KCs as it expressed the *FoxP* marker gene (Figure 5.12c and 5.14). This cluster also had a high expression of the diuretic hormone receptor, *Dh44-R2*. As mentioned in Chapter 4, Dh44-R2 expression in the *Drosophila* malpighian tube is required for salt tolerance (Hector et al. 2009), but *Dh44-R2* is also necessary for nutritious sugar selection in starved flies (Dus et al. 2015). This cluster might be involved in nutrient sensing and feeding behaviours.

Cluster 6 Might be Class-II KCs

Cluster 6 was unannotated. These cells may be Class-II KCs, but there are currently no good marker genes to annotate Class-II KCs discretely. Cluster 6 had a high expression of marker genes such as *Mipp1*, *PIG-K* and *Ekar* (Figure 5.14). However, many of the genes highly expressed in cluster 6 compared to the other clusters were uncharacterised.

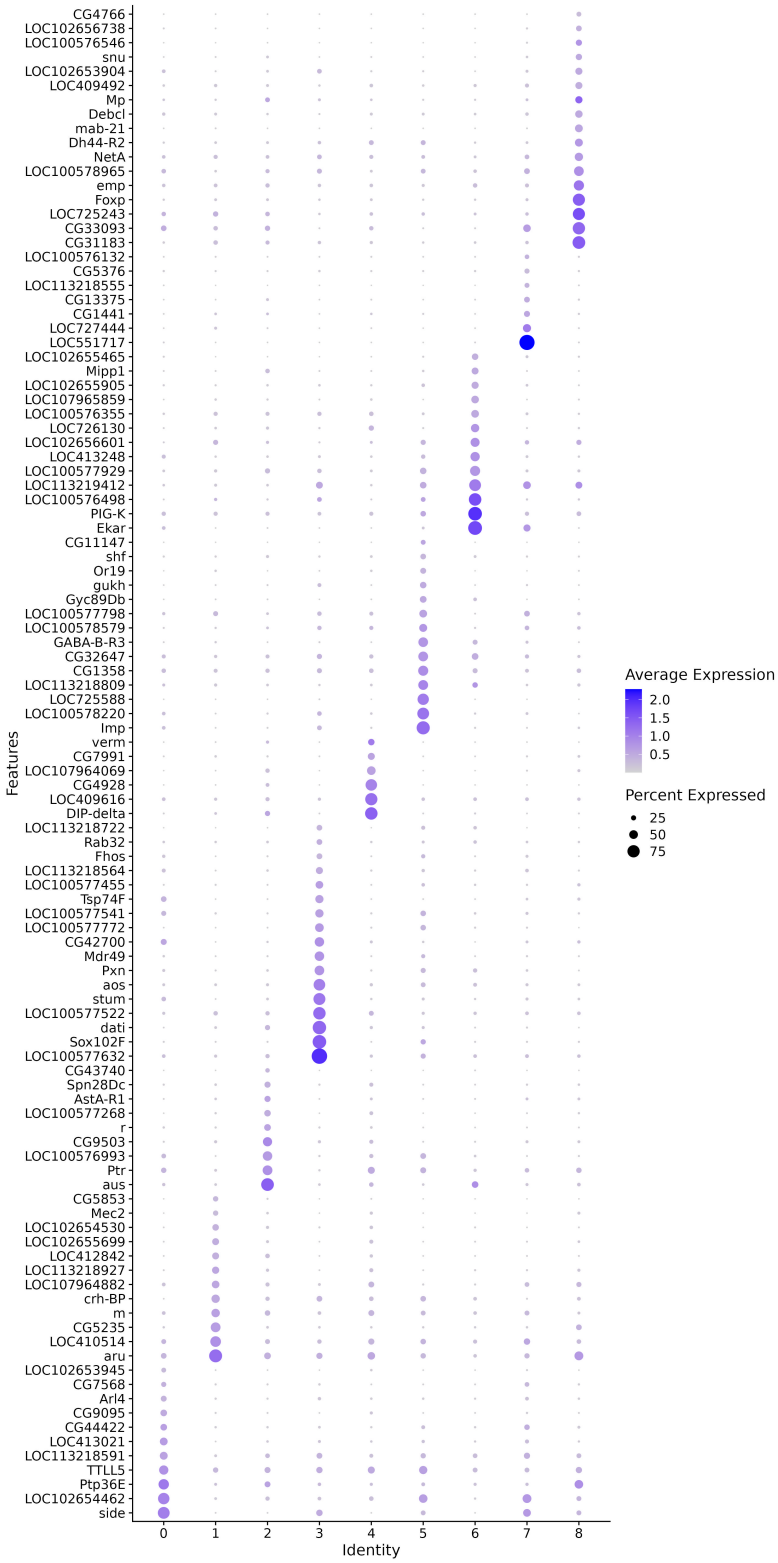
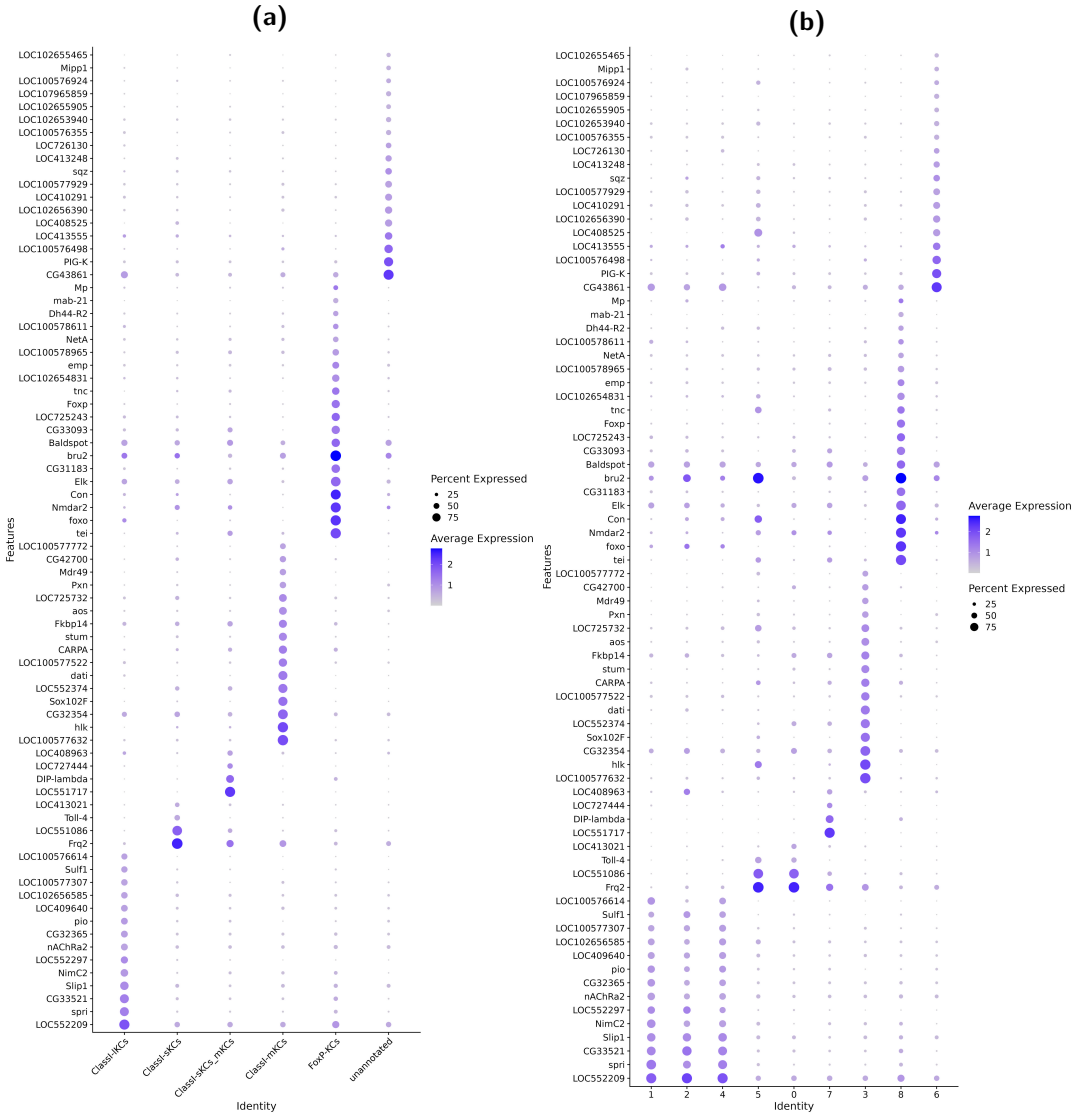


Figure 5.14: Top unique marker genes for each Kenyon Cell cluster. Analysis of differentially expressed genes identified between six and seventeen unique marker genes per cluster.


(a)
(b)
Features
Identity
Percent Expressed
25
50
75
Average Expression
2
1
Classl-lKCs
Classl-sKCs
Classl-sKCs_mKCs
Classl-mKCs
FoxP-KCs
unannotated
1
2
4
5
0
7
3
8
6

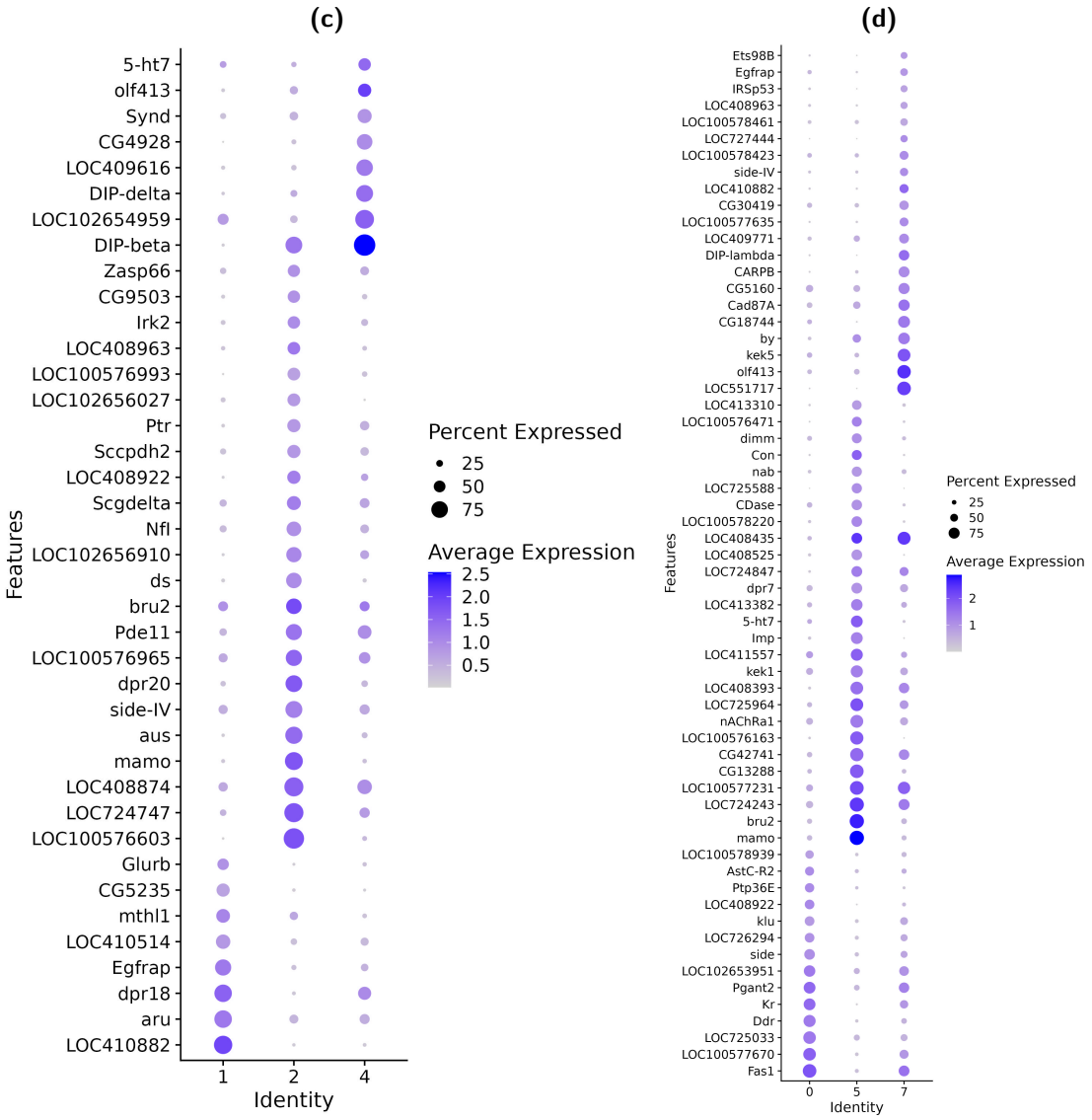


Figure 5.15: Top unique marker genes between different Kenyon Cell subtypes from the literature and between the clusters in the Class-I sKCs and Class-I lKCs.
a - b. Top marker genes between different Kenyon Cell subtypes. a. Grouped by Kenyon Cell subtype. b. Grouped by cluster. Clusters that make up Class-I lKCs have a greater expression of genes such as *nAChRa1* and *Sulphated1* compared the other groups. Clusters that make up Class-I sKCs have a greater expression of *Frq2* and *chinmo* (gene ID LOC551086) compared to the other groups. c. Top unique differentially expressed genes between clusters of Class-I lKCs. Cluster 1 has the highest expression of *Glurb*, cluster 2 has the highest expression of *mamo* and cluster 4 has the greatest expression of *5-ht7*. d. Top unique differentially expressed genes between clusters of Class-I sKCs. Cluster 0 had a greater expression of *AstC-R*, *Fas1* and *Ptp36E*, and cluster 5 had a greater expression of *5-ht7*, *sNPF* (gene ID LOC100576163) and *Imp*.

5.2.12 Expression Patterns of Neurotransmitters, Neuropeptides and their Receptors were Sufficient for Kenyon Cell Classification

This report identified further KC clusters than the KC types previously published in the literature (Suenami et al. 2018). To understand if the additional clusters are functionally distinct KC subtypes, I explored the expression of different neurotransmitters (NT), neuropeptides (NP) and their receptors (NTR and NPR, respectively). For this, I employed a random forest machine learning model to understand the importance of NT, NP, NTR and NPR genes for cell annotation and which genes were most important. Gene names for NT, NP, NTR and NPR were taken from FlyBase (see Methods) and used to filter the cell-by-feature count matrix of the KC clusters extracted from the *Apis* dorsal protocerebra scRNA-seq data. This gave a matrix containing 80250 barcodes and 187 features. Training the model on 80% of the data and testing on 20% gave an accuracy score of 0.97, suggesting the model can accurately classify KC clusters. The model mis-classified cluster 7 as cluster 0, 181 times (Figure 5.16a), but classified cluster 7 correctly 671 times. The misclassification of cells from cluster 7 was likely due to the cells being positioned within the boundary between cluster 0 and 7 and therefore, have acquired similar transcriptional profiles.

The random forest model was repeated using 187 random features from the complete barcode feature matrix. This was done to determine if the features associated with NT, NP, and their receptors had greater clustering weight than a random subset. The accuracy score from using 187 random features was 0.73.

After reviewing the top sixteen NT, NP, NTR and NPR features that were the most important for KC cluster annotation, none annotated a discrete cluster (Figure 5.16b). This suggests the combination of features was necessary to identify KC clusters accurately. The neuropeptide receptor for *RYamide* (*RYa-R*) was expressed in lKCs (cluster 1, 2 and 4) with some expression in cluster 8. *olf413* encodes a dopamine beta-monooxygenase enzyme that is involved in the octopamine biogenesis pathway (Ramya et al. 2024), this

gene was expressed in part of cluster 4 and 8, and broad expression in clusters 7 and 3. Expression of *olf413* in clusters 4 and 8 could suggest some further subdivision of these clusters (Figure 5.16b). The dopamine and ecdysone responsive GPCR *DopEcR* is expressed in lKCs, but also in clusters 6, 0 and part of cluster 8. *DopEcR* is expressed in all *Drosophila* KC types (Croset et al. 2018), which suggest some differences in function between *Apis* and *Drosophila* KCs.

Many features from FlyBase that were used in this model do not directly relate to NT, NP, or their receptors. For example, *Fascilin 2* (*Fas2*) is a cell adhesion molecule that is essential for correct temporal development of adult *Drosophila* MB lobes (Fushima et al. 2007). Although not relevant for this section of this chapter, it is interesting to observe that *Fas2* is highly expressed in Class I KCs and not in the unannotated cluster (likely Class II KCs, Figure 5.16b). As *Fas2* is expressed in $\alpha\beta$ and γ *Drosophila* KCs, this could suggest some developmental similarities between $\alpha\beta$ and γ of *Drosophila* KCs and Class I type *Apis* KCs. An additional filtering step would be beneficial to remove features not directly related to NT, NP or their receptors for a more accurate annotation model.

Overall, the results from this section suggest that the expression pattern of NT, NP and their receptors was sufficient to accurately identify each KC cluster, with few incorrect predictions, suggesting each cluster has a unique expression profile (5.16a). However, Figure 5.16b shows that the most important features in this model for annotating KC clusters do not discretely annotate one cluster; a combination of gene expression is likely required for cluster classification. Further analysis exploring the anatomical position and function of each KC cluster would be important for verifying this additional clustering (see section 5.4).

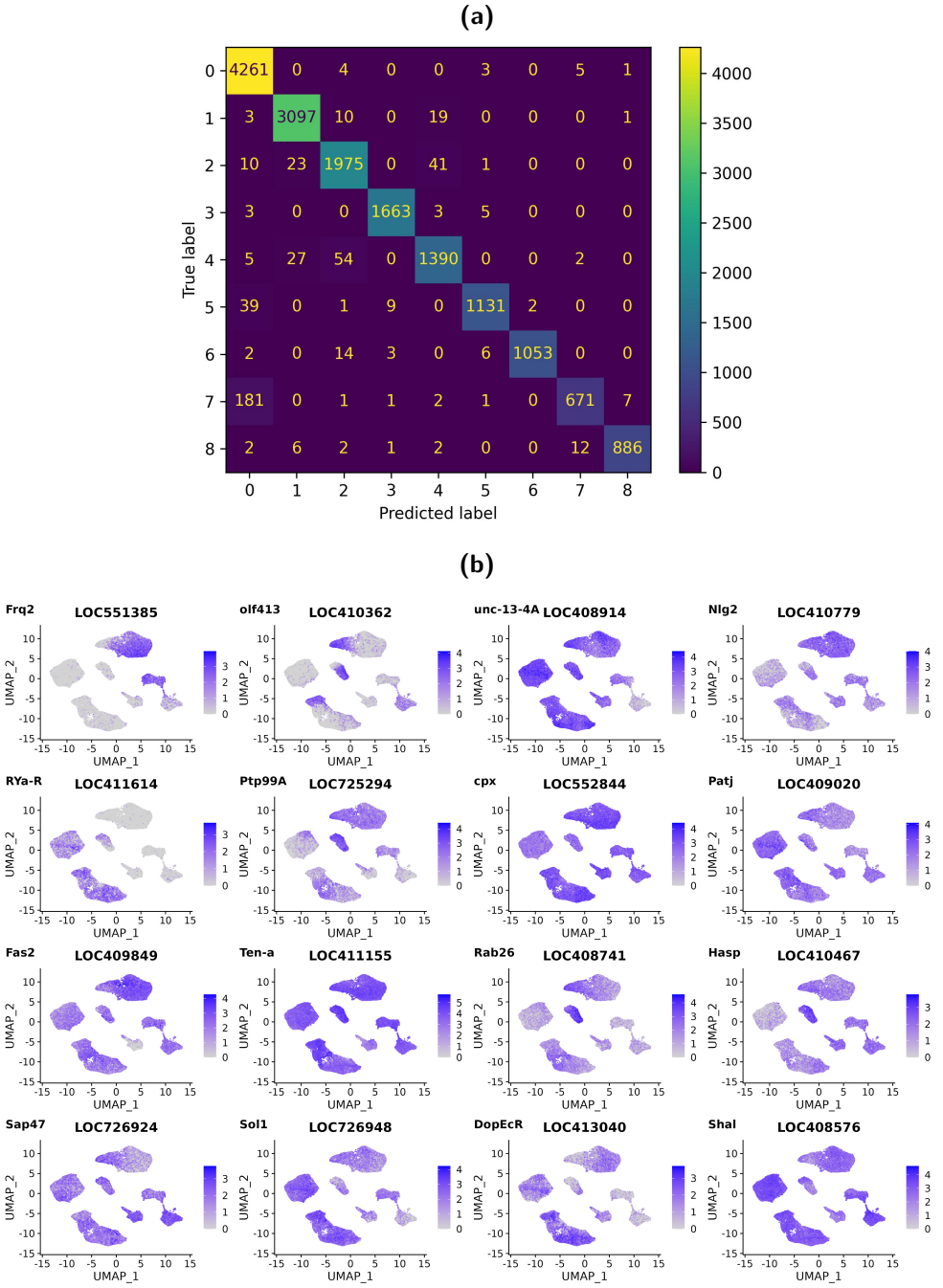
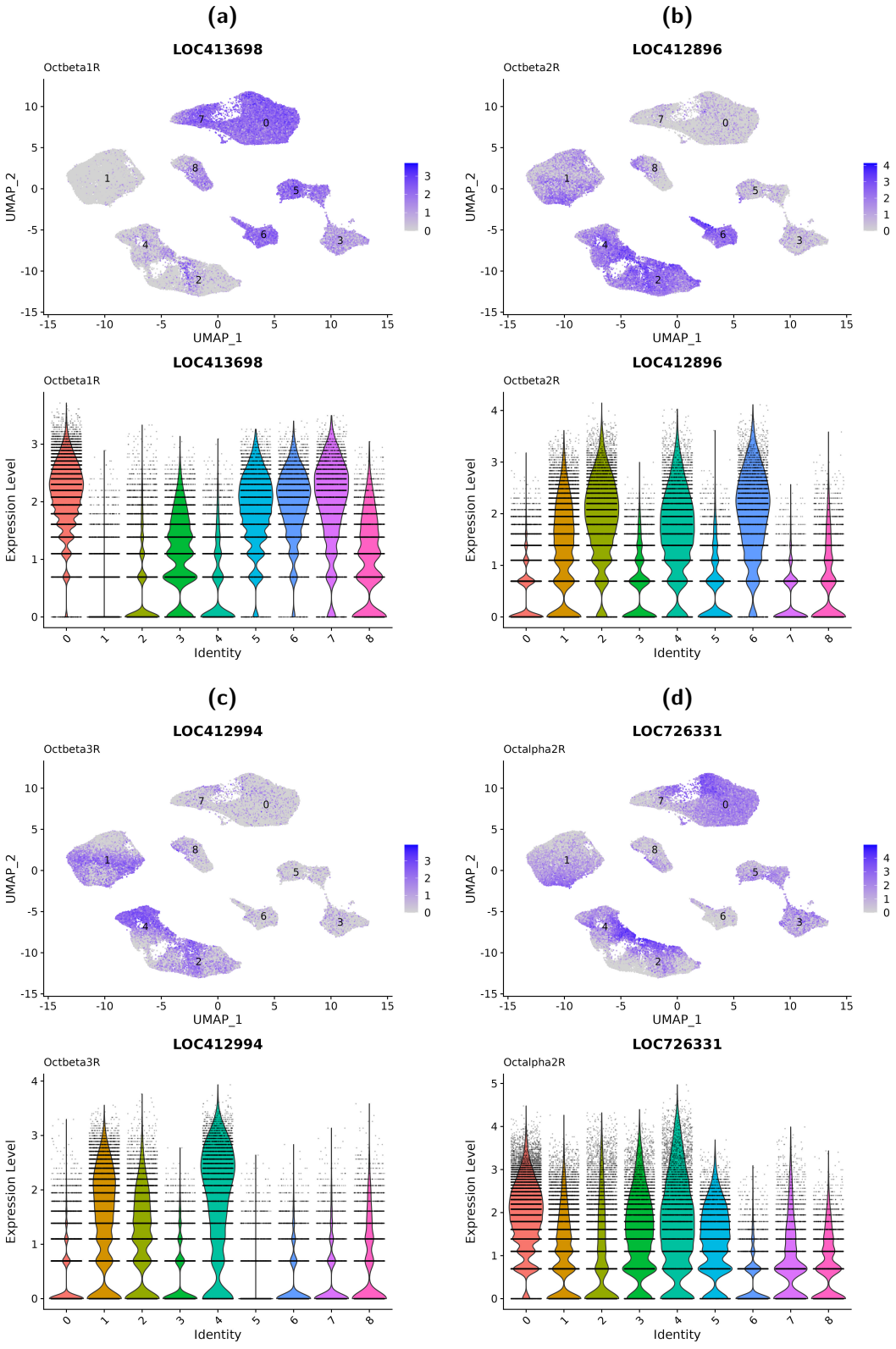
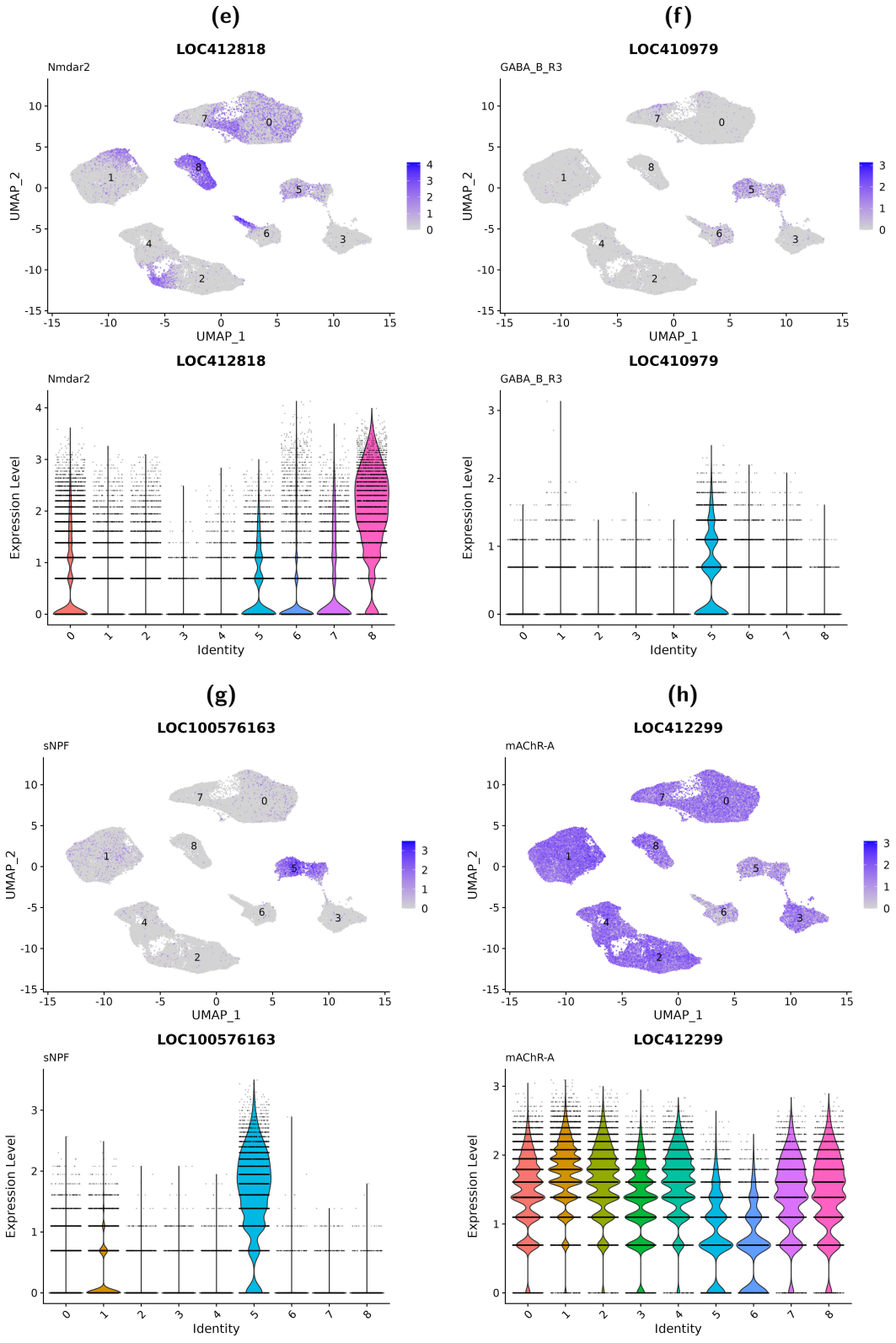


Figure 5.16: A random forest (RF) machine learning model using features associated with neurotransmitters (NT), neuropeptides (NP) and their receptors (NTR and NPR, respectively) was sufficient for accurately classifying Kenyon Cell clusters. The RF model using a feature-barcode matrix from Kenyon Cell clusters, filtered by features associated with NT, NP, NTR and NPR was sufficient to classify Kenyon Cell clusters, with an accuracy score of 0.97. a. Confusion matrix showing the FR model's predictions. The model was effective at predicting the correct labels for most clusters. The model miss-classified cluster 7 as cluster 0 181 times; this was likely a reflection of cells that sit between the boundary between clusters 0 and 7. b. Feature plots of the features most important for classifying Kenyon Cell clusters in rank order starting from the top left.

5.2.13 Octopaminergic Receptors were Expressed in Different Subsets of Kenyon Cell Clusters

Based on the result that the gene expression signature from NP, NT and their receptors was sufficient to classify the KC clusters, marker genes of interest, such as octopaminergic receptors, were plotted to understand their expression across the different clusters (Figures 5.17a - 5.17g). Octopaminergic receptor gene *Octbeta1R* had the lowest expression levels in clusters 1, 2 and 4, moderate expression levels in clusters 3 and 8 and high expression in clusters 0, 5, 6 and 7 (Figure 5.17a). The reverse was true for *Octbeta2R*, however, cluster 6 appeared to express both receptor genes (Figure 5.17d). *Octbeta3R* expression was greatest in clusters 1, 2 and 4 (Figure 5.17c), whereas *Octalpha2R* was expressed in all clusters except cluster 6 and a small amount was expressed in clusters 7 and 8. When plotting other genes of interest, few were cluster-specific. *Nmdar2* expression appeared to be greatest in cluster 8 (Figures 5.17e), an expression pattern that is different from *Nmdar2* expression in *Drosophila* KCs, as scRNA-seq analysis shows a broad expression across all clusters (Shih et al., 2018). The expression of *sNPF* and *GABA-B-3R* was greatest in cluster 5 (Figures 5.17f and 5.17g, see Figure 5.19b for a comparison between *sNPF* expression in *Apis* compared to *Drosophila* KCs). Muscarinic acetylcholine receptors A and B (*mAChR-A* and *mAChR-B*) have broad expression across the different KC clusters. *mAChR-A* was highly expressed in all clusters except clusters 5 and 6 (Figure 5.17h). *mAChR-B* has a lower overall expression, but the highest expression levels in clusters 3 and 6 (Figure 5.17i).


(a)
LOC413698
Octbeta1R
UMAP_2
UMAP_1
LOC413698
Octbeta1R
Expression Level
Identity
(b)
LOC412896
Octbeta2R
UMAP_2
UMAP_1
LOC412896
Octbeta2R
Expression Level
Identity
(c)
LOC412994
Octbeta3R
UMAP_2
UMAP_1
LOC412994
Octbeta3R
Expression Level
Identity
(d)
LOC726331
Octalpha2R
UMAP_2
UMAP_1
LOC726331
Octalpha2R
Expression Level
Identity


(e)
LOC412818
Nmdar2
UMAP_2
UMAP_1
LOC412818
Nmdar2
Expression Level
Identity
(f)
LOC410979
GABA_B_R3
UMAP_2
UMAP_1
LOC410979
GABA_B_R3
Expression Level
Identity
(g)
LOC100576163
sNPF
UMAP_2
UMAP_1
LOC100576163
sNPF
Expression Level
Identity
(h)
LOC412299
mAChR-A
UMAP_2
UMAP_1
LOC412299
mAChR-A
Expression Level
Identity

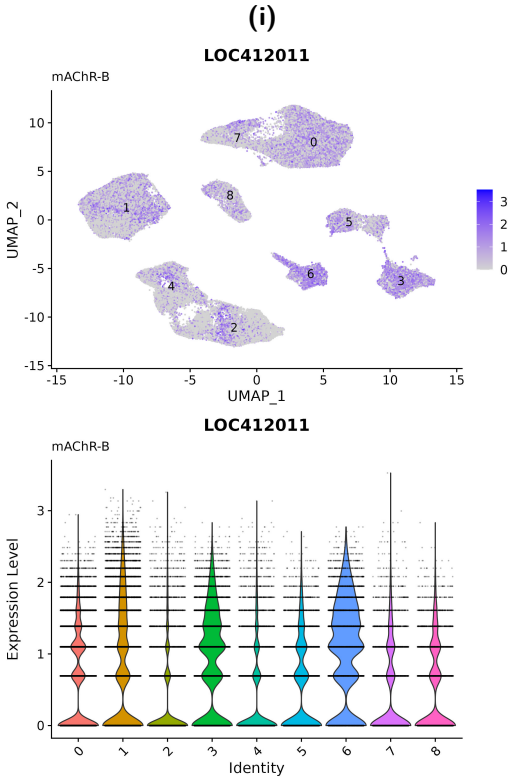


Figure 5.17: Expression pattern of the different octopamine and muscarinic acetylcholine receptors varied across Kenyon Cell clusters, and expression of *GABA-B-R3*, *Nmdar2*, and *sNPF* was highest in one cluster.
Feature and violin plots for the expression of features of interest. a. *Octbeta1R* expression was greatest in clusters 0, 3, 5, 6, 7 and 8. b. *Octbeta2R* expression was greatest in clusters 1, 2, 4 and 6. c. *Octbeta3R* expression was greatest in clusters 1, 2 and 4. d. *Octalpha2R* 0, 1, 2, 3, 4 and 5. e. *Nmdar2* expression was highest in cluster 8. f. *GABA-B-R3* expression was highest in cluster 5. g. *sNPF* expression was greatest in cluster 5. h. *mAChR-A* expression was high in all clusters except 5 and 6, where the expression was lower. i. *mAChR-B* expression was low in all clusters, with some expression in clusters 3 and 6.

5.2.14 The γ and $\alpha\beta$ *Drosophila melanogaster* Marker Gene, *sNPF*, Annotates One *Apis mellifera* Kenyon Cell Cluster

To perform the comparison of transcriptomic profiles between KCs of *Drosophila* and *Apis*, *Drosophila* KC clusters were extracted from the scRNA-seq data published by Park et al., (2022) (Figures 5.18a - 5.18d). As reported by Park et al., (2022), the KCs were identified by the expression of *Mub* and *ey* (Figures 5.18c and 5.18b). The clusters for *Drosophila* KCs, $\alpha\beta$, $\alpha'\beta'$ and γ KCs were identified by markers *ab* for γ KCs, *sNPF* for γ and $\alpha\beta$ KCs, *Ca-alpha1T* and *Eip93F* for $\alpha\beta$, and *CG8641* for $\alpha'\beta'$ KCs. In this section, only *Drosophila* genes with an *Apis* homologue were used. In a first pass of comparing *Drosophila* and *Apis* KCs, the marker genes for *Drosophila* KC sub-clusters were plotted on the *Apis* KC clusters within a UMAP space (Figure 5.19a - 5.19d). Surprisingly, marker genes that annotate discrete subtypes of KCs in *Drosophila* do not annotate specific *Apis* KC clusters. The transcription factor *ab*, appeared to be expressed in all *Apis* KC clusters, but with greater expression in clusters 0, 3 and 7 (Figure 5.19a). *Ca-alpha1T*, was also broadly expressed in *Apis* KC clusters, with slightly lower expression in clusters 5 (Figure 5.19c). *CG8641* had some expression in all clusters but greater expression in clusters 4, 3 and 6 (Figure 5.19d). The marker genes that discretely annotate *Drosophila* KC clusters have broad and overlapping expression patterns in *Apis* KC clusters. This suggests no direct cell type correspondence between *Drosophila* and *Apis* KC clusters. The marker gene for γ and $\alpha\beta$ KCs, *sNPF*, however, had discrete expression within cluster 5 (Figure 5.19b), suggesting some functional division in this cluster compared to other *Apis* KC clusters.

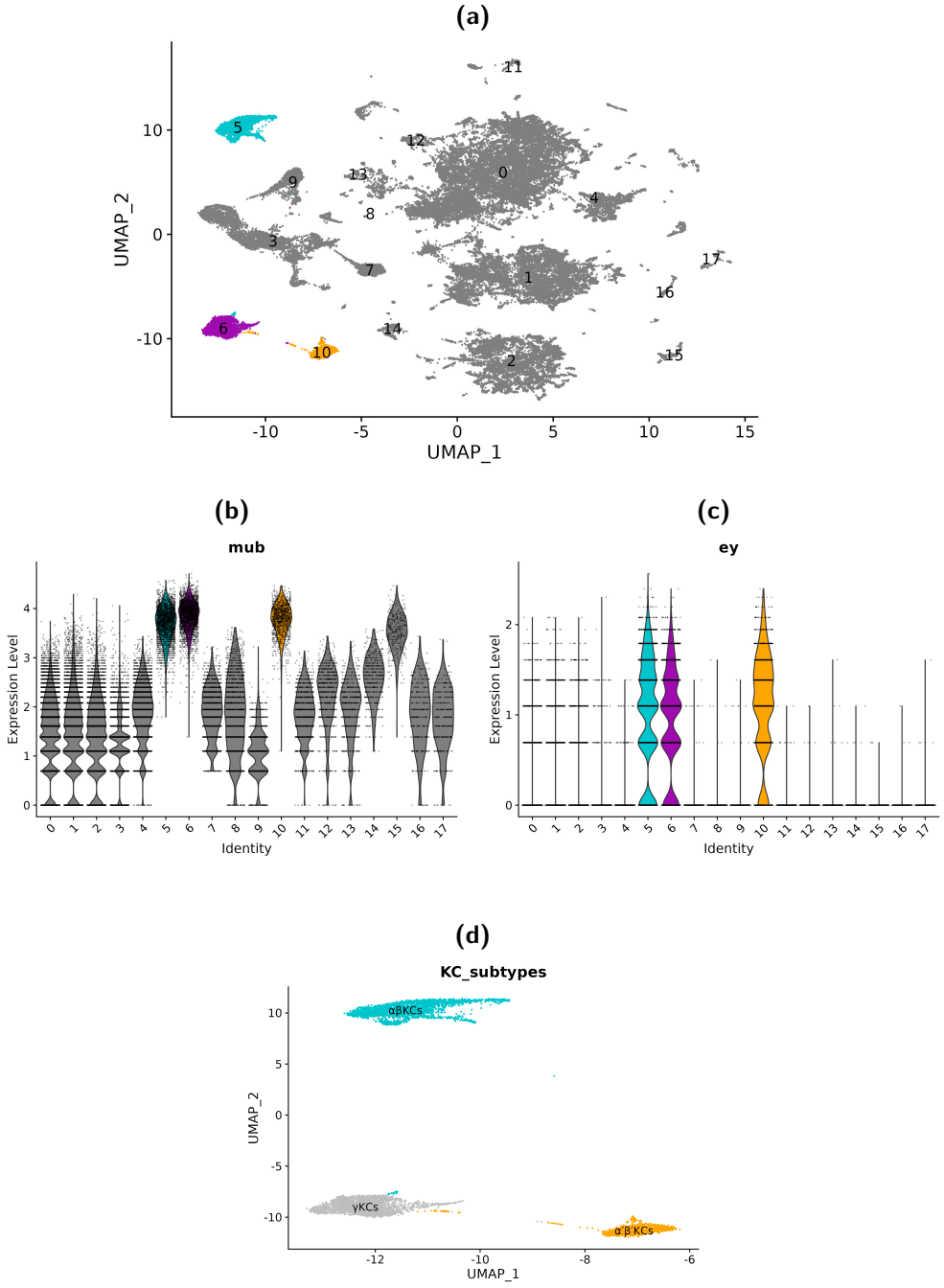


Figure 5.18: Extracting Kenyon Cell clusters from the published single-cell RNA sequencing object of the *Drosophila melanogaster* midbrain from Park et al., (2022). a. UMAP plot of clusters from the full Park et al., (2022) dataset showing eighteen clusters at a cluster resolution of 0.1. Kenyon Cell clusters identified in blue, orange and purple. b. Expression levels of the Kenyon Cell marker gene *mub* and c. expression levels of the Kenyon Cell marker *ey*. d. UMAP plot of the Kenyon Cell sub-clusters, identified by the expression of marker genes *sNPF* for γ and $\alpha\beta$, *ab* for γ, *Ca-alpha1T* for $\alpha\beta$ and *CG8641* $\alpha'\beta'$ Kenyon Cells as shown in Park et al., (2022).

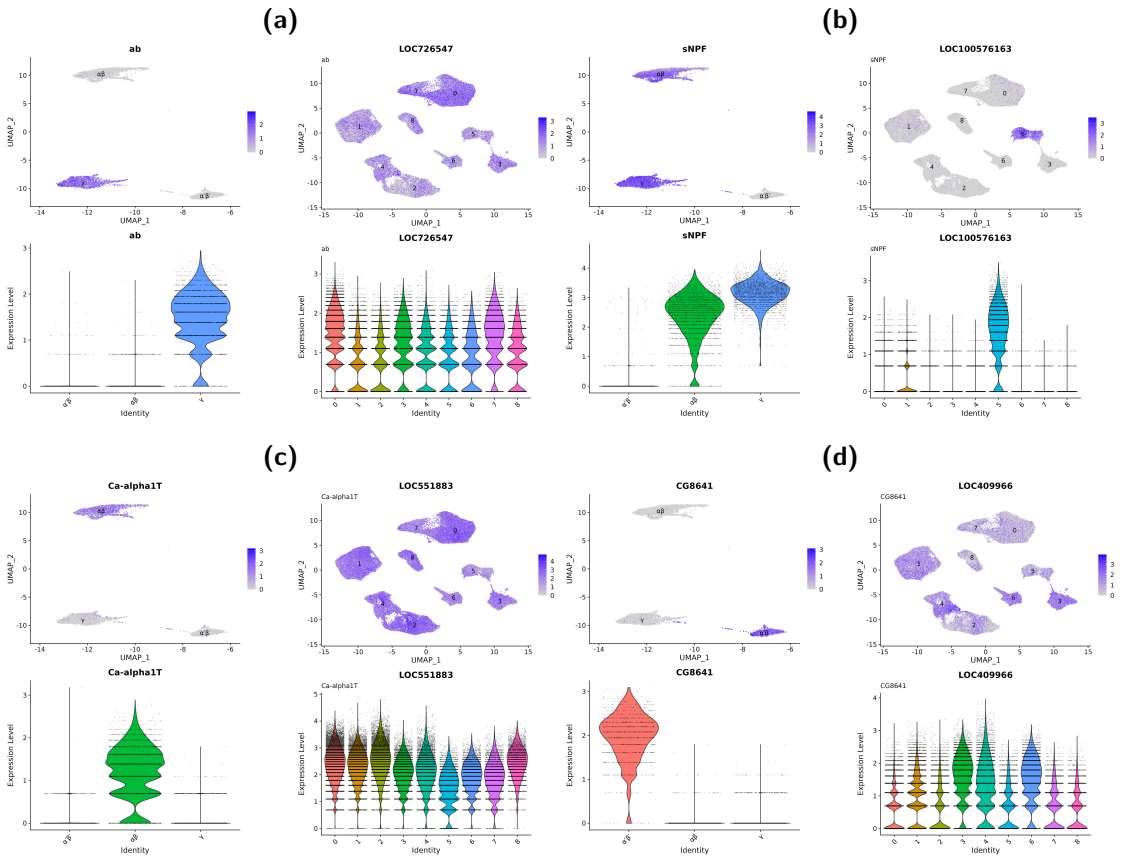


Figure 5.19: Expression pattern of *Drosophila melanogaster* Kenyon Cell marker genes in *Apis mellifera* Kenyon Cell clusters in UMAP space.
a. *ab* the marker gene for *Drosophila* γ Kenyon Cells is broadly expressed in *Apis* Kenyon Cells with the highest expression in clusters 0, 3 and 7. b. *sNPF* the marker gene for *Drosophila* γ and $\alpha\beta$ Kenyon Cells is expressed in *Apis* Kenyon Cell cluster number 5. c. *Ca-alpha1T* the marker gene for *Drosophila* $\alpha\beta$ Kenyon Cells are broadly expressed in all *Apis* Kenyon Cell clusters, with cluster 5 having the lowest expression level. d. *CG8641* the marker gene for *Drosophila* $\alpha'\beta'$ Kenyon Cells is lowly expressed in all *Apis* clusters with the greatest expression in clusters 3, 4 and 6.

5.2.15 Marker Gene Sets for *Drosophila melanogaster* Kenyon Cell Subclusters do not Annotate Discrete Subsets of *Apis mellifera* Kenyon Cells

To further compare *Drosophila* with *Apis* KC clusters, additional marker genes for each *Drosophila* KC cluster were identified using FindAllMarkers() (Figure 5.20), much like the work published in Croset et al., (2018, Figure 2F). Module scores for each *Drosophila* KC cluster were plotted in both *Drosophila* and *Apis* KC clusters in UMAP space. Only features with *Apis* homologues were used. Again, there appeared to be a broad expression profile for all *Drosophila* KC markers in *Apis* KC clusters. However, like the expression of *ab* alone, there was greater expression of γ KC marker genes in clusters 0, 7 and 3 (Figures 5.21a). Like the expression of *Ca-alpha1T* alone, *Drosophila* $\alpha\beta$ KC marker genes appear to annotate all clusters, but with the lowest expression in cluster 5. Like the expression of *CG8641*, clusters 4, 3, 6, but also cluster 1 appear to have a greater *Drosophila* $\alpha'\beta'$ KC marker expression. This could suggest some functional similarities in the *Drosophila* and *Apis* KCs.

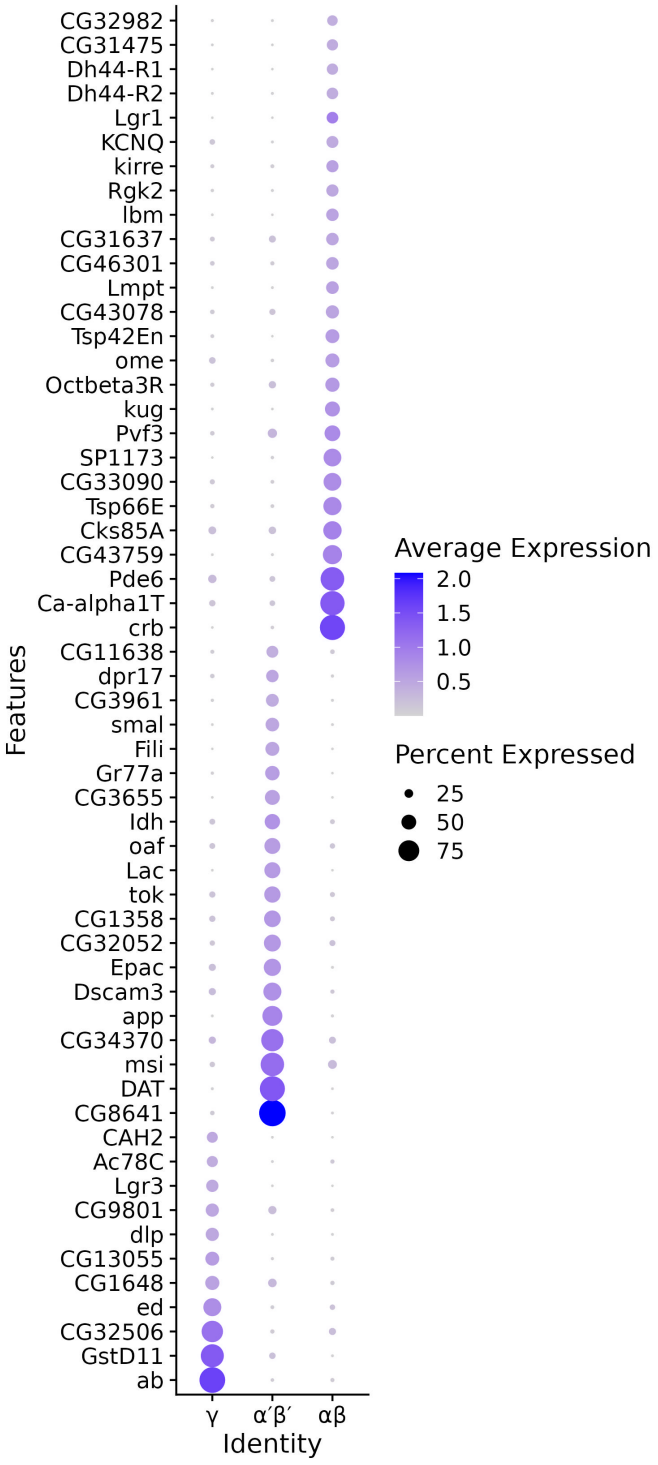


Figure 5.20: Top differentially expressed genes for each *Drosophila melanogaster* Kenyon Cell types. *Drosophila* Kenyon Cell markers were identified using FindAllMarkers(). Genes were filtered for a log2FC change $>$ 2, with $>$30% of cells within the cluster expressing the gene, and $<$20% of cells in the other clusters expressing the gene.

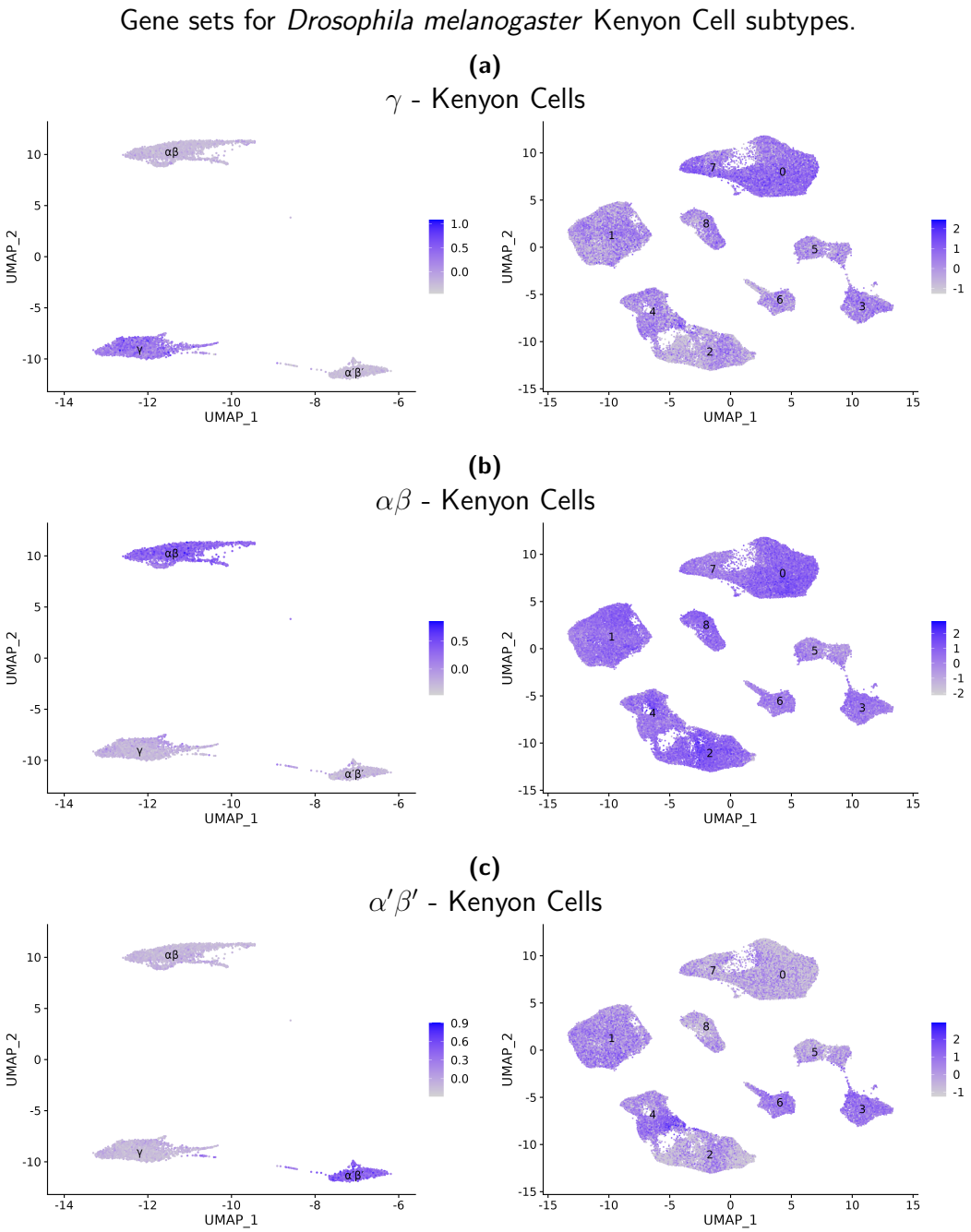


Figure 5.21: Expression pattern of the top marker gene sets for *Drosophila melanogaster* Kenyon Cell subtypes mapped onto *Apis mellifera* Kenyon Cell clusters in UMAP space.
a. Gene set module scores for *Drosophila* γ Kenyon Cells are broadly expressed in all *Apis* Kenyon Cell clusters, with higher expression in clusters 0 and 7. b. Gene set module scores for *Drosophila* $\alpha\beta$ Kenyon Cells are broadly expressed in all *Apis* Kenyon Cell clusters, with lower expression in cluster 5. c. Gene set module scores for *Drosophila* $\alpha'\beta'$ Kenyon Cells are lowly expressed in all *Apis* Kenyon Cell clusters, with higher expression in clusters 1, 3, 4 and 6.

5.2.16 Common Genes that are Important for *Apis mellifera* and *Drosophila melanogaster* Kenyon Cell Type Identity are Involved in Neuronal Organisation and Gene Regulation

To understand if there are common features that underlie *Apis* and *Drosophila* KC clustering, a machine-learning Random Forest model was used to identify which features were most important for cell-type annotation for both species (see Methods). The top 150 features were compared, and only *Apis*, *Drosophila* homologous genes were used. Out of 150 features, 24 (9%) were common between the two species (Figure 5.22). Of these common features, many are involved in gene regulation (*pros*, *mamo*, *Rbp6*, *Imp*, as mentioned in Croset et al. 2018), and cell adhesion, fasciculation and axonal patterning (*Fas2* and *olf413*, Fushima et al. 2007; Ramya et al. 2024). Many of these genes are identified marker genes for *Drosophila* KC sub-types. *mamo* and *msi* are known marker genes for *Drosophila* $\alpha'\beta'$ KCs (Croset et al. 2018; Liu et al. 2019; Li et al. 2022). Figure 5.23b shows that there is indeed high expression of these genes in the *Drosophila* $\alpha'\beta'$ KCs, however, *mamo* also has high expression in γ KCs. Expression of *Fas2* signalling is high in the $\alpha\beta$ lobes of *Drosophila* KCs, with some signalling in the γ lobes (Shih et al. 2018). The reported expression pattern for *Fas2* is echoed in the *Fas2* expression pattern in Figure 5.23b. This suggests that the features identified in this analysis could be good markers for identifying KC subtypes in the *Apis* brain.

This analysis is limited by the homologues between *Apis* and *Drosophila* documented in BioMart. For example, *LOC100576163*, the *Apis* gene ID for *sNPF*, was missing from the homology data taken from BioMart. It is possible that other important features were missing from this analysis. Future work should also address the features that were uncommon between the two species; this would further our understanding of differences in KC development and function.

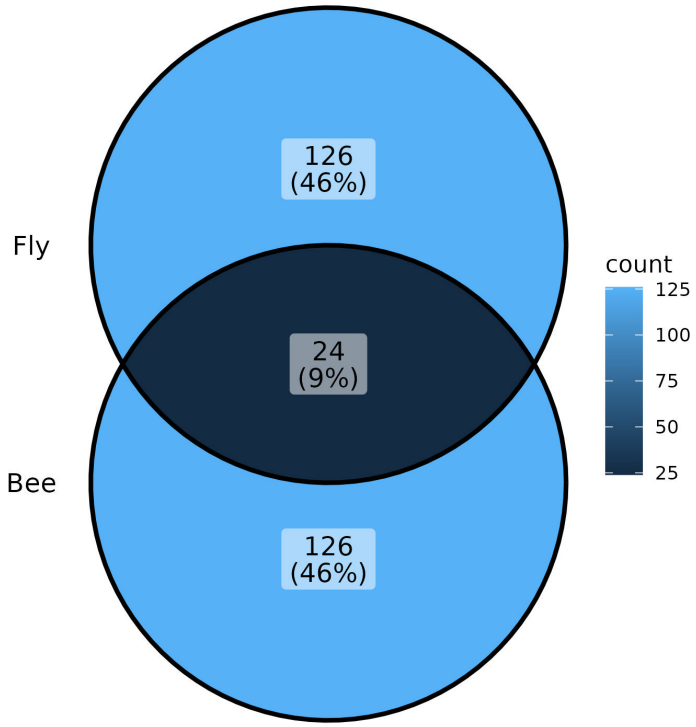
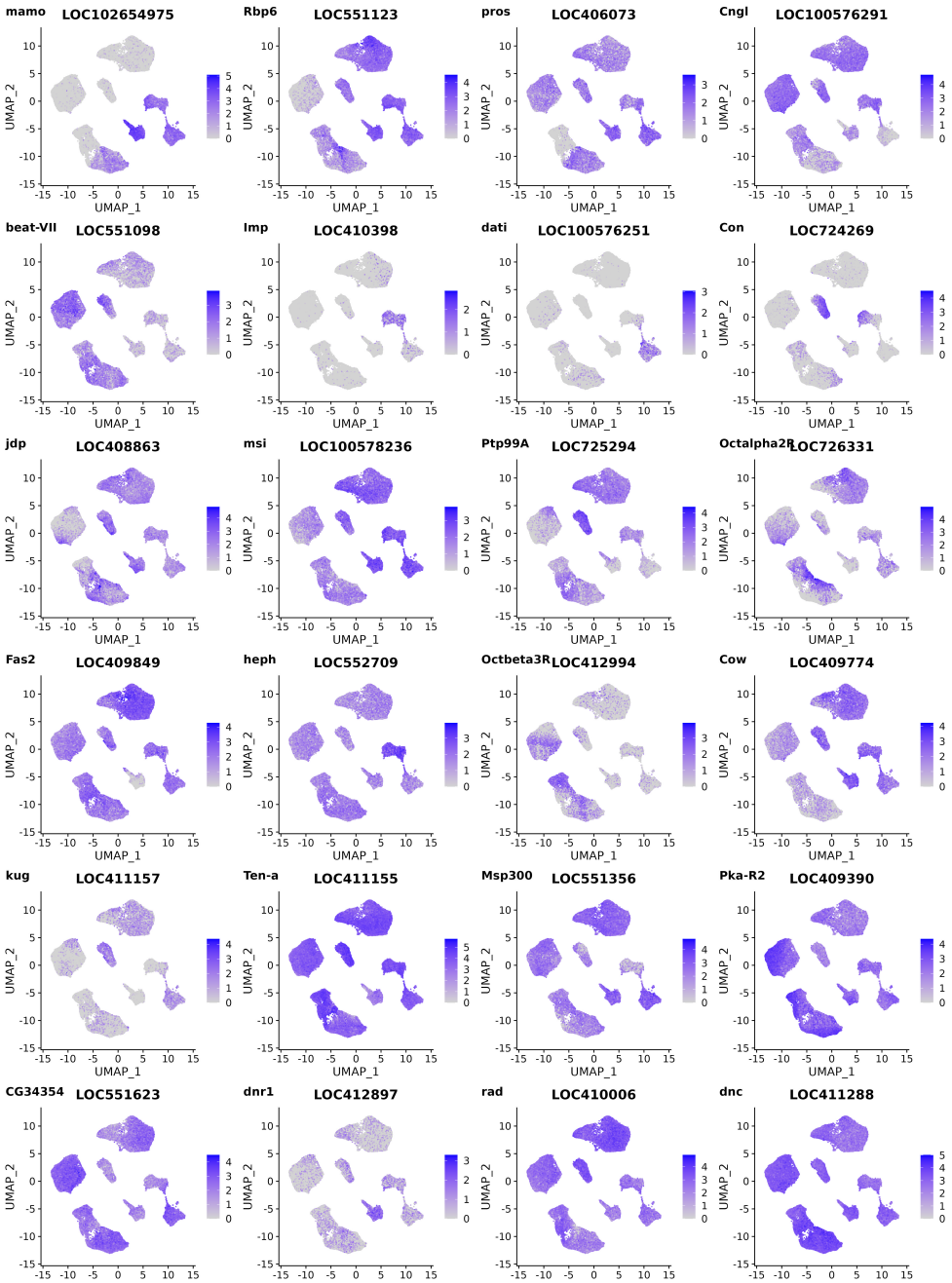


Figure 5.22: Venn diagram representing the number of common features which are important for *Apis mellifera* and *Drosophila melanogaster* Kenyon Cell identity when taking the top-ranked 150 features.
Out of 150 features, only 24 features (9%) are common between *Apis* and *Drosophila* KC clusters.

(a)


mamo LOC102654975
Rbp6 LOC551123
pros LOC406073
Cngl LOC100576291
beat-VII LOC551098
Imp LOC410398
dati LOC100576251
Con LOC724269
jdp LOC408863
msi LOC100578236
Ptp99A LOC725294
Octalpha2R LOC726331
Fas2 LOC409849
heph LOC552709
Octbeta3R LOC412994
Cow LOC409774
kug LOC411157
Ten-a LOC411155
Msp300 LOC551356
Pka-R2 LOC409390
CG34354 LOC551623
dnr1 LOC412897
rad LOC410006
dnc LOC411288
UMAP_1
UMAP_2

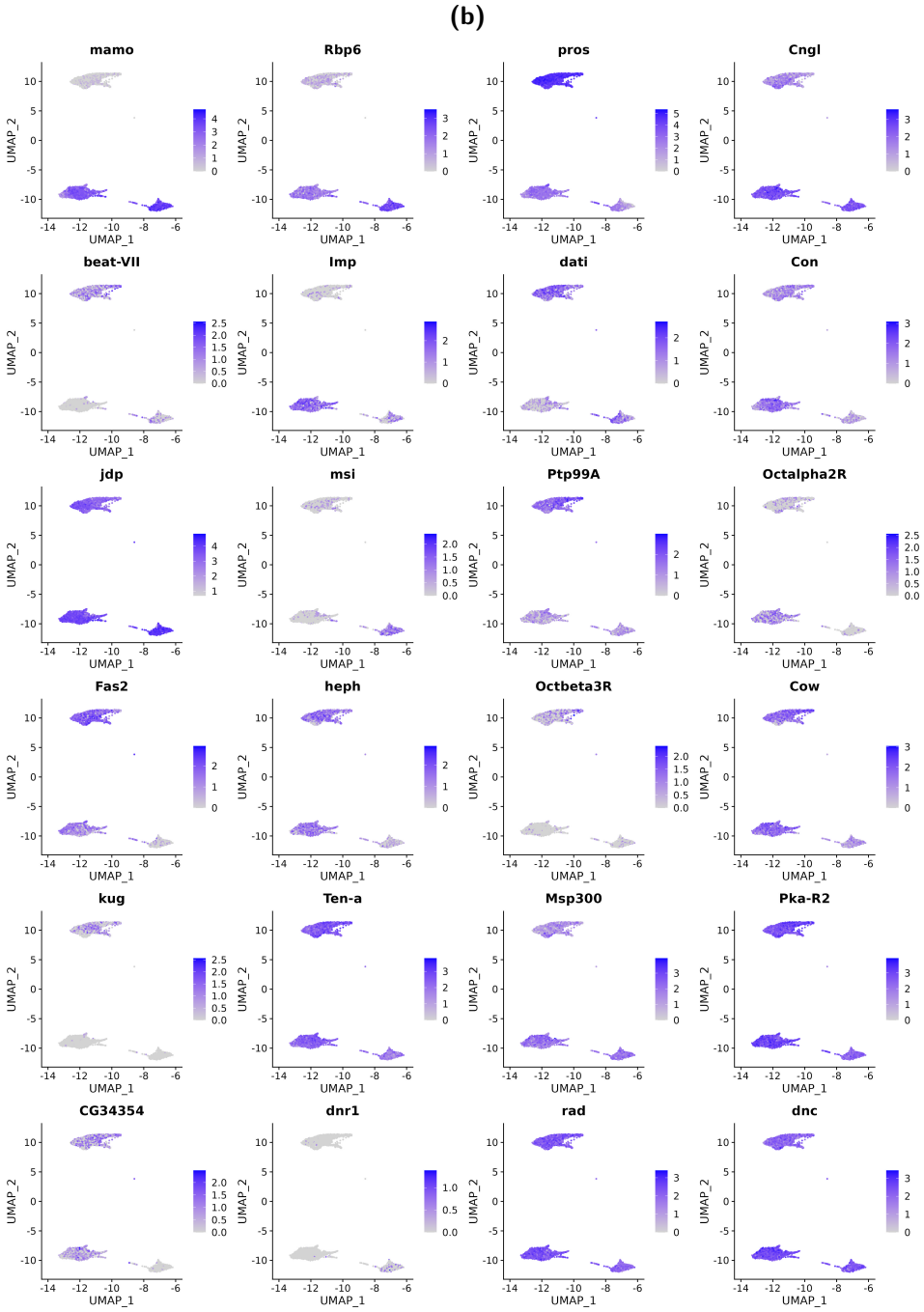
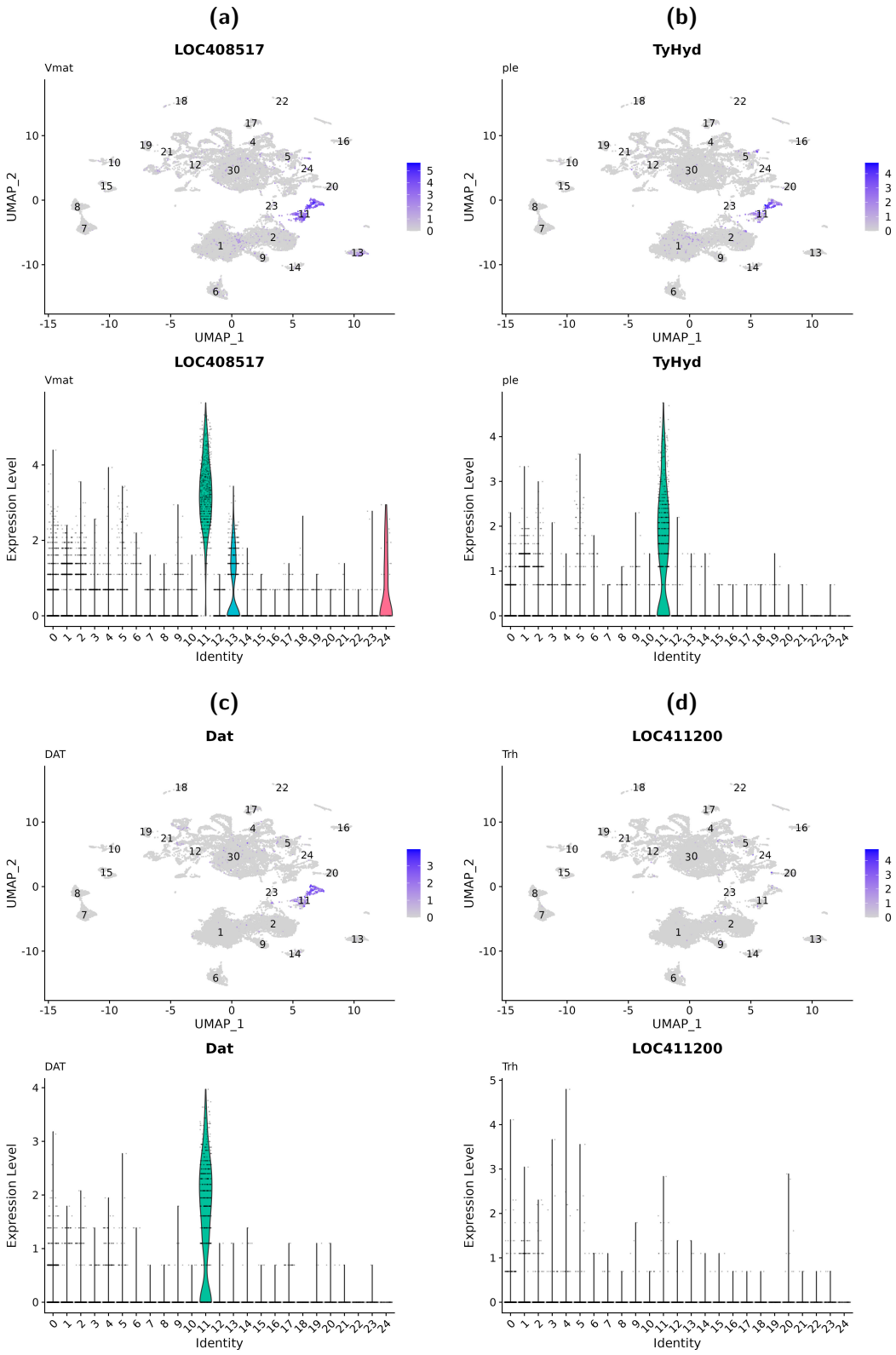


Figure 5.23: Top common ranked features for annotating *Apis mellifera* (a) and *Drosophila melanogaster* (b) Kenyon Cell clusters from a Random Forest machine learning model of a cell-by-feature count matrix using all genes.

5.2.17 Monoaminergic Cells

The monoaminergic cell marker gene *Vmat* identified clusters 11, 13 and 24 as monoaminergic (Figure 5.24a). Dopaminergic marker genes, tyrosine hydroxylase (*TyHyd*) and *DAT*, were highly expressed in cluster 11 (Figures 5.24b and 5.24c). Tryptophan hydroxylase (*Trh*), a serotonergic marker gene and tyramine β-hydroxylase (*Tbh*), an octopaminergic marker gene, were lowly expressed in general, with few cells annotated in cluster 11 (Figures 5.24d and 5.24e). Two *Apis* genes are closely related to the *Drosophila* gene, tyrosine decarboxylase 2 (*Tdc2*, an enzyme required for the production of octopamine and tyramine), *LOC410947* and *LOC410948*, which I will refer to as *Tdc2.1* and *Tdc2.2*, respectively. *Tdc2.1* was lowly expressed in cluster 9 and cluster 24. *Tdc2.2* was lowly expressed in all clusters. Considering octopaminergic neurones VUMmx1 are positioned in the SEZ of the *Apis* brain (Sinakevitch et al. 2013), and this analysis used cells from the dorsal protocerebrum, many octopaminergic neurones were likely removed before sequencing. A comparison between the proportion of octopaminergic cells in the scRNA-seq data in this report and scRNA-seq data collected by Zhang et al., (2022) using the whole pollen-forager *Apis*, brain would be useful for understanding how many octopaminergic cells were lost.


(a)
LOC408517
Vmat
UMAP_2
UMAP_1
LOC408517
Vmat
Expression Level
Identity
(b)
TyHyd
ple
UMAP_2
UMAP_1
TyHyd
ple
Expression Level
Identity
(c)
Dat
DAT
UMAP_2
UMAP_1
Dat
DAT
Expression Level
Identity
(d)
LOC411200
Trh
UMAP_2
UMAP_1
LOC411200
Trh
Expression Level
Identity

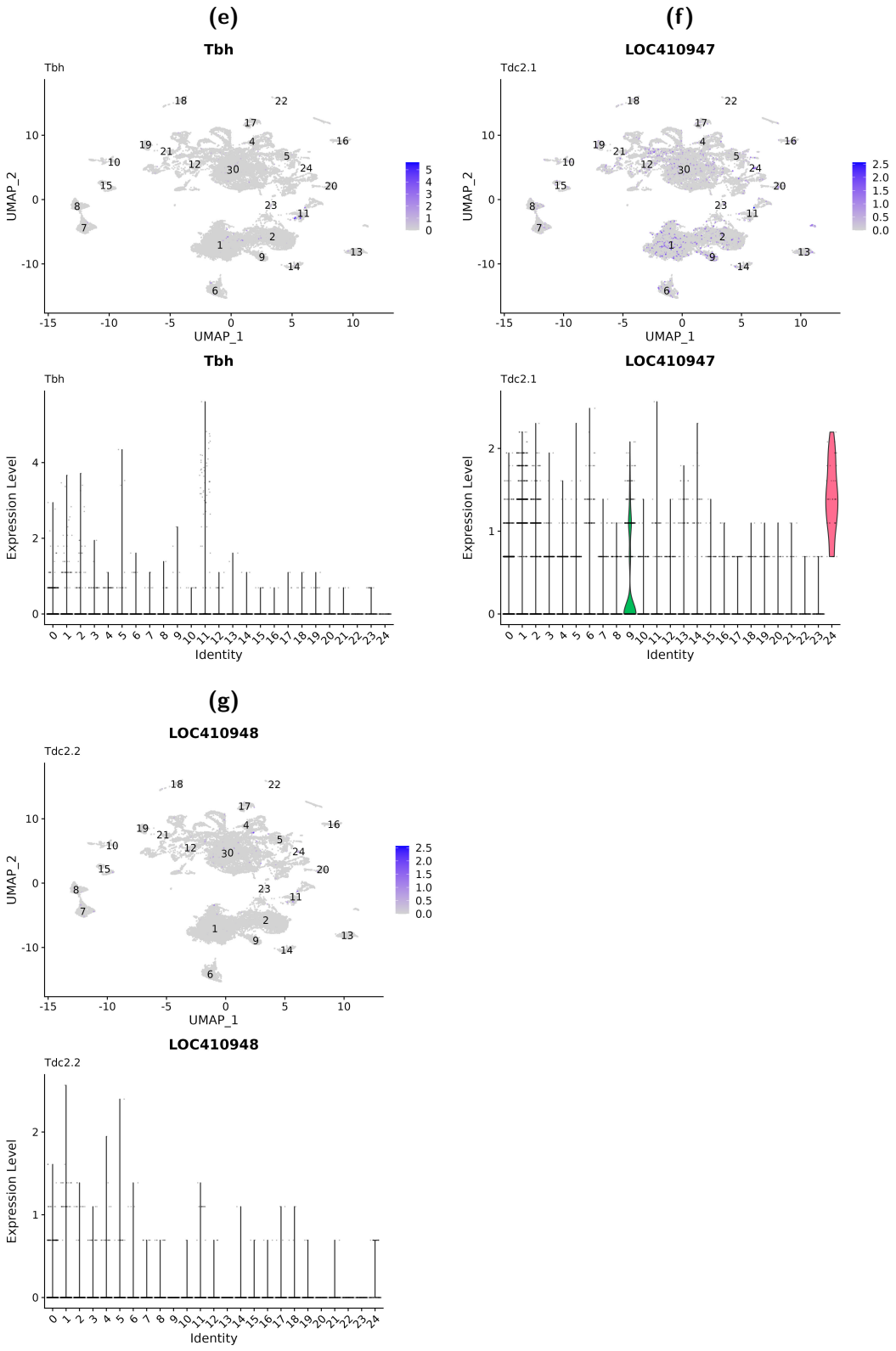


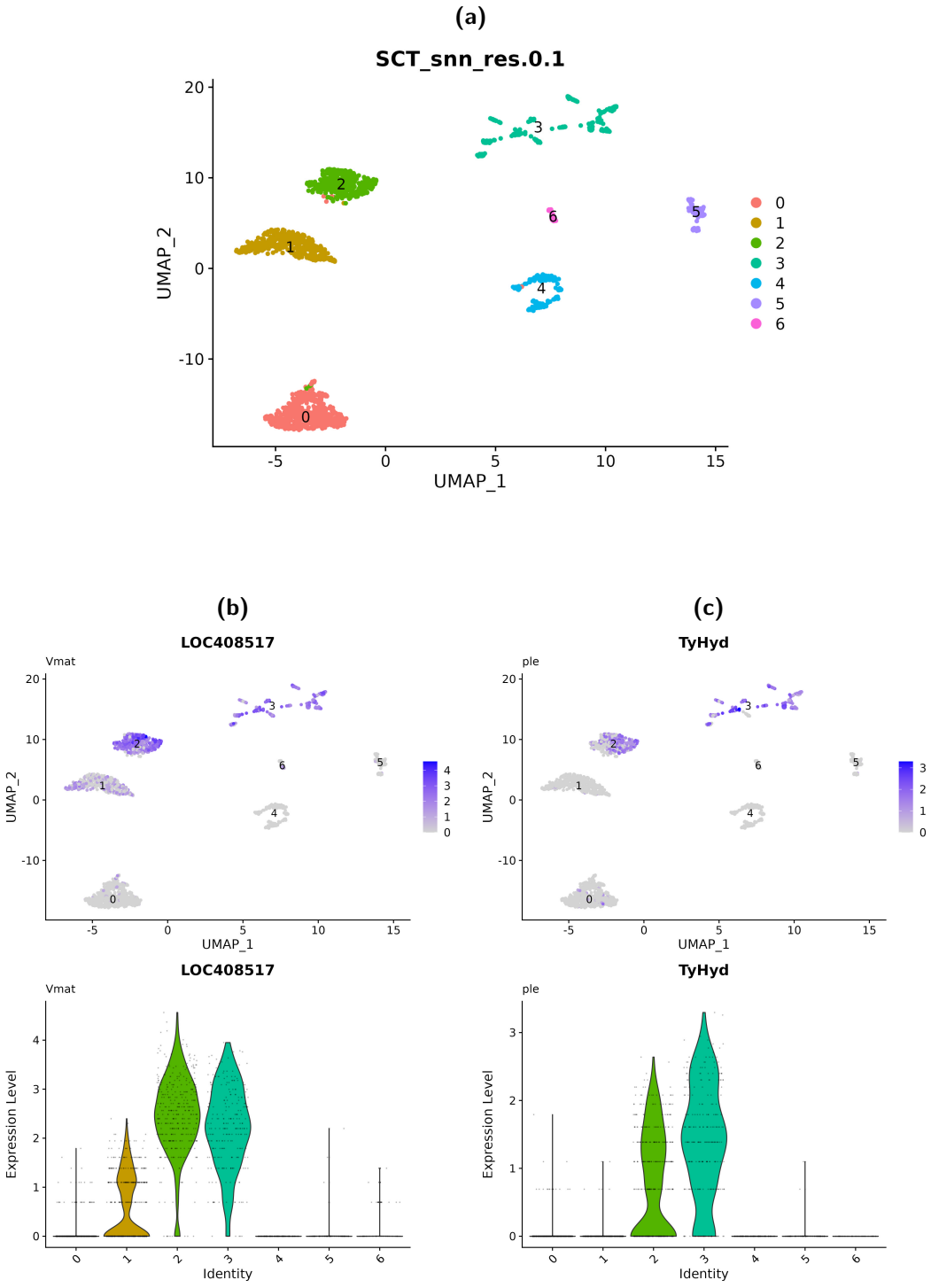
Figure 5.24: Identification of monoaminergic clusters.
a - g. Feature and violin plots for expression of monoaminergic cell marker genes, *Vmat*, *TyHyd*, *DAT*, *Trh*, *Tbh* and two genes associated with *Tdc2*.

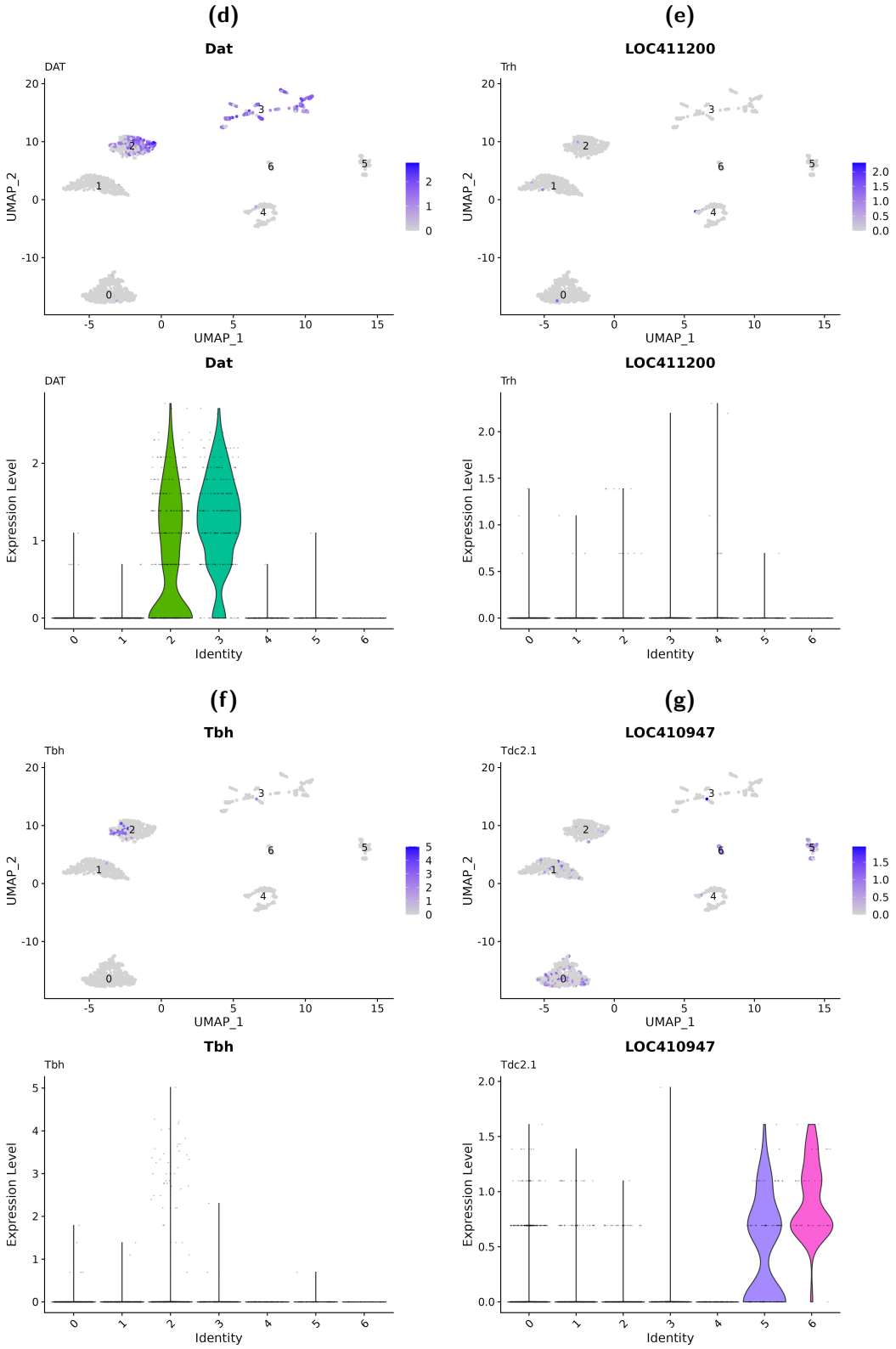
Two *Apis mellifera* Dopaminergic Clusters Expressed *Drosophila melanogaster* Marker Genes for PAM and PPL1-DANs

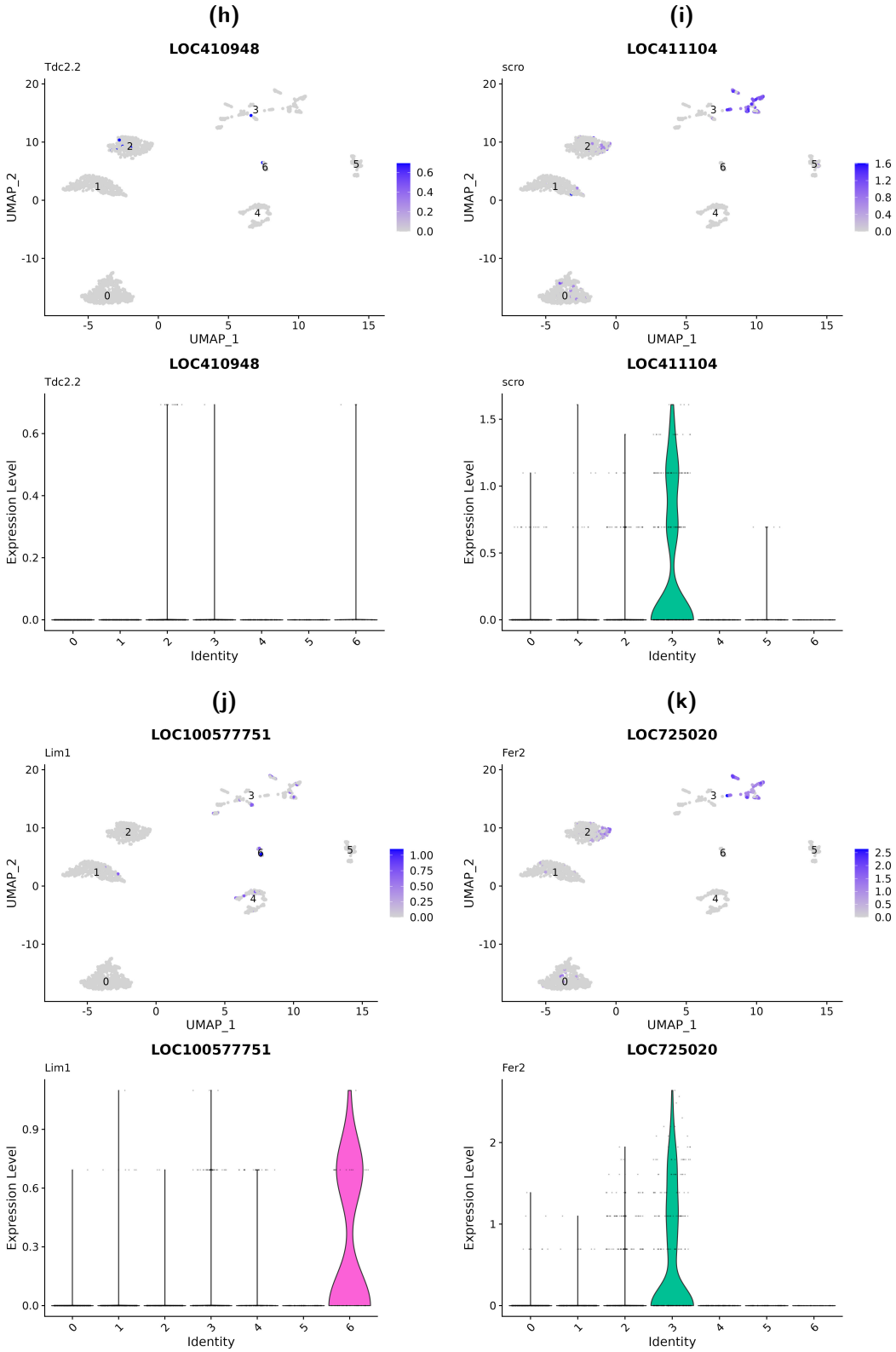
Subsetting and reprocessing clusters 9, 11, 13 and 24 created seven clusters from 1884 cells (Figure 5.25a). Clusters 2 and 3 were *Vmat* positive, with lower expression in cluster 1 (Figure 5.25b). Clusters 3 and half of cluster 2 were also *TyHyd* and *DAT* positive, suggesting these are dopaminergic clusters, respectively (Figure 5.25c and 5.25d). *Trh* was lowly expressed in all clusters (5.25e). *Tbh* was expressed in the portion of cluster 2 that was not positive for dopaminergic markers (5.25f). Clusters 5 and 6 were *Tdc2.1* positive (Figure 5.25g), suggesting these are octopaminergic/tyraminergic cells. *Tdc2.2* was lowly expressed in general (Figure 5.25h).

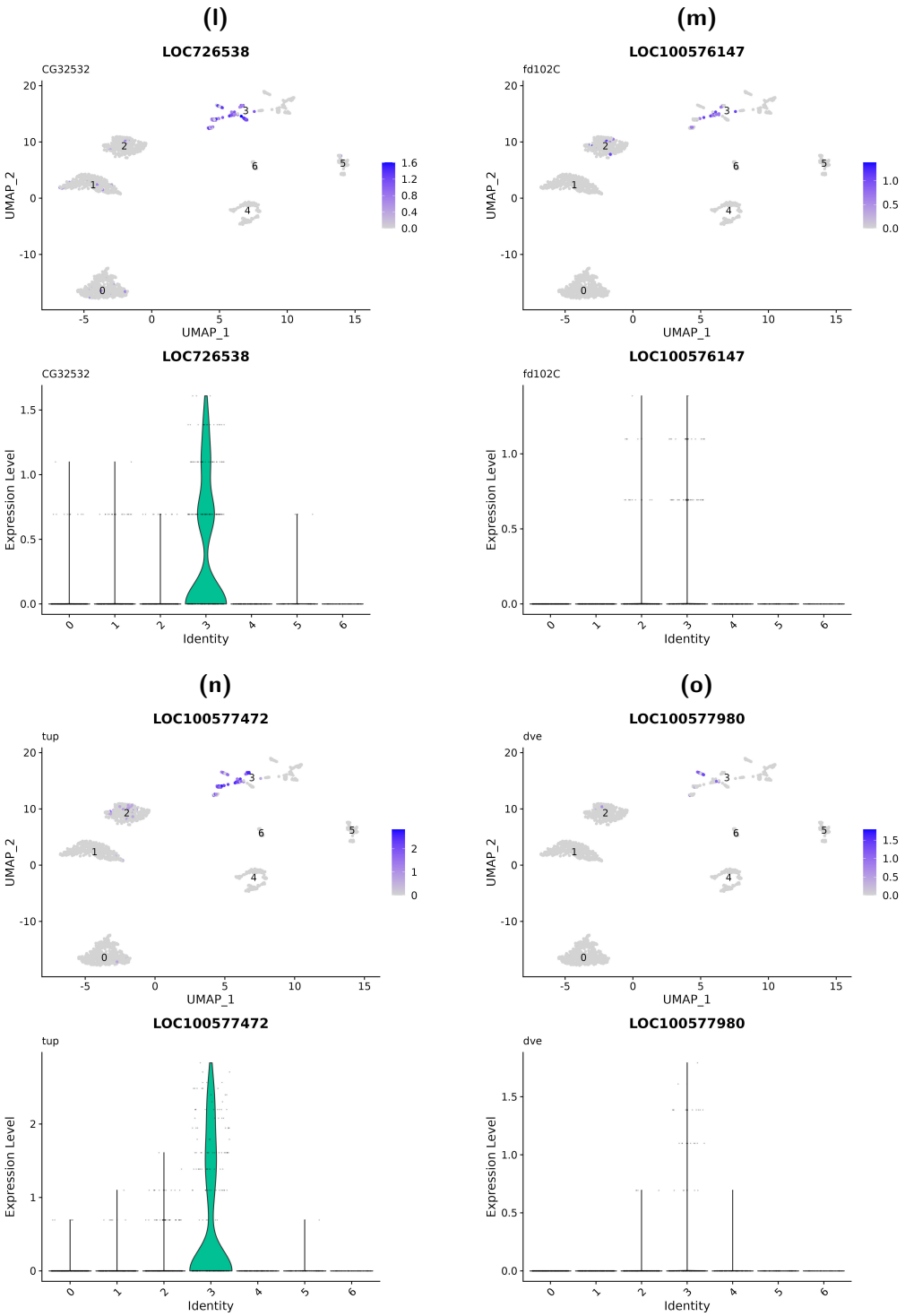
In the *Drosophila* brain, there are two major clusters of dopaminergic neurones that project to the MB, the protocerebral anterior medial (PAM) and protocerebral posterior lateral 1 (PPL1) clusters. The PAM and PPL1 clusters have been shown to provide teaching signals to reinforce appetitive and aversive memories, respectively, during olfactory learning tasks (see section 1.2). PPL1 clusters can be identified from other monoaminergic cells within scRNA-seq data of the *Drosophila* brain by high expression of the marker genes *CG32532*, *fd102C*, *tup* and *dve*. The PAM cluster could be identified with marker genes *scro*, *Lim1* and *Fer2* (personal communication from Dr Christoph Treiber, University of Oxford, 2025 and Bou Dib et al. 2014). The orthologs of PAM marker genes *scro* and *Fer2* were highly expressed in a portion of cluster 3 (Figure 5.25i and 5.25k) and markers *CG32532* and *tup* for PPL1 were highly expressed in a portion of cluster 3 that does not express PAM marker genes (Figures 5.25l and 5.25n). *fd102C* and *dve* were also expressed in the region of cluster 3 that did not express PAM marker genes, but at a lower expression level (Figures 5.25m and 5.25o). *Lim* was lowly expressed in general, but expression was the highest in cluster 6 (Figure 5.25j). As marker genes for dopamine were not highly expressed in cluster 6 (Figure 5.25c and 5.25d), it is likely *Lim1* does not label a subtype of *Apis* DANs.

After reviewing the clustree diagram (Appendix Figure A.8), I selected a new clustering resolution of 0.2, which separates cluster 3 into the PPL1 and PAM expressing clusters (clusters 4 and 5, respectively. See Figure 5.25p), as evidenced by the expression of *CG32532* and *scro* (Figure 5.25q and 5.25r). Cluster 4 (PPL1) has 140 cells and cluster 5 (PAM) has 105 cells.


(a)
SCT_snn_res.0.1
UMAP_2
UMAP_1
0
1
2
3
4
5
6
(b)
LOC408517
Vmat
(c)
TyHyd
ple
Expression Level
Identity


(d)
Dat
DAT
UMAP_2
UMAP_1
Dat
DAT
Expression Level
Identity
(e)
LOC411200
Trh
UMAP_2
UMAP_1
LOC411200
Trh
Expression Level
Identity
(f)
Tbh
Tbh
UMAP_2
UMAP_1
Tbh
Tbh
Expression Level
Identity
(g)
LOC410947
Tdc2.1
UMAP_2
UMAP_1
LOC410947
Tdc2.1
Expression Level
Identity


(h)
LOC410948
Tdc2.2
UMAP_2
UMAP_1
LOC410948
Tdc2.2
Expression Level
Identity
(i)
LOC411104
scro
UMAP_2
UMAP_1
LOC411104
scro
Expression Level
Identity
(j)
LOC100577751
Lim1
UMAP_2
UMAP_1
LOC100577751
Lim1
Expression Level
Identity
(k)
LOC725020
Fer2
UMAP_2
UMAP_1
LOC725020
Fer2
Expression Level
Identity


(l)
LOC726538
CG32532
UMAP_2
UMAP_1
LOC726538
CG32532
Expression Level
Identity
(m)
LOC100576147
fd102C
UMAP_2
UMAP_1
LOC100576147
fd102C
Expression Level
Identity
(n)
LOC100577472
tup
UMAP_2
UMAP_1
LOC100577472
tup
Expression Level
Identity
(o)
LOC100577980
dve
UMAP_2
UMAP_1
LOC100577980
dve
Expression Level
Identity

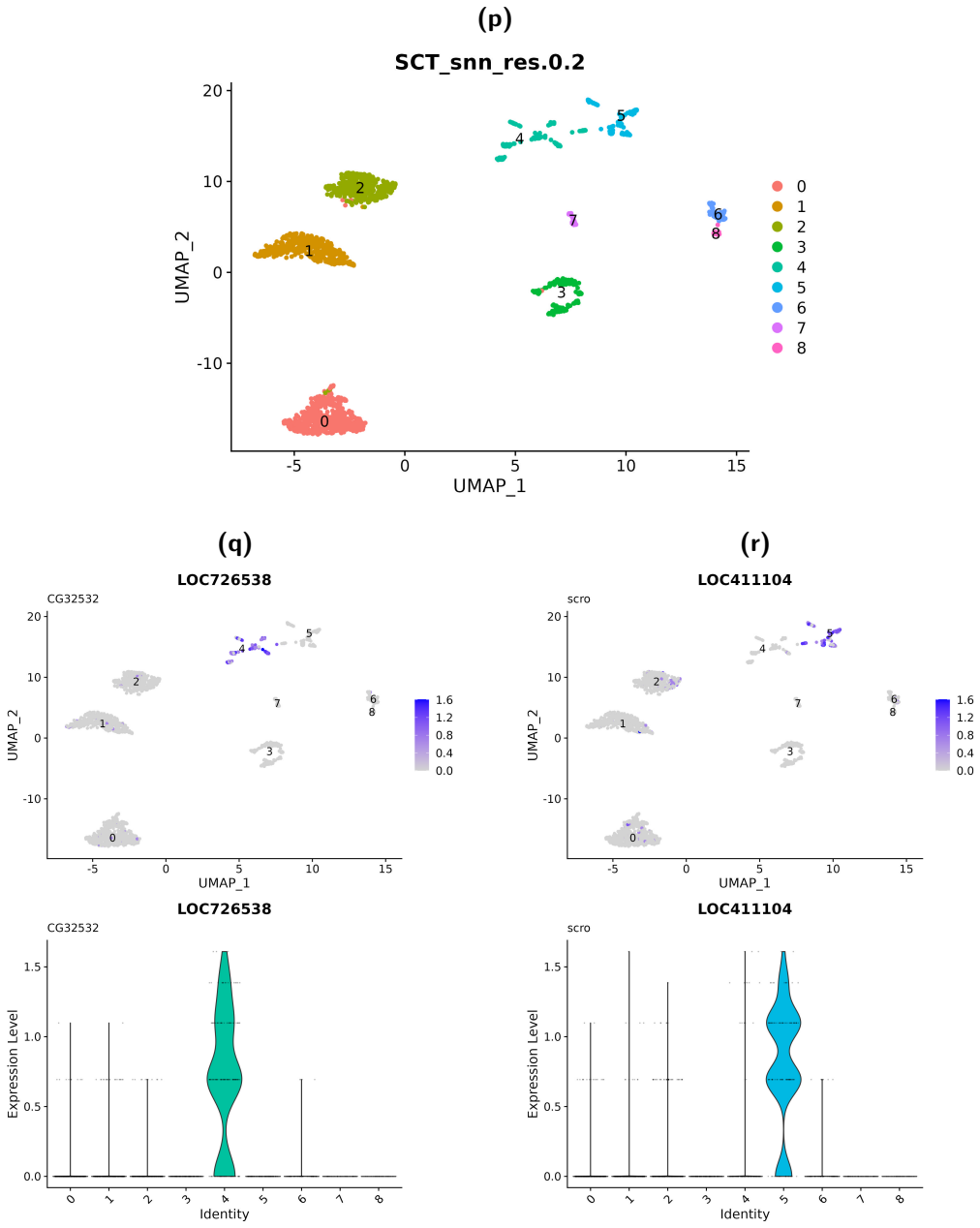


Figure 5.25: Identification of *Apis* monoaminergic clusters and the expression pattern of marker genes associated with *Drosophila melanogaster* protocerebral anterior medial (PAM) and protocerebral posterior lateral 1 (PPL1) DANs.
a. Reintegrating the monoaminergic clusters produced seven clusters from 1884 cells. b - h. Feature and violin plots for the expression of monoaminergic marker genes, *Vmat*, *ple*, *DAT*, *Trh*, *Tbh*, *Tdc2.1* and *Tdc2.2*. i - k. Feature and violin plots for the expression of *Drosophila* PAM-DAN marker genes, *scro*, *Lim1* and *Fer2*. l - o. Feature and violin plots for the expression of PPL1-DANs marker genes, *CG32532*, *fd102C*, *tup* and *dve*. p. Monoaminergic clusters at a clustering resolution of 0.2 produce 9 clusters. q. Cluster 4 was annotated by *Drosophila* PPL1-DAN marker gene *CG32532* r. Cluster 5 was annotated by *Drosophila* PAM-DAN marker gene *scro*.

5.3 Discussion

5.3.1 Results Summary

The research reported in this chapter identified additional granularity to *Apis mellifera* KCs compared to previous published literature. Comparative analysis found no direct correspondence between *Apis* and *Drosophila* KCs; however, the opposite was true for DANs, as marker genes for *Drosophila* PAM and PPL1-DANS labelled subsets of *Apis* DAN clusters.

This report utilised scRNA-seq data from twelve *Apis mellifera* dorsal protocerebra to identify the transcriptomic profiles of the cells involved in learning and memory, and to compare these cells with those from the *Drosophila melanogaster* midbrain. Extracting and reintegrating clusters identified as KCs revealed nine transcriptomically distinct clusters. Of the nine KC clusters identified, three distinct clusters belong to the Class-I lKCs, two clusters belong to the Class-I sKCs, one cluster belongs to the Class-I mKCs, one cluster belongs to both Class-I sKCs and mKCs, and one cluster belongs to the FoxP-KCs. One cluster did not have a high expression of any Class-I marker genes and therefore, could be Class II KCs. A machine learning model based on genes associated with neurotransmitters, neuropeptides and their receptors was sufficient for identifying the KC clusters. Surprisingly, plotting the most important genes for cluster classification did not discretely annotate one cluster; many genes annotated different subsets of KC clusters. This suggests that cluster identity is based on combinations of these features. Reclustering the data from monoaminergic cells, found nine clusters, with two DAN clusters expressing marker genes from the *Drosophila* PAM and PPL1-DAN clusters, cell types that are known for their role in reinforcement learning in *Drosophila*.

5.3.2 Transcriptomic Analysis Identified Nine Transcriptomically Distinct Kenyon Cell Subclusters

Apis KCs are divided into two main classes: Class I somata sit in the calyx cup, and Class II somata are positioned outside the calyx cup. Class I KCs project their axons

to medial and vertical MB lobes. Class II project axons to the γ lobe (Malun 1998; Strausfeld 2002). One KC cluster was not annotated by any Class I KC marker genes, suggesting this cluster could be Class II KCs. This unannotated cluster did not express *Fas2*, whereas all the other clusters annotated with Class I marker genes did. As *Fas2* expression is required in *Drosophila* for the correct development of the MB lobes (Fushima et al. 2007), it suggests that the unannotated cluster developed differently from the other clusters. As the Class II KC develop first (Farris et al. 2004), this gives additional evidence the unannotated cluster is Class II KCs. The scRNA-seq data from this project could be integrated with spatial data of the adult foraging *Apis mellifera* at 22 days, from Mu et al., (2025), using integration tools such as SpaGE (Abdelaal et al. 2020). This will allow us to map marker genes from this analysis onto the spatial data, thus identifying the location of the unannotated cluster. If the unannotated cluster is Class II KCs, I would expect to see expression of marker genes such as *Ekar* and *PIG-K* located outside the calyx cup. This logic could also be extended to identify and confirm the existence of additional granularity to the Class I KCs. This report identified eight Class I KC clusters; this additional granularity is supported by anatomical evidence from Strausfeld, (2002) and Strausfeld et al., (2000) that suggests the vertical lobe is striated with layers containing axons from different KC groups that have dendrites projecting to different zones within each region of the calyx. Immunostaining of FMRFamide and gastrin cholecystokinin (summarised by Strausfeld 2002, Figure 7) shows thirteen layers corresponding to different zones of the calyx. There might be additional granularity to the Class I KCs than found in this report. Identifying the spatial location of the KC subclusters in this report will help understand their function, especially concerning visual and odour learning, as KCs innervating different calyx regions will receive different sensory information. The calyx lip receives olfactory information from the antennal lobe (AL), the collar gets visual information from the optic lobes (OL), and the basal ring receives information from both the AL and OL (Ehmer et al. 2002; Mobbs et al. 1982; Abel et al. 2001; Gronenberg 2001).

5.3.3 Kenyon Cell Clusters and Learning and Memory

In addition to understanding the spatial information of different *Apis* KCs, identifying the expression patterns of neurotransmitters, neuropeptides, and their receptors could also indicate their function. This report found that the expression of many of these genes was not restricted to one cluster. For example, the octopaminergic signalling in *Apis* is suggested to provide the reinforcing signal for appetitive olfactory learning (Hammer 1993; Hammer et al. 1998). The receptors for octopamine were tiled across different KC clusters. Considering that different octopamine receptors have various roles in learning and memory in *Drosophila*, it is likely that this is also true for the octopamine receptors and KC types in *Apis*. For example, Octα1-R (or OAMB) expression in *Drosophila* $\alpha\beta$ and γ KCs rescued the loss in appetitive learning in an *OAMB* mutant *Drosophila* (Kim et al. 2013), but octβ1R expression in $\alpha\beta$ KCs restored aversive memory in *octβ1R* mutant *Drosophila* (Sabandal et al. 2020). *Apis* mutants for different octopamine receptors could be tested with different learning tasks and compared to wild type *Apis*. These experiments could be followed up with calcium imaging of subsets of KCs that express the relevant octopamine receptor, with stimulus presentation (US) and with presentation of the conditioned odour (CS+) (as performed by Szyszka et al. 2008). KCs that are involved in specific associative learning/memory tasks would show a response to the CS+.

Muscarinic acetylcholine receptors A and B were expressed in *Apis* KC clusters (*mAChR-A* and *mAChR-B*, respectively). The expression of *mAChR-A* was broad across most *Apis* KC clusters, with a lower expression in clusters 5 and 6 (Figure 5.17h). The expression of *mAChR-B* was broadly low, with the greatest expression in clusters 3 and 6 (Figure 5.17i). In *Drosophila* scRNA-seq data, the expression of *mAChR-A* was greatest in γ and $\alpha\beta$ KCs (Bielopolski et al. 2019). However, only the expression of *mAChR-A* in γ KCs was required for short-term aversive olfactory learning (Bielopolski et al. 2019). This suggests that this receptor may have different functions in different KC clusters. To understand if muscarinic acetylcholine receptors also have a role in aversive olfactory learning, *mAChR* antagonists (such as scopolamine) could be applied to the brain when

testing for aversive learning in *Apis*, using the sting extension reflex (SER) assay (Vergoz et al. 2007). I would expect that aversive learning would be impaired after applying the antagonist. Muscarinic receptors appear to be involved in other types of memory in *Apis*; application of scopolamine impaired memory recall in PER appetitive olfactory associative learning tasks (Gauthier et al. 1994). Taken together, this evidence suggests that different clusters expressing the *mAChR* receptors have different roles in learning and memory.

Neuropeptides have also been implicated in insect learning and memory. In this report, cluster 5 had high expression of *sNPF*, a neuropeptide expressed in *Drosophila* γ and $\alpha\beta$ KCs, but not in $\alpha'\beta'$ (Johard et al. 2008), and is required for appetitive olfactory memory (Knapek et al. 2013). In *Apis*, topical application of 10 μg μl^{-1} sNPF to the thorax in experiments by Bestea et al., (2022) found that *Apis* will seek visual cues that were previously paired with a sucrose reward, similar to starved *Apis*. Similar results are seen in *Drosophila* when RNAi against sNPF in KCs (tested in all, γ and $\alpha\beta$ KCs) impaired appetitive olfactory learning to sucrose (Knapek et al. 2013). Taken together, this evidence suggests that cluster 5 expressing *sNPF* could likely have a role in appetitive learning in *Apis*.

5.3.4 Expression Patterns of *Apis mellifera* Kenyon Cells Appear to be More Similar to *Monomorium pharaonis* than *Drosophila melanogaster*

The KCs of the *Monomorium pharaonis* ant, another eusocial insect, are split into two main categories, Class-A and Class-B. Class-A KCs, like the lKCs of *Apis* (clusters 1, 2 and 4), have a high expression of *mblk-1*. Class-B, however, does not have strong expression of *Mblk-1*, but has a high expression of *Cow*, *msi* and *dati* (Li et al. 2022); expression which was also seen in *Apis* KC clusters 0, 3, 5, 6, 7 and 8 with expression of *dati* being restricted to cluster 3 (Figure 5.23a). Expression of *cow*, *dati* and *msi* in *Drosophila* do not appear to annotate a specific cluster; expression of each feature annotates a different combination of clusters (Figure 5.23b). This suggests some overlap

in KC composition between *Apis* and *Monomorium*, which are both more distant from *Drosophila*. This reflects the evolutionary history of these species as hymenoptera (*Apis mellifera* and *Monomorium pharaonis*) diverged from Diptera (*Drosophila melanogaster*) nearly 300 million years ago (Maier et al. 2008).

5.3.5 Marker Genes for Dopaminergic Cells Involved in Reinforcement Learning in *Drosophila melanogaster* Annotate Subsets of *Apis mellifera* Dopaminergic Clusters

The current understanding of the neurotransmitters required for olfactory reward or punishment learning differs between *Apis* and *Drosophila*. It was initially thought that insects use octopamine for reward and dopamine for punishment learning, however, this view was challenged in *Drosophila* (as reviewed by Waddell 2013). Although octopamine is required for short-term reward learning for sweet sugars, octopamine acts upstream of the *Drosophila* PAM-DANs, and it is the PAM-DANs that control short and long-term reward learning in *Drosophila* (Waddell 2013). The octopaminergic neurone VUMmx1 arborises on the *Apis* MB and is thought to be responsible for reward signalling to reinforce appetitive memories (Hammer 1993; Hammer et al. 1998). A study by Vieira et al., (2018) also challenges this view in *Apis* as microinjections of antagonists against octopamine or dopamine before reward PER-conditioning experiments reduced the PER response to visual cues paired/ previously paired with sucrose in harnessed *Apis*. This could suggest that the olfactory and visual reward-learning pathways are different, with olfactory reward-learning using octopamine and visual reward learning using octopamine and dopamine. Considering this report identified an *Apis* dopaminergic cell cluster that expressed *Drosophila* marker genes for PAM-DANs, this group of cells may be responsible for signalling reward during visual reward-learning in *Apis*. Electrophysiological recordings of these cells during visual reward learning would be essential for understanding the circuits involved in this type of memory.

5.3.6 Future Work and Conclusions

ScRNA-seq analysis of the dorsal protocerebra from twelve pollen-foraging *Apis* honey bees has allowed for an in-depth analysis of key cells involved in reinforcement learning,

such as the KCs. By having a greater representation of KCs in this analysis, it has identified nine KC subclusters with a unique expression pattern of key genes such as those for neurotransmitters, neuropeptides and their receptors. This suggests additional granularity than previously proposed in the literature (Suenami et al. 2018). To confirm this greater granularity, *in situ* techniques such as *FISH* could be applied in future experiments to understand if the KC subclusters identified in this study match discrete subtypes in *Apis* dorsal protocerebrum. This could be performed using new marker genes identified within this report. Utilising new marker genes from known subtypes, such as *Frq2* which appears to annotate the sKC subtype along with markers for just one of the sKC sub clusters (such as *sNPF*) would identify if there is anatomical subdivision in the sKCs. In addition, the scRNA-seq data from *Apis* dorsal protocerebrum could be combined with spatial data collected by Mu et al., (2025), utilising tools such as SpaGE (Abdelaal et al. 2020) to combine the scRNA-seq data with spatial information. This will help with understanding where the KC subtype cell bodies are positioned.

5.3.7 Study Limitations

The analysis in this report is limited by the incomplete annotation of the *Apis* genome and a direct comparison between *Apis* and *Drosophila* is limited by the annotation of orthologous genes from Biomart. The *Apis* and *Drosophila* KC transcriptomic comparison could also be limited by differences in cell number. For the *Apis* KC dataset, there were 80768 cells, whereas there were only 6561 cells in the *Drosophila* KC data. This analysis was also limited by comparing the *Apis* KCs with the three *Drosophila* KC types rather than the seven subtypes. The KCs of the *Drosophila* MBs can be further subdivided by morphology and genetic markers (Tanaka et al. 2008; Aso et al. 2014; Li et al. 2020a). The γ KCs are split into dorsal and main subtypes; $\alpha\beta$ KCs are split into posterior, middle, surface and core subtypes, and $\alpha'\beta'$ KCs can be split into anterior-posterior-1 and -2, and middle subtypes. Further investigation utilising a combination of published *Drosophila* scRNA-seq data to create a large KC object would be necessary to compare each KC subtype sufficiently. The combined KC scRNA-seq data could be used in parallel with the annotations identified by single-nuclear sequencing analysis of the

Drosophila KC subtypes (Shih et al. 2018). It is possible that the markers for the further *Drosophila* KC sub-clusters could annotate more discrete subsets of *Apis* KC clusters.

Chapter 6

Discussion

Contents

6.1 Thesis Summary

This thesis outlines evidence for a state-dependent memory for NaCl in adult *Drosophila*. The memory of a low concentration of NaCl (0.05 M NaCl) in nutrient-deprived *Drosophila* could be detected immediately, but this memory was not detectable long-term. The memory of high NaCl (0.2 M NaCl) was not detectable immediately, but was detectable after 6 or 24 hours depending on the nutrient-status of the animal before testing. In nutrient-deprived *Drosophila*, the 24-hour memory was appetitive. However, the 6-hour memory in *Drosophila* fed a standard diet was aversive; a behaviour that was more pronounced in females. A differential gene expression analysis of scRNA-seq analysis of female *Drosophila* midbrains found an increase in expression of *Myosuppressin* in nutrient-deprived *Drosophila* compared to those fed 0.2 M NaCl. However, the scRNA-seq data suffered from a high level of adapter contamination throughout reads. More data is required for understanding the role of *Myosuppressin* in signalling NaCl status.

Analysis of scRNA-seq data from twelve pollen-foraging *Apis* protocerebra identified additional granularity to KC subtypes than was previously suggested in the literature. A transcriptomic comparison between *Apis* and *Drosophila* KC clusters found no direct correspondence, suggesting KC clusters between the two species function differently. In

contrast, marker genes for PAM and PPL1 DANs annotated two *Apis* DAN clusters, suggesting these groups of cells could have a similar function.

6.2 A Novel Memory for NaCl in Adult *Drosophila*

As mentioned in previous chapters, a memory for NaCl has been studied in rats, crickets and *Drosophila* larvae (Cone et al. 2016; Robinson et al. 2013; Mizunami et al. 2010; Russell et al. 2011; Niewalda et al. 2008). This report has now identified a memory for NaCl in adult *Drosophila*. Experiments with rodents found a state-dependent memory for NaCl (Cone et al. 2016). Dopamine was detected in the nucleus accumbens (NAc), a part of the brain involved in motivation and reward, with US presentation (intraoral infusion of 0.45 M NaCl) and with the paired CS cue, when sodium deprived (Cone et al. 2016). Zhang et al. (2023) performed scRNA-seq analysis on sections of subfornical organ (SFO) of the lamina terminalis (LT), a region of the brain activated by sodium-deprivation (Hiyama et al. 2004; Ch'ng et al. 2019), in satiated and salt-deprived mice. *Fos* was differentially expressed in a group of SFO excitatory neurones from salt-deprived mice that express the prostaglandin E2 (PGE2) receptor, *Ptger3*. Activation of *Ptger3* expressing SFO neurones in thirsty mice increased consumption of a normally aversive salt solution (0.5 M NaCl) and increased the amount of dopamine released from the NAc. Considering similar logic used for the scRNA-seq experiments from Zhang et al., (2023) to the scRNA-seq experiments outlined in Chapter 4, it is possible that better quality of scRNA-seq data may also allow the identification of cells that are active under sodium or nutrient-deprivation compared to NaCl-satiated *Drosophila*. This could also lead to the identification of cells that modulate MB extrinsic DAN activity during sodium-deprivation. Considering that neuropeptides such as *sNPF* and *Lk* modulate the activity of subsets of dopaminergic neurones to gate state-dependent memories (Krashes et al. 2009; Senapati et al. 2019), it is likely that similar mechanisms gate the expression and/or valence of NaCl memories. Future experiments for this project could be focused towards identifying the populations of DANs involved in reinforcing state-dependent NaCl memories in adult *Drosophila*. Given the identification

of sodium-responsive anterior enteric neurones, INSO neurones (Kim et al. 2024) in *Drosophila*, it would be important to understand how these neurones interact with different subsets of DANs and how this might influence the expression of NaCl memories.

6.3 Greater Heterogeneity of Kenyon Cells in *Apis*

Chapter 5 showed that there is greater heterogeneity in the KCs of *Apis* compared to what was previously thought in the literature (Suenami et al. 2018). The analysis in this report found nine KC clusters, however, it is possible there could be further division to these clusters, as Strausfeld, (2002) found thirteen layers if KC axons in the MB vertical lobe that correspond to different zones of the calyx. An important next step would be to combine the scRNA-seq data from this experiment onto spatial data of the *Apis* brain (Mu et al. 2025). Knowing the position of the KC cell bodies in the brain will be helpful for identifying their function, as visual and olfactory inputs are in different sites in the calyx (Ehmer et al. 2002; Mobbs et al. 1982; Abel et al. 2001; Gronenberg 2001). This could also provide information about what type of inputs could be expected from these regions based on receptor genes expressed by the KC clusters in those locations, such as *Octβ3R* expressed in cells from lKC clusters (Figure 5.17c). Ultimately this could help us understand the KCs and circuits involved in olfactory and visual learning in *Apis*.

6.4 Dopaminergic Signalling in Insects

In *Apis*, DANs provide information to reinforce aversive memories and octopaminergic neurones signal reward. Interestingly, marker genes for *Drosophila* PAM-DANs, a group of cells required for appetitive reinforcement learning in *Drosophila*, annotated one of the three *Apis* DAN clusters, in scRNA-seq data outlined in Chapter 5. Given that Vieira et al., (2018) found that dopamine and octopamine is involved in associative visual reward learning, it is possible that this PAM-like cluster is involved in this type

of memory in *Apis*. Exploring the expression of *Drosophila* PAM-DAN marker genes in scRNA-seq data from other social insects, such as the dataset from Li et al., (2022) of the *Monomorium pharaonis* ant, would be interesting to see if this is a common cell type between social and solitary insects. Tedjakumala et al., (2017) suggested, based on anatomy, that *Apis* DAN cluster C2 corresponds with *Drosophila* PPL1, PPM3 and PPL2ab-DANs. Marker genes for *Drosophila* PPL1-DANs annotated an *Apis* DAN cluster, different to the PAM-like cluster. Plotting *Drosophila* marker genes for PPM3 and PPL2ab-DANs would be beneficial for understanding if this annotates the PPL1-like DANs cluster also. The third *Apis* DAN cluster was not labelled by *Drosophila* PAM or PPL-DAN marker genes. It would be interesting to compare marker genes for this cluster with those from *Drosophila* to understand the correspondence of transcripts from this cluster between the two species. In addition, further granularity to the *Apis* PAM and PPL1-like DAN clusters could be investigated by utilising the transcriptional profiles identified by Aso et al. (2019) from sequencing subtypes of *Drosophila* PAM and PPL1-DANs.

6.5 Conclusion

The research outlined in this report gives an excellent foundation for future experimental work. The work presented in Chapter 3 provides evidence for a novel memory for NaCl in adult *Drosophila* that is affected by internal-state (sex and nutrient-status) and by the concentration of the NaCl solution used as the US. The memory for low NaCl (0.05 M NaCl) in nutrient-deprived flies is detectable immediately after training but the memory was not long-lasting. The memory for high NaCl (0.2 M NaCl) was only detectable 6 and 24 hours after training; the valence of the memory to high NaCl was dependent on the animal's internal state. The memory for high NaCl at 24 hours was appetitive in nutrient-deprived flies, whereas, the memory for high NaCl at 6 hours was aversive in satiated flies, with learning scores being more pronounced in female flies.

Unfortunately, due to adapter contamination in the scRNA-seq data analysed in Chapter 4, it is not possible to draw conclusions from this analysis. This chapter aimed to understand transcriptional changes that occur in cells of the brain in nutrient-deprived compared to NaCl-satiated *Drosophila*. This analysis was extended further to understand if any changes were female specific or different in mated vs. unmated animals. Due to low cell number and UMI count per cell, the ability to detect meaningful differential expression was impaired. The only differentially expressed gene was found in a subset of cholinergic cells of female flies; *Myosuppressin*, was upregulated in nutrient-deprived female flies. Considering that other neuropeptides are involved in signalling nutrient-status and thus influence state-dependent learning, it is possible that this neuropeptide is involved in signalling NaCl status. However, experimental evidence is required to understand if *Myosuppressin* plays a role in state-dependent NaCl learning and memory.

The analysis from Chapter 5 identified nine transcriptionally distinct KC clusters from scRNA-seq data of twelve pollen-foraging *Apis mellifera* protocerebra. Further analysis combining the *Apis* scRNA-seq data from this report with spatial data or smFISH analysis using identified marker genes for *Apis* KC clusters would be helpful to confirm that these clusters are indeed KC subtypes. When comparing the transcriptional profiles of the KC clusters between *Apis* and *Drosophila* there was little correspondence between the two species. However, marker genes for *Drosophila* PAM and PPL1-DAN annotated two *Apis* DAN clusters, suggesting these clusters may have similar functions across species. Further experiments addressing the function of PAM and PPL1 like DAN clusters in *Apis* would be an important next step.

Appendices

Appendix A

Appendix chapters

Contents

A.1 Appendix 1

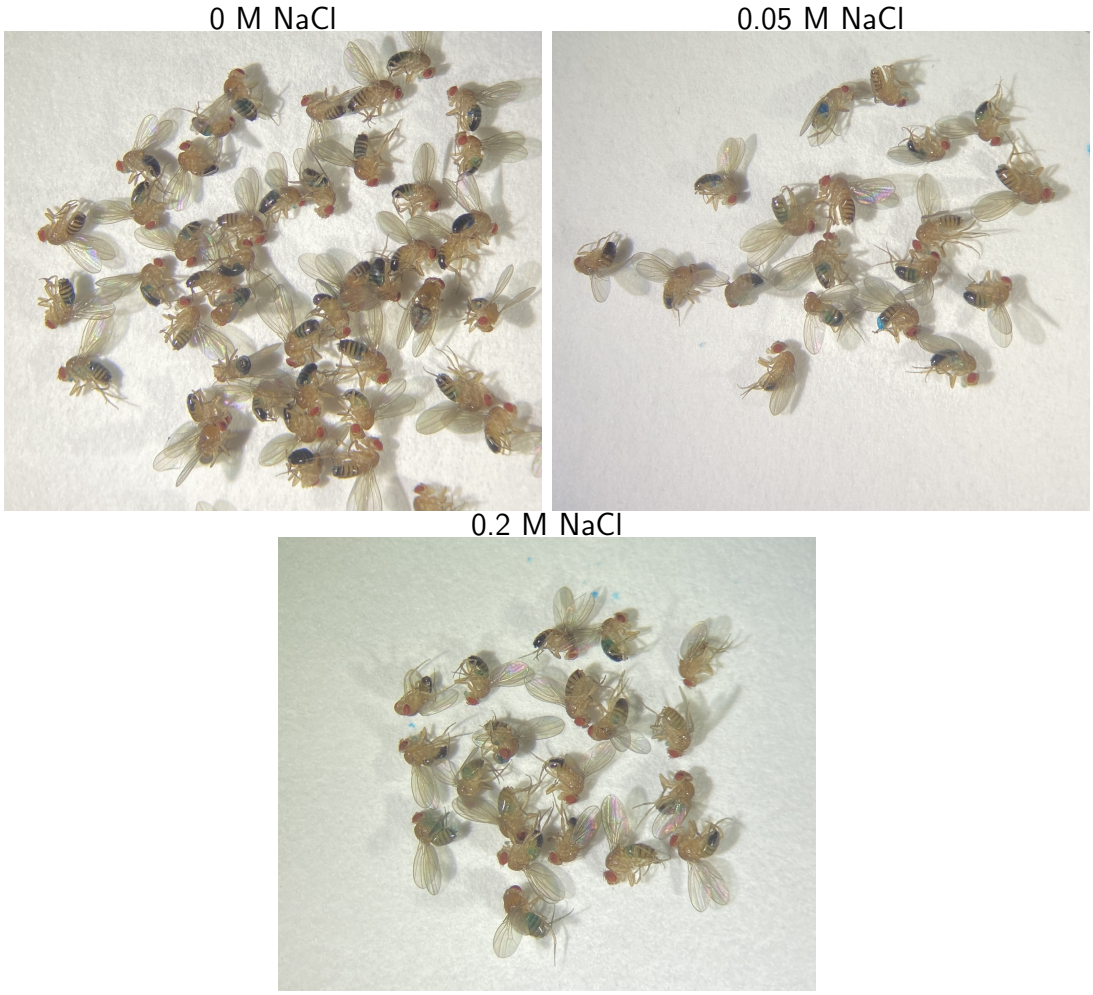
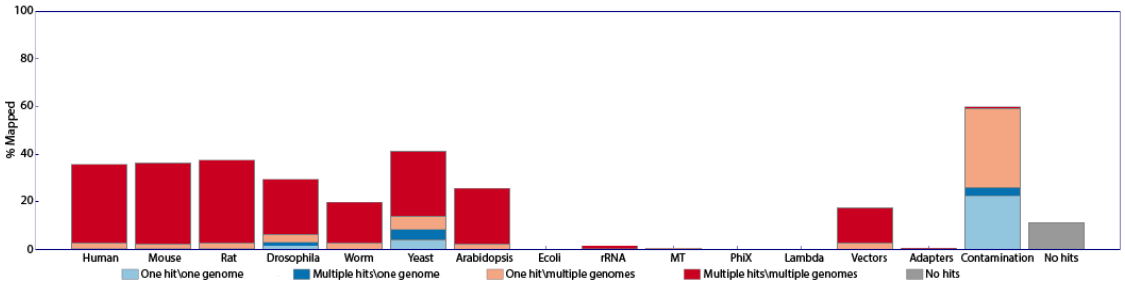


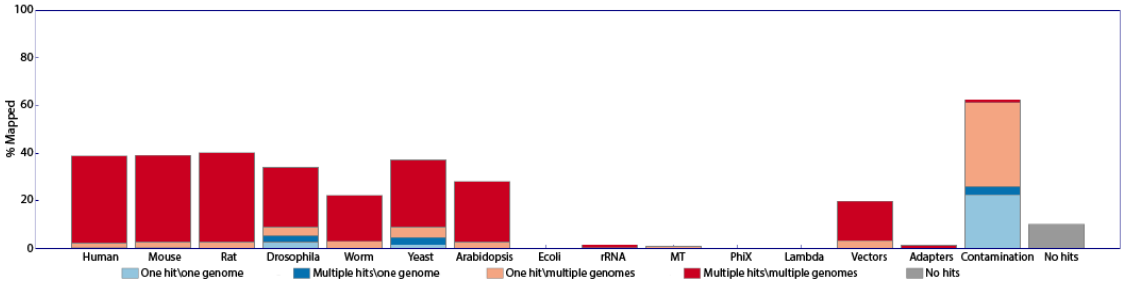
Figure A.1: NaCl feeding assay. 20-hour nutrient-deprived *Drosophila* were fed a solution containing 1% agar, 0.6% FD&C Blue dye No.1, MQ water and NaCl at 0 M, 0.05 M or 0.2 M for 2 minutes in the dark. Data collected by Dr. Bhagyashree Senapati.

A.2 Appendix 2

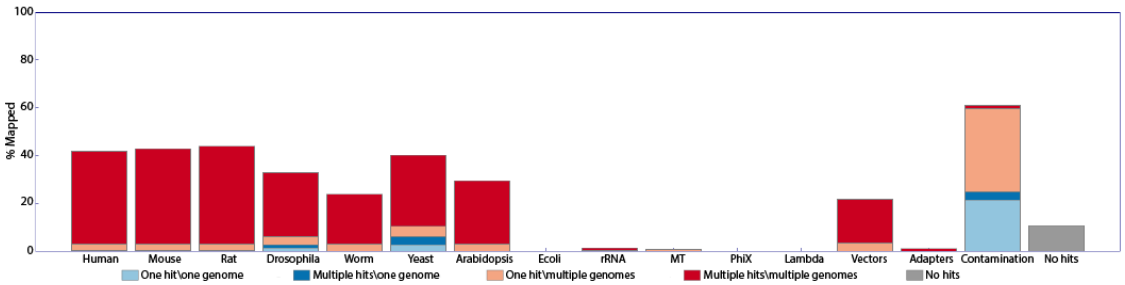
Sample 5



Sample 6



Sample 7



Sample 8

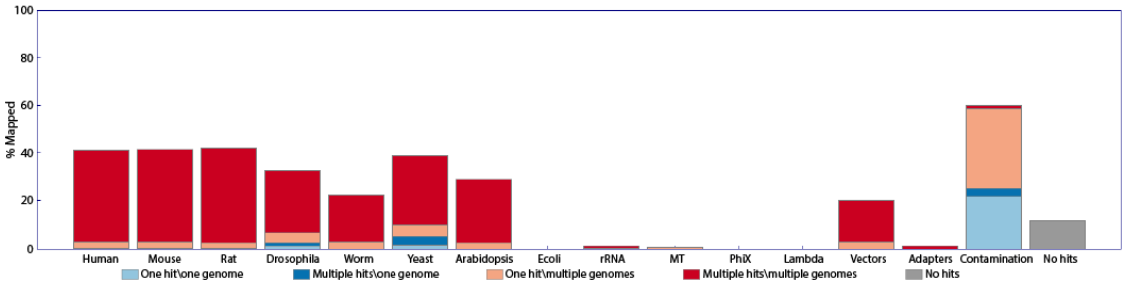


Figure A.2: High level of sequence contamination in reads samples 5-8.

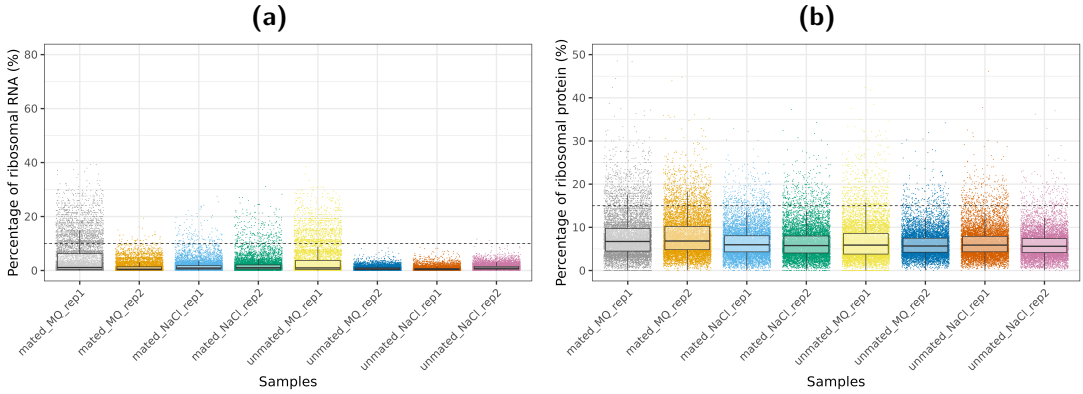


Figure A.3: Percentage of reads from ribosomal RNA and ribosomal protein (%). a. Percentage of reads from ribosomal RNA for each cell per sample. The dotted line shows the cutoff value of 10%. b. Percentage of reads from ribosomal protein for each cell per sample. The dotted line shows the cutoff value of 15%.

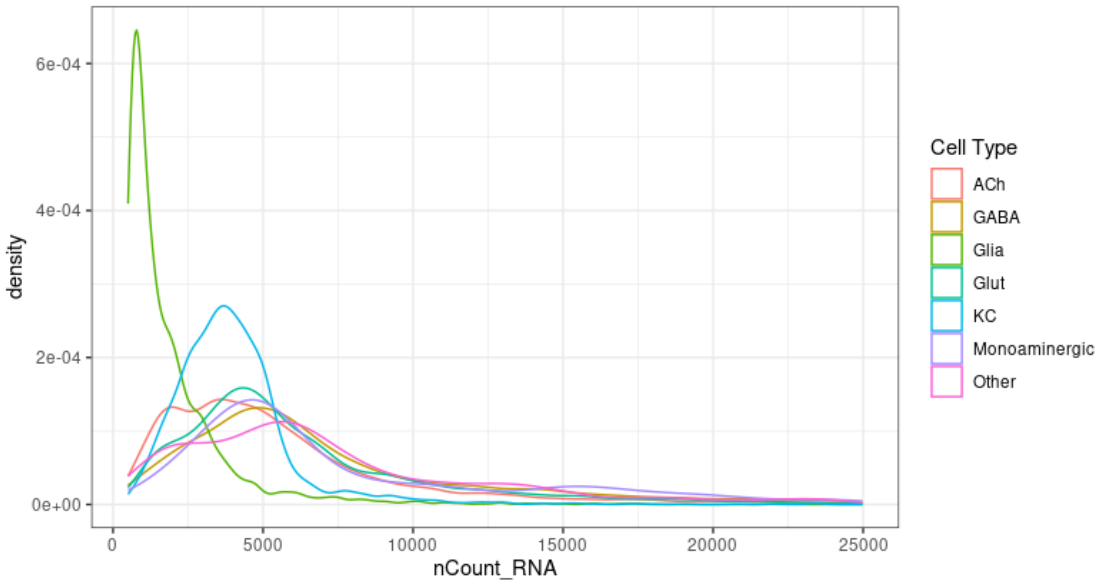


Figure A.4: Distribution plots of UMI counts for each of the major cell *Drosophila* types from the Park et al., 2022 dataset. Glial cells have a greater proportion of cells with lower UMI counts compared to the other cell types.

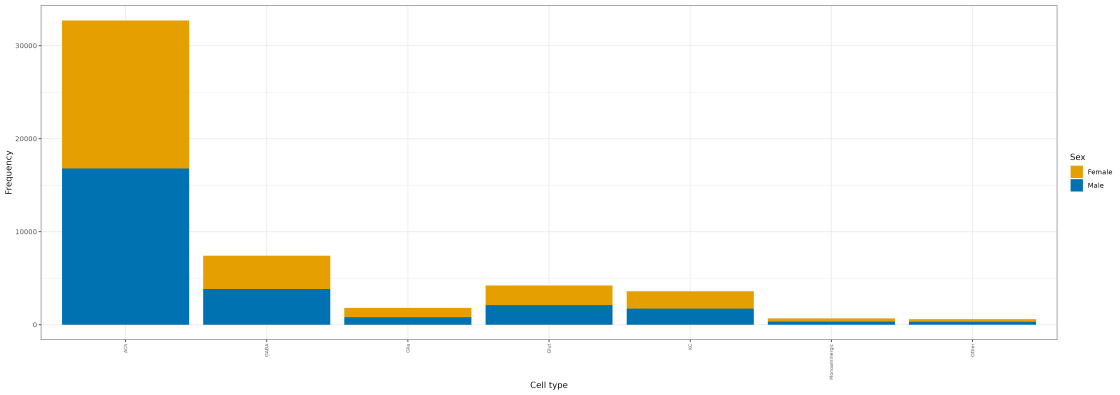


Figure A.5: The distribution of cells classified as male or female in each of the major *Drosophila* cell types.

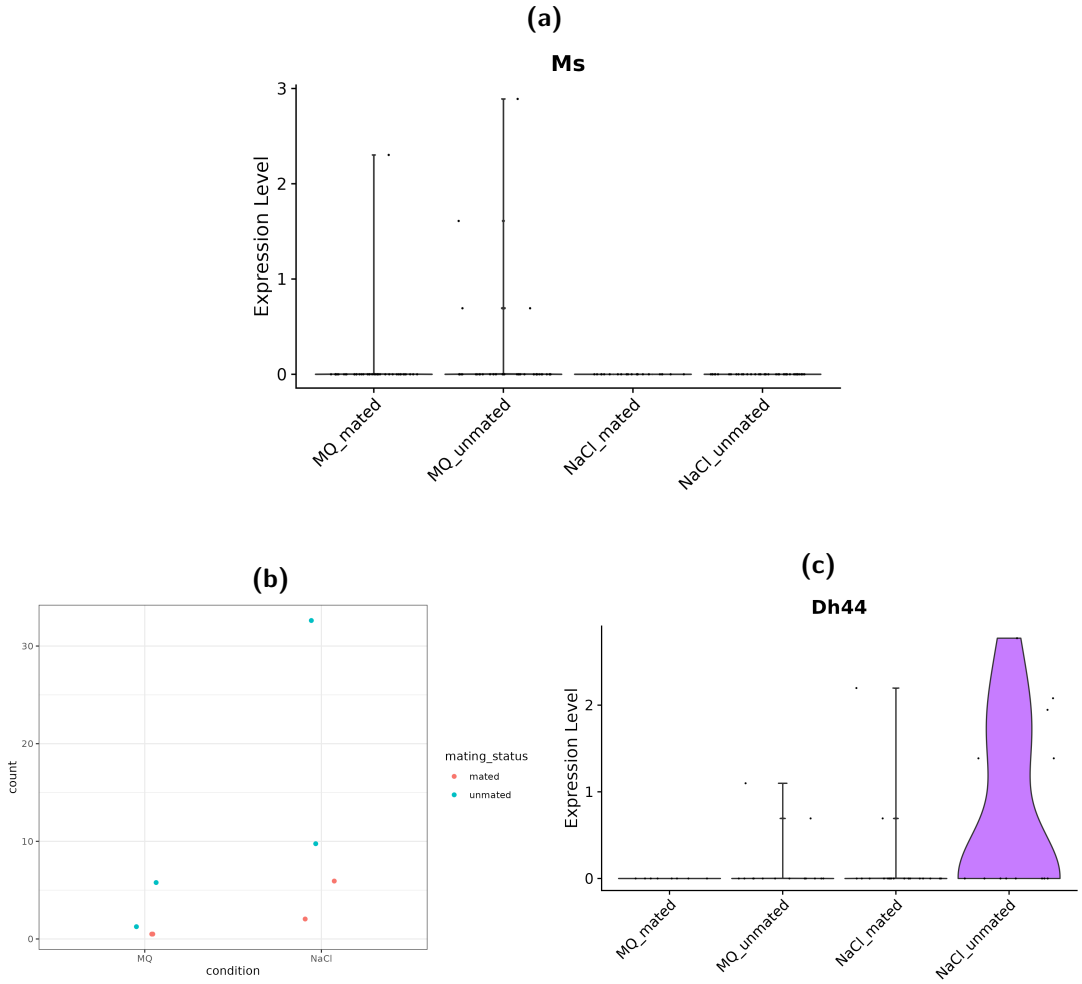


Figure A.6: Pseudobulk counts and SCTransformed counts for *Ms* and *Dh44* in clusters 22-ACh and 20-Glut respectfully from female *Drosophila*. a. SCTransformed counts of *Ms* in the cluster 22-ACh, split by condition and mating status. A few cells in the nutrient-deprived condition (MQ) had a greater expression of *Ms* compared to the NaCl-fed condition. b. Pseudobulk counts for *Dh44* in cluster 20-Glut. The value of the pseudobulk counts for *Dh44* were greater in unmated NaCl-fed compared to nutrient-deprived female *Drosophila*. c. SCTransformed counts for *Dh44* were greater in cluster 20-Glut in NaCl-fed female *Drosophila*, especially in unmated female *Drosophila*, compared to nutrient-deprived *Drosophila*.

A.3 Appendix 3

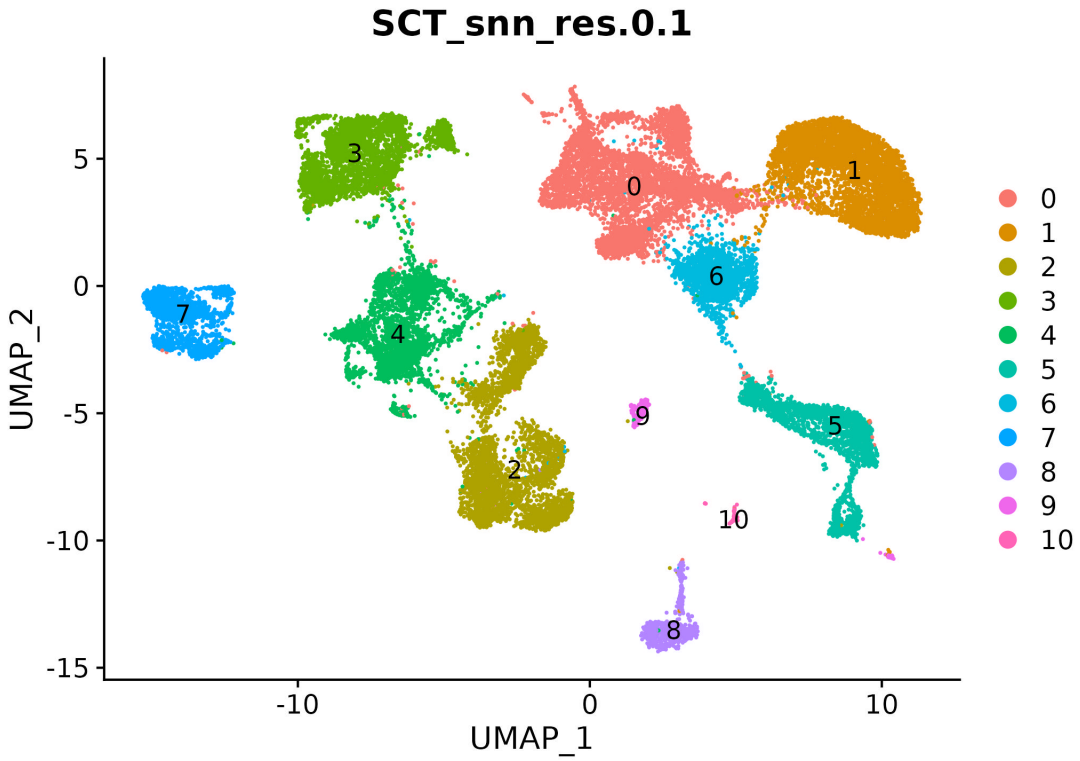


Figure A.7: Reclustering the glial cells at a clustering resolution of 0.1 created 11 clusters.

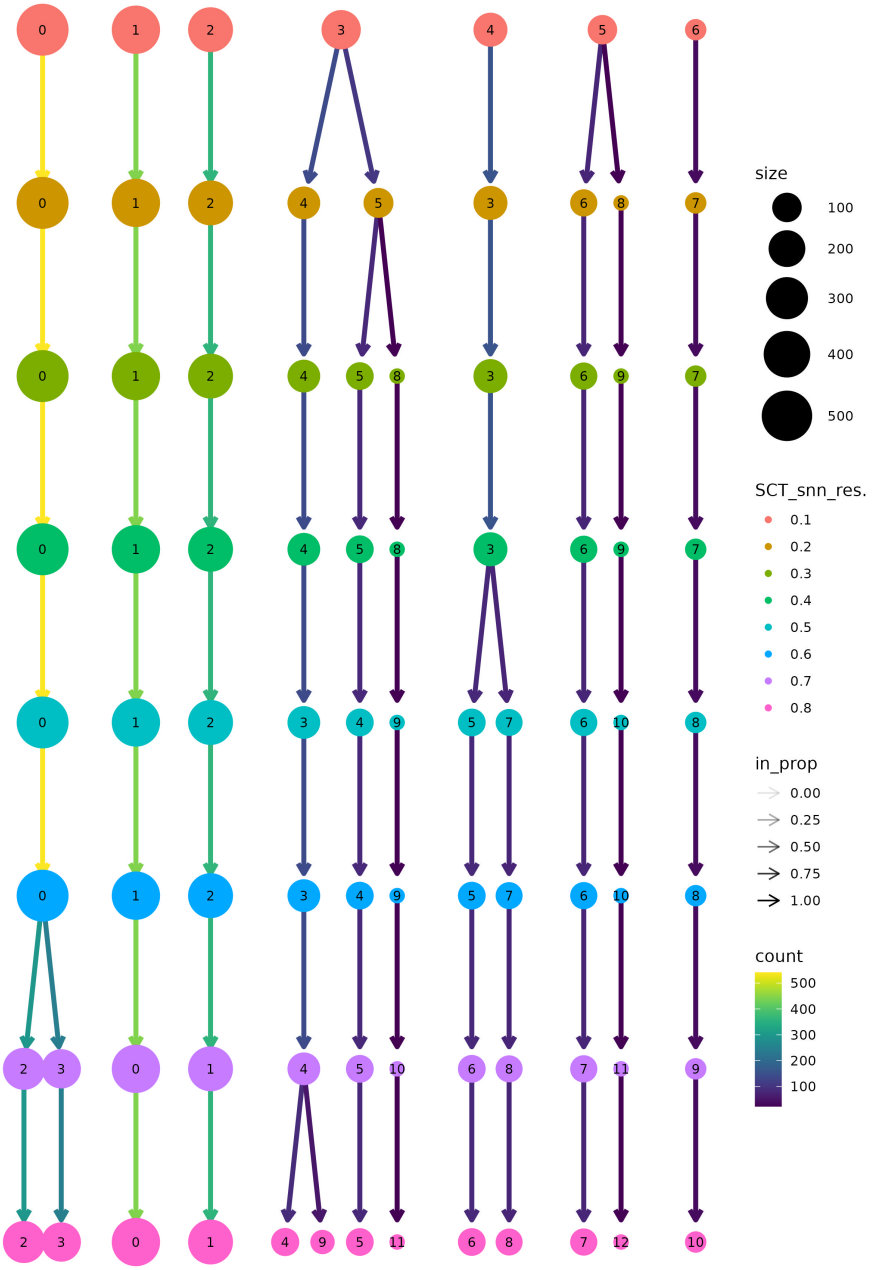


Figure A.8: Clustering resolution analysis for monoaminergic cells.
Cluster 3 at cluster resolution 0.1, which expressed *Drosophila* PAM and PPL cluster markers, splits into two clusters at the cluster resolution of 0.2.

References

Abdelaal, Tamim, Soufiane Mourragui, Ahmed Mahfouz, and Marcel J.T. Reinders (Oct. 2020). "SpaGE: Spatial Gene Enhancement using scRNA-seq". In: *Nucleic Acids Research* 48.18, e107–e107. URL: https://dx.doi.org/10.1093/nar/gkaa740.

Abel, Regina, Jrgen Rybak, and Randolf Menzel (Aug. 2001). "Structure and response patterns of olfactory interneurons in the honeybee, Apis mellifera". In: *Journal of Comparative Neurology* 437.3, pp. 363–383. URL: /doi/pdf/10.1002/cne.1289%20https://onlinelibrary.wiley.com/doi/abs/10.1002/cne.1289%20https://onlinelibrary.wiley.com/doi/10.1002/cne.1289.

Andrews, Simon (2010). "FastQC". In: *Babraham Bioinformatics*.

Aran, Dvir, Agnieszka P. Looney, Leqian Liu, Esther Wu, Valerie Fong, Austin Hsu, Suzanna Chak, Ram P. Naikawadi, Paul J. Wolters, Adam R. Abate, Atul J. Butte, and Mallar Bhattacharya (Feb. 2019). "Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage". In: *Nature immunology* 20.2, pp. 163–172. URL: https://pubmed.ncbi.nlm.nih.gov/30643263/.

Aso, Yoshinori, Daisuke Hattori, Yang Yu, Rebecca M Johnston, Nirmala A Iyer, Teri-Tb Ngo, Heather Dionne, L F Abbott, Richard Axel, Hiromu Tanimoto, and Gerald M Rubin (2014). "The neuronal architecture of the mushroom body provides a logic for associative learning". In: 3, p. 4577.

Aso, Yoshinori, Andrea Herb, Maite Ogueta, Igor Siwanowicz, Thomas Templier, Anja B. Friedrich, Kei Ito, Henrike Scholz, and Hiromu Tanimoto (July 2012). "Three Dopamine pathways induce aversive odor memories with different stability". In: *PLoS Genetics* 8.7. URL: https://pubmed.ncbi.nlm.nih.gov/22807684/.

Aso, Yoshinori, Robert P. Ray, Xi Long, Daniel Bushey, Karol Cichewicz, Teri Tb Ngo, Brandi Sharp, Christina Christoforou, Amy Hu, Andrew Lemire, Paul Tillberg, Jay Hirsh, Ashok Litwin-Kumar, and Gerald M. Rubin (Nov. 2019). "Nitric oxide acts as a cotransmitter in a subset of dopaminergic neurons to diversify memory dynamics". In: *eLife* 8.

Awasaki, Takeshi, Mai Saito, Masaki Sone, Emiko Suzuki, Ryoko Sakai, Kei Ito, and Chihiro Hama (2000). "The Drosophila Trio Plays an Essential Role in Patterning of Axons by Regulating Their Directional Extension". In: *Neuron* 26, pp. 119–131.

Bestea, Louise, Emmanuelle Briard, Julie Carcaud, Jean Christophe Sandoz, Rodrigo Velarde, Martin Giurfa, and Maria Gabriela De Brito Sanchez (2022). "The short neuropeptide F (sNPF) promotes the formation of appetitive visual memories in honey bees". In: *Biology Letters* 18.2. URL: /doi/pdf/10.1098/rsbl.2021.0520.

Bielopolski, Noa, Hoger Amin, Anthi A Apostolopoulou, Eyal Rozenfeld, Hadas Lerner, Wolf Huetteroth, Andrew C Lin, and Moshe Parnas (June 2019). "Inhibitory muscarinic acetylcholine receptors enhance aversive olfactory learning in adult Drosophila". In: *eLife* 8, e48264. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC6641838/.

Bitterman, M. E., R. Menzel, A. Fietz, and S. Schäfer (1983). "Classical conditioning of proboscis extension in honeybees (Apis mellifera)." In: *Journal of comparative psychology (Washington, D.C. : 1983)* 97.2, pp. 107–119.

Bornstein, Bavat, Hagar Meltzer, Ruth Adler, Idan Alyagor, Victoria Berkun, Gideon Cummings, Fabienne Reh, Hadas Keren-Shaul, Eyal David, Thomas Riemensperger, and Oren Schuldiner (June 2021). " Transneuronal Dpr12/DIP-δ interactions facilitate compartmentalized dopaminergic innervation of Drosophila

mushroom body axons ". In: *The EMBO Journal* 40.12, p. 105763. URL: /doi/pdf/10.15252/embj.2020105763?download=true.

Bou Dib, Peter, Bettina Gnägi, Fiona Daly, Virginie Sabado, Damla Tas, Dominique A. Glauser, Peter Meister, and Emi Nagoshi (Oct. 2014). "A conserved role for p48 homologs in protecting dopaminergic neurons from oxidative stress". In: *PLoS genetics* 10.10. URL: https://pubmed.ncbi.nlm.nih.gov/25340742/.

Burke, Christopher J., Wolf Huetteroth, David Owald, Emmanuel Perisse, Michael J. Krashes, Gaurav Das, Daryl Gohl, Marion Silies, Sarah Certel, and Scott Waddell (Dec. 2012). "Layered reward signalling through octopamine and dopamine in Drosophila". In: *Nature* 492.7429, pp. 433–437. URL: https://www.nature.com/articles/nature11614.

Burke, Christopher J. and Scott Waddell (May 2011). "Remembering nutrient quality of sugar in drosophila". In: *Current Biology* 21.9, pp. 746–750. URL: https://pubmed.ncbi.nlm.nih.gov/21514159/.

Ch'ng, Sarah S. and Andrew J. Lawrence (Aug. 2019). "The subfornical organ in sodium appetite: Recent insights". In: *Neuropharmacology* 154, pp. 107–113. URL: https://www.sciencedirect.com/science/article/abs/pii/S0028390818305069.

Claridge-Chang, Adam, Robert D. Roorda, Eleftheria Vrontou, Lucas Sjulson, Haiyan Li, Jay Hirsh, and Gero Miesenböck (Oct. 2009). "Writing Memories with Light-Addressable Reinforcement Circuitry". In: *Cell* 139.2, pp. 405–415. URL: https://pubmed.ncbi.nlm.nih.gov/19837039/.

Cohn, Raphael, Ianessa Morantte, and Vanessa Ruta (Dec. 2015). "Coordinated and Compartmentalized Neuromodulation Shapes Sensory Processing in Drosophila". In: *Cell* 163.7, pp. 1742–1755. URL: https://www.cell.com/action/showFullText?pii=S0092867415014993%20https://www.cell.com/action/showAbstract?pii=S0092867415014993%20https://www.cell.com/cell/abstract/S0092-8674(15)01499-3.

Cone, Jackson J., Samantha M. Fortin, Jenna A. McHenry, Garret D. Stuber, James E. McCutcheon, and Mitchell F. Roitman (Feb. 2016). "Physiological state gates acquisition and expression of mesolimbic reward prediction signals". In: *Proceedings of the National Academy of Sciences of the United States of America* 113.7, pp. 1943–1948. URL: https://www.pnas.org/content/113/7/1943%20https://www.pnas.org/content/113/7/1943.abstract.

Corrales-Carvajal, Verónica María, Aldo A Faisal, and Carlos Ribeiro (Oct. 2016). "Internal states drive nutrient homeostasis by modulating exploration-exploitation trade-off". In: *eLife* 5.

Crittenden, Jill R, Efthimios M C Skoulakis, Kyung-An Han, Daniel Kalderon, and Ronald L Davis (1998). "Tripartite Mushroom Body Architecture Revealed by Antigenic Markers". In.

Croset, Vincent, Christoph D. Treiber, and Scott Waddell (Apr. 2018). "Cellular diversity in the Drosophila midbrain revealed by single-cell transcriptomics". In: *eLife* 7.

Das, Gaurav, Martín Klappenbach, Eleftheria Vrontou, Emmanuel Perisse, Christopher M. Clark, Christopher J. Burke, and Scott Waddell (Aug. 2014). "Drosophila learn opposing components of a compound food stimulus". In: *Current Biology* 24.15, pp. 1723–1730. URL: /pmc/articles/PMC4131107/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4131107/.

De Sousa, Raquel T., Robyn Darnell, and Geraldine A. Wright (2022). "Behavioural regulation of mineral salt intake in honeybees: a self-selection approach". In: *Philosophical*

Transactions of the Royal Society B: Biological Sciences 377.1853, p. 20210169. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC9058550/.

Dus, Monica, Jason Sih Yu Lai, Keith M. Gunapala, Soohong Min, Timothy D. Tayler, Anne C. Hergarden, Eliot Geraud, Christina M. Joseph, and Greg S.B. Suh (July 2015). "Nutrient Sensor in the Brain Directs the Action of the Brain-Gut Axis in Drosophila". In: *Neuron* 87.1, pp. 139–151. URL: https://pubmed.ncbi.nlm.nih.gov/26074004/%20https://pubmed.ncbi.nlm.nih.gov/26074004/?dopt=Abstract.

Dus, Monica, Soo Hong Min, Alex C. Keene, Ga Young Lee, and Greg S.B. Suh (July 2011). "Taste-independent detection of the caloric content of sugar in Drosophila". In: *Proceedings of the National Academy of Sciences of the United States of America* 108.28, pp. 11644–11649. URL: https://pubmed.ncbi.nlm.nih.gov/21709242/.

Ehmer, Birgit and Wulfila Gronenberg (Sept. 2002). "Segregation of visual input to the mushroom bodies in the honeybee (Apis mellifera)". In: *Journal of Comparative Neurology* 451.4, pp. 362–373. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cne.10355?casa_token=Z67wn0iT0acAAAAA:9iy9QBiwzgemrolsSMM32pillpI1gUEB84sXSahivgc7ra_RMezb2D7yfewCPvUHNjfOPLO3jjN1WVsi%20https://onlinelibrary.wiley.com/doi/abs/10.1002/cne.10355%20https://onlinelibrary.wiley.com/doi/10.1002/cne.10355.

Eisenhardt, Dorothea (Apr. 2006). "Learning and memory formation in the honeybee (Apis mellifera) and its dependency on the cAMP-protein kinase A pathway". In: *Animal Biology* 56.2, pp. 259–278.

Erber, J., TH Masuhr, and R. Menzel (1980). "Localization of short-term memory in the brain of the bee, Apis mellifera". In: *Physiological Entomology* 5.4, pp. 343–358.

Ewels, Philip, Måns Magnusson, Sverker Lundin, and Max Käller (Oct. 2016). "MultiQC: summarize analysis results for multiple tools and samples in a single report". In: *Bioinformatics* 32.19, pp. 3047–3048. URL: https://dx.doi.org/10.1093/bioinformatics/btw354.

Fahrbach, Susan E. (Jan. 2006). "Structure of the mushroom bodies of the insect brain". In: *Annual Review of Entomology* 51.Volume 51, 2006, pp. 209–232. URL: https://www.annualreviews.org/content/journals/10.1146/annurev.ento.51.110104.150954.

Farris, Sarah M., Andrew I. Abrams, and Nicholas J. Strausfeld (June 2004). "Development and morphology of Class II Kenyon cells in the mushroom bodies of the honey bee, Apis mellifera". In: *Journal of Comparative Neurology* 474.3, pp. 325–339. URL: https://pubmed.ncbi.nlm.nih.gov/15174077/.

Felsenberg, Johannes (Apr. 2021). "Changing memories on the fly: the neural circuits of memory re-evaluation in Drosophila melanogaster". In: *Current Opinion in Neurobiology* 67, pp. 190–198. URL: https://www.sciencedirect.com/science/article/pii/S0959438820301884.

Feng, Kai, Mark T. Palfreyman, Martin Häsemeyer, Aaron Talsma, and Barry J. Dickson (July 2014). "Ascending SAG Neurons Control Sexual Receptivity of Drosophila Females". In: *Neuron* 83.1, pp. 135–148. URL: http://www.cell.com/article/S0896627314004036/fulltext%20http://www.cell.com/article/S0896627314004036/abstract%20https://www.cell.com/neuron/abstract/S0896-6273(14)00403-6.

Fushima, Kazuma and Hidenobu Tsujimura (Apr. 2007). "Precise control of fasciclin II expression is required for adult mushroom body development in Drosophila". In: *Development Growth and Differentiation* 49.3, pp. 215–227. URL: https://pubmed.ncbi.nlm.nih.gov/17394600/.

Ganguly, Ishani, Emily L. Heckman, Ashok Litwin-Kumar, E. Josephine Clowney, and Rudy Behnia (July 2024). "Diversity of visual inputs to Kenyon cells of the Drosophila mushroom body". In: *Nature Communications 2024 15:1* 15.1, pp. 1–18. URL: https://www.nature.com/articles/s41467-024-49616-z.

Gauthier, Monique, Valérie Cano-Lozano, Asma Zaoujal, and Daniel Richard (Aug. 1994). "Effects of intracranial injections of scopolamine on olfactory conditioning retrieval in the honeybee". In: *Behavioural Brain Research* 63.2, pp. 145–149. URL: https://www.sciencedirect.com/science/article/pii/016643289490085X.

Gronenberg, Wulfila (July 2001). "Subdivisions of hymenopteran mushroom body calyces by their afferent supply". In: *Journal of Comparative Neurology* 435.4, pp. 474–489. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cne.1045?casa_token=o7xWNyab874AAAAA:3ssLB5kmfzrDLBl6zzoPgO28mnwJA9P0-agDrT-IlyT-XR_ieQlc21dwFZcwKkyqjE4PUS6iUDIMkIC-%20https://onlinelibrary.wiley.com/doi/abs/10.1002/cne.1045%20https://onlinelibrary.wiley.com/doi/10.1002/cne.1045.

Grüter, Christoph and Francis L.W. Ratnieks (May 2011). "Honeybee foragers increase the use of waggle dance information when private information becomes unrewarding". In: *Animal Behaviour* 81.5, pp. 949–954. URL: https://www.sciencedirect.com/science/article/abs/pii/S0003347211000170.

Hammer, Martin (1993). "An identified neuron mediates the unconditioned stimulus in associative olfactory learning in honeybees". In: *Nature* 366.6450, pp. 59–63. URL: https://www.nature.com/articles/366059a0.

Hammer, Martin and Randolf Menzel (1995). "Learning and Memory in the Honeybee". In: *The Journal of Neuroscience* 15.3.

— (1998). "Multiple Sites of Associative Odor Learning as Revealed by Local Brain Microinjections of Octopamine in Honeybees". In: *Learning & Memory* 5.1, p. 146. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC311245/.

Handler, Annie, Thomas G.W. Graham, Raphael Cohn, Ianessa Morantte, Andrew F. Siliciano, Jianzhi Zeng, Yulong Li, and Vanessa Ruta (June 2019). "Distinct Dopamine Receptor Pathways Underlie the Temporal Sensitivity of Associative Learning". In: *Cell* 178.1, pp. 60–75. URL: https://www.cell.com/action/showFullText?pii=S0092867419306117%20https://www.cell.com/action/showAbstract?pii=S0092867419306117%20https://www.cell.com/cell/abstract/S0092-8674(19)30611-7.

Hector, Clare E., Colin A. Bretz, Yan Zhao, and Erik C. Johnson (Oct. 2009). "Functional differences between two CRF-related diuretic hormone receptors in Drosophila". In: *Journal of Experimental Biology* 212.19, pp. 3142–3147. URL: https://pubmed.ncbi.nlm.nih.gov/19749107/.

Heisenberg, Martin (1998). "What Do the Mushroom Bodies Do for the Insect Brain? An Introduction". In: *Learning & Memory* 5.1, p. 1. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC311238/.

Heisenberg, Martin, Alexander Borst, Sibylle Wagner, and Duncan Byers (1985). "drosophila mushroom body mutants are deficient in olfactory learning: Research papers". In: *Journal of Neurogenetics* 2.1, pp. 1–30. URL: https://pubmed.ncbi.nlm.nih.gov/4020527/.

Hige, Toshihide, Yoshinori Aso, Gerald M. Rubin, and Glenn C. Turner (Oct. 2015). "Plasticity-driven individualization of olfactory coding in mushroom body output neurons". In: *Nature* 526.7572, pp. 258–262. URL: https://pubmed.ncbi.nlm.nih.gov/26416731/.

Hiroi, Makoto, Nicolas Meunier, Frédéric Marion-Poll, and Teiichi Tanimura (Dec. 2004). "Two antagonistic gustatory receptor neurons responding to sweet-salty and bitter taste in Drosophila". In: *Journal of Neurobiology* 61.3, pp. 333–342. URL: https://pubmed.ncbi.nlm.nih.gov/15389687/.

Hiyama, Takeshi Y., Eiji Watanabe, Haruo Okado, and Masaharu Noda (Oct. 2004). "The Subfornical Organ is the Primary Locus of Sodium-Level Sensing by Nax Sodium Channels for the Control of Salt-Intake Behavior". In: *Journal of Neuroscience* 24.42, pp. 9276–9281. URL: https://www.jneurosci.org/content/24/42/9276%20https://www.jneurosci.org/content/24/42/9276.abstract.

Huetteroth, Wolf, Emmanuel Perisse, Suewei Lin, Martín Klappenbach, Christopher Burke, and Scott Waddell (Mar. 2015). "Sweet taste and nutrient value subdivide rewarding dopaminergic neurons in drosophila". In: *Current Biology* 25.6, pp. 751–758.

Jaeger, Alexandria H., Molly Stanley, Zachary F. Weiss, Pierre Yves Musso, Rachel C.W. Chan, Han Zhang, Damian Feldman-Kiss, and Michael D. Gordon (Oct. 2018). "A complex peripheral code for salt taste in drosophila". In: *eLife* 7.

Johard, Helena A.D., Lina E. Enell, Elisabeth Gustafsson, Pierre Trifilieff, Jan A. Veenstra, and Dick R. Nässel (Apr. 2008). "Intrinsic neurons of Drosophila mushroom bodies express short neuropeptide F: Relations to extrinsic neurons expressing different neurotransmitters". In: *Journal of Comparative Neurology* 507.4, pp. 1479–1496. URL: https://pubmed.ncbi.nlm.nih.gov/18205208/.

Jovic, Dragomirka, Xue Liang, Hua Zeng, Lin Lin, Fengping Xu, and Yonglun Luo (Mar. 2022). "Single-cell RNA sequencing technologies and applications: A brief overview". In: *Clinical and Translational Medicine* 12.3, e694. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC8964935/.

Kaneko, Kumi, Tsubomi Ikeda, Mirai Nagai, Sayaka Hori, Chie Umatani, Hiroto Tadano, Atsushi Ugajin, Takayoshi Nakaoka, Rajib Kumar Paul, Tomoko Fujiyuki, Kenichi Shirai, Takekazu Kunieda, Hideaki Takeuchi, and Takeo Kubo (Aug. 2013). "Novel Middle-Type Kenyon Cells in the Honeybee Brain Revealed by Area-Preferential Gene Expression Analysis". In: *PLOS ONE* 8.8, e71732. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0071732.

Kelley, Richard L. and Mitzi I. Kuroda (June 2003). "The Drosophila roX1 RNA gene can overcome silent chromatin by recruiting the male-specific lethal dosage compensation complex". In: *Genetics* 164.2, pp. 565–574. URL: https://pubmed.ncbi.nlm.nih.gov/12807777/.

Kendroud, Sarah, Ali Asgar Bohra, Philipp A. Kuert, Bao Nguyen, Oriane Guillermin, Simon G. Sprecher, Heinrich Reichert, Krishnaswamy VijayRaghavan, and Volker Hartenstein (Jan. 2018). "Structure and Development of the Subesophageal Zone of the Drosophila Brain. II. Sensory Compartments". In: *The Journal of comparative neurology* 526.1, p. 33. URL: /pmc/articles/PMC5971197/%20/pmc/articles/PMC5971197/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5971197/.

Kim, Byoungsoo, Gayoung Hwang, Sung Eun Yoon, Meihua Christina Kuang, Jing W. Wang, Young Joon Kim, and Greg S.B. Suh (May 2024). "Postprandial sodium sensing by

enteric neurons in Drosophila". In: *Nature Metabolism* 6.5, pp. 837–846. URL: https://pubmed.ncbi.nlm.nih.gov/38570627/.

Kim, Young Cho, Hyun Gwan Lee, Junghwa Lim, and Kyung An Han (Jan. 2013). "Appetitive learning requires the alpha1-like octopamine receptor oamb in the drosophila mushroom body neurons". In: *Journal of Neuroscience* 33.4, pp. 1672–1677. URL: https://pubmed.ncbi.nlm.nih.gov/23345239/.

Knapek, Stephan, Lily Kahsai, Åsa M.E. Winther, Hiromu Tanimoto, and Dick R. Nässel (2013). "Short neuropeptide F acts as a functional neuromodulator for olfactory memory in kenyon cells of Drosophila mushroom bodies". In: *Annals of Internal Medicine* 158.6, pp. 5340–5345. URL: https://pubmed.ncbi.nlm.nih.gov/23516298/.

Korsunsky, Ilya, Nghia Millard, Jean Fan, Kamil Slowikowski, Fan Zhang, Kevin Wei, Yuriy Baglaenko, Michael Brenner, Po ru Loh, and Soumya Raychaudhuri (Nov. 2019). "Fast, sensitive and accurate integration of single-cell data with Harmony". In: *Nature Methods 2019 16:12* 16.12, pp. 1289–1296. URL: https://www.nature.com/articles/s41592-019-0619-0.

Krashes, Michael J., Shamik DasGupta, Andrew Vreede, Benjamin White, J. Douglas Armstrong, and Scott Waddell (Oct. 2009). "A Neural Circuit Mechanism Integrating Motivational State with Memory Expression in Drosophila". In: *Cell* 139.2, pp. 416–427. URL: https://pubmed.ncbi.nlm.nih.gov/19837040/.

Krashes, Michael J. and Scott Waddell (Mar. 2008). "Rapid consolidation to a radish and protein synthesis-dependent long-term memory after single-session appetitive olfactory conditioning in Drosophila". In: *Journal of Neuroscience* 28.12, pp. 3103–3113. URL: /pmc/articles/PMC2516741/%20/pmc/articles/PMC2516741/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2516741/.

— (May 2011a). "Drosophila appetitive olfactory conditioning". In: *Cold Spring Harbor Protocols* 6.5. URL: https://pubmed.ncbi.nlm.nih.gov/21536767/.

— (May 2011b). "Drosophila aversive olfactory conditioning". In: *Cold Spring Harbor Protocols* 6.5. URL: https://pubmed.ncbi.nlm.nih.gov/21536768/.

Kubo, T and Takeo Kubo (2012). "Neuroanatomical Dissection of the Honey Bee Brain Based on Temporal and Regional Gene Expression Patterns". In: *Honeybee Neurobiology and Behavior*, pp. 341–357. URL: https://link.springer.com/chapter/10.1007/978-94-007-2099-2_26.

Lee, Gahbien, Heejin Jang, and Yangkyun Oh (2023). "The role of diuretic hormones (DHs) and their receptors in Drosophila". In: *BMB Reports* 56.4, pp. 209–215. URL: https://pubmed.ncbi.nlm.nih.gov/36977606/.

Lee, Tzumin, Arthur Lee, and Liqun Luo (Sept. 1999). "Development of the Drosophila mushroom bodies: sequential generation of three distinct types of neurons from a neuroblast". In: *Development* 126.18, pp. 4065–4076. URL: https://dx.doi.org/10.1242/dev.126.18.4065.

Li, Feng, Jack Lindsey, Elizabeth C. Marin, Nils Otto, Marisa Dreher, Georgia Dempsey, Ildiko Stark, Alexander Shakeel Bates, Markus William Pleijzier, Philipp Schlegel, Aljoscha Nern, Shinya Takemura, Nils Eckstein, Tansy Yang, Audrey Francis, Amalia Braun, Ruchi Parekh, Marta Costa, Louis Scheffer, Yoshinori Aso, Gregory S.X.E. Jefferis, L. F. Abbott, Ashok Litwin-Kumar, Scott Waddell, and Gerald M. Rubin (Dec. 2020a). "The connectome of the adult drosophila mushroom body provides insights into function". In: *eLife* 9, pp. 1–217.

— (Dec. 2020b). "The connectome of the adult drosophila mushroom body provides insights into function". In: *eLife* 9, pp. 1–217.

Li, Qiye, Mingyue Wang, Pei Zhang, Yang Liu, Qunfei Guo, Yuanzhen Zhu, Tinggang Wen, Xueqin Dai, Xiafang Zhang, Manuel Nagel, Bjarke Hamberg Dethlefsen, Nianxia Xie, Jie Zhao, Wei Jiang, Lei Han, Liang Wu, Wenjiang Zhong, Zhifeng Wang, Xiaoyu Wei, Wei Dai, Longqi Liu, Xun Xu, Haorong Lu, Huanming Yang, Jian Wang, Jacobus J. Boomsma, Chuanyu Liu, Guojie Zhang, and Weiwei Liu (Aug. 2022). "A single-cell transcriptomic atlas tracking the neural basis of division of labour in an ant superorganism". In: *Nature Ecology and Evolution* 6.8, pp. 1191–1204. URL: https://pubmed.ncbi.nlm.nih.gov/35711063/.

Lin, Suewei (Jan. 2023). "The making of the Drosophila mushroom body". In: *Frontiers in Physiology* 14, p. 1091248.

Lin, Suewei, David Owald, Vikram Chandra, Clifford Talbot, Wolf Huetteroth, and Scott Waddell (Oct. 2014). "Neural correlates of water reward in thirsty Drosophila". In: *Nature Neuroscience* 17.11, pp. 1536–1542. URL: https://pubmed.ncbi.nlm.nih.gov/25262493/.

Liu, Chang, Pierre Yves Plaaais, Nobuhiro Yamagata, Barret D. Pfeiffer, Yoshinori Aso, Anja B. Friedrich, Igor Siwanowicz, Gerald M. Rubin, Thomas Preat, and Hiromu Tanimoto (Aug. 2012). "A subset of dopamine neurons signals reward for odour memory in Drosophila". In: *Nature* 488.7412, pp. 512–516. URL: https://www.nature.com/articles/nature11304.

Liu, Ling Yu, Xi Long, Ching Po Yang, Rosa L. Miyares, Ken Sugino, Robert H. Singer, and Tzumin Lee (Sept. 2019). "Mamo decodes hierarchical temporal gradients into terminal neuronal fate". In: *eLife* 8.

Love, Michael I., Wolfgang Huber, and Simon Anders (Dec. 2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2". In: *Genome Biology* 15.12, pp. 1–21. URL: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8.

Maier, Dieter, Anna X. Chen, Anette Preiss, and Manuela Ketelhut (2008). "The tiny Hairless protein from Apis mellifera: a potent antagonist of Notch signaling in Drosophila melanogaster". In: *BMC Evolutionary Biology* 8.1, p. 175. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC2440387/.

Malun, Dagmar (1998). "Early Development of Mushroom Bodies in the Brain of the Honeybee Apis mellifera as Revealed by BrdU Incorporation and Ablation Experiments". In: *Learning & Memory* 5.1, p. 90. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC311246/.

Mao, Zhengmei and Ronald L. Davis (July 2009). "Eight different types of dopaminergic neurons innervate the Drosophila mushroom body neuropil: Anatomical and physiological heterogeneity". In: *Frontiers in Neural Circuits* 3.JUL. URL: https://pubmed.ncbi.nlm.nih.gov/19597562/.

Matsumura, Yasuhiro, Taiko Kim To, Takekazu Kunieda, Hiroki Kohno, Tetsuji Kakutani, and Takeo Kubo (Dec. 2022). "Mblk-1/E93, an ecdysone related-transcription factor, targets synaptic plasticity-related genes in the honey bee mushroom bodies". In: *Scientific Reports* 12.1. URL: https://pubmed.ncbi.nlm.nih.gov/36494426/.

McGinnis, Christopher S., Lyndsay M. Murrow, and Zev J. Gartner (Apr. 2019). "DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors". In: *Cell Systems* 8.4, pp. 329–337. URL: https://pubmed.ncbi.nlm.nih.gov/30954475/.

McKim, Theresa H, Jayati Gera, Ariana J Gayban, Nils Reinhard, Giulia Manoli, Selina Hilpert, Charlotte Helfrich-Förster, and Meet Zandawala (Nov. 2024). "Synaptic

connectome of a neurosecretory network in the Drosophila brain". In: *eLife* 13. URL: https://elifesciences.org/reviewed-preprints/102684.

Miyamoto, Tetsuya and Hubert Amrein (2013). "Diverse roles for the Drosophila fructose sensor Gr43a". In: *Fly* 8.1, p. 19. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC3974889/.

Miyamoto, Tetsuya, Jesse Slone, Xiangyu Song, and Hubert Amrein (Nov. 2012). "A fructose receptor functions as a nutrient sensor in the drosophila brain". In: *Cell* 151.5, pp. 1113–1125. URL: https://pubmed.ncbi.nlm.nih.gov/23178127/.

Mizunami, Makoto and Yukihisa Matsumoto (Nov. 2010). "Roles of aminergic neurons in formation and recall of associative memory in crickets". In: *Frontiers in Behavioral Neuroscience* 4.NOV, p. 2163. URL: www.frontiersin.org.

Mobbs, P. G., Mobbs, and P. G. (July 1982). "The Brain of the Honeybee Apis Mellifera. I. The Connections and Spatial Organization of the Mushroom Bodies". In: *RSPTB* 298.1091, pp. 309–354. URL: https://ui.adsabs.harvard.edu/abs/1982RSPTB.298..309M/abstract.

Mohammad, Farhan, James C Stewart, Stanislav Ott, Katarina Chlebikova, Jia Yi Chua, Tong-Wey Koh, Joses Ho, and Adam Claridge-Chang (Jan. 2017). "Optogenetic inhibition of behavior with anion channelrhodopsins". In: *Nature Methods 2017 14:3* 14.3, pp. 271–274. URL: https://www.nature.com/articles/nmeth.4148.

Mu, Xiaohuan, Zijing Zhang, Qun Liu, Jie Ma, Yating Qin, Haoyu Lang, Yingying Zhang, Nannan Zhang, Qunfei Guo, Pei Zhang, Denghui Li, Ruihua Zhang, Qianyue Ji, Aijun Jiang, Yang Wang, Shanshan Pan, Xiawei Liu, Xuemei Liu, Jiahui Sun, Yan Liu, Hao Chen, Li Zheng, Liang Meng, Haorong Lu, He Zhang, Yifan Zhai, Qiye Li, Junnian Liu, Huanming Yang, Jian Wang, Xiaosong Hu, Xun Xu, Shanshan Liu, and Hao Zheng (Apr. 2025). "Single-nucleus and spatial transcriptomics identify brain landscape of gene regulatory networks associated with behavioral maturation in honeybees". In: *Nature Communications 2025 16:1* 16.1, pp. 1–16. URL: https://www.nature.com/articles/s41467-025-58614-8.

Musso, Pierre-Yves, Paul Tchenio, and Thomas Preat (Feb. 2015). "Delayed Dopamine Signaling of Energy Level Builds Appetitive Long-Term Memory in Drosophila". In: *Cell Reports* 10.7, pp. 1023–1031. URL: http://www.cell.com/article/S2211124715000613/fulltext%20http://www.cell.com/article/S2211124715000613/abstract%20https://www.cell.com/cell-reports/abstract/S2211-1247(15)00061-3.

Niewalda, Thomas, Nidhi Singhal, André Fiala, Timo Saumweber, Stephanie Wegener, and Bertram Gerber (Oct. 2008). "Salt Processing in Larval Drosophila: Choice, Feeding, and Learning Shift from Appetitive to Aversive in a Concentration-Dependent Way". In: *Chemical Senses* 33.8, pp. 685–692. URL: https://dx.doi.org/10.1093/chemse/bjn037.

O'Donnell, Sean, Nicole A. Donlan, and Theresa A. Jones (Feb. 2004). "Mushroom body structural change is associated with division of labor in eusocial wasp workers (Polybia aequatorialis, Hymenoptera: Vespidae)". In: *Neuroscience Letters* 356.3, pp. 159–162. URL: https://pubmed.ncbi.nlm.nih.gov/15036619/.

Okada, Ryuichi, Jürgen Rybak, Gisela Manz, and Randolf Menzel (Oct. 2007). "Learning-related plasticity in PE1 and other mushroom body-extrinsic neurons in the honeybee brain". In: *Journal of Neuroscience* 27.43, pp. 11736–11747. URL: https://pubmed.ncbi.nlm.nih.gov/17959815/.

Otto, Nils, Markus W Pleijzier, Isabel C Morgan, Gerald M Rubin, Gregory S X E Jefferis, and Scott Waddell Correspondence (2020). "Input Connectivity Reveals Additional Heterogeneity of Dopaminergic Reinforcement in Drosophila". In: URL: https://doi.org/10.1016/j.cub.2020.05.077.

Owald, David, Johannes Felsenberg, Clifford B. Talbot, Gaurav Das, Emmanuel Perisse, Wolf Huetteroth, and Scott Waddell (Apr. 2015a). "Activity of defined mushroom body output neurons underlies learned olfactory behavior in Drosophila". In: *Neuron* 86.2, pp. 417–427. URL: https://pubmed.ncbi.nlm.nih.gov/25864636/.

Owald, David, Suewei Lin, and Scott Waddell (Sept. 2015b). "Light, heat, action: Neural control of fruit fly behaviour". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 370.1677. URL: https://pubmed.ncbi.nlm.nih.gov/26240426/.

Park, Annie, Vincent Croset, Nils Otto, Eleonora Meschi, David Sims, and Scott Waddell Correspondence (2022). "Gliotransmission of D-serine promotes thirst-directed behaviors in Drosophila". In: URL: https://doi.org/10.1016/j.cub.2022.07.038.

Paul, R. K., H. Takeuchi, Y. Matsuo, and T. Kubo (Jan. 2005). "Gene expression of ecdysteroid-regulated gene E74 of the honeybee in ovary and brain". In: *Insect Molecular Biology* 14.1, pp. 9–15. URL: https://pubmed.ncbi.nlm.nih.gov/15663771/.

Perisse, Emmanuel, David Owald, Oliver Barnstedt, Clifford B B. Talbot, Wolf Huetteroth, and Scott Waddell (June 2016). "Aversive Learning and Appetitive Motivation Toggle Feed-Forward Inhibition in the Drosophila Mushroom Body". In: *Neuron* 90.5, pp. 1086–1099. URL: https://pubmed.ncbi.nlm.nih.gov/27210550/.

Plaçais, Pierre Yves, Séverine Trannoy, Guillaume Isabel, Yoshinori Aso, Igor Siwanowicz, Ghislain Belliart-Guérin, Philippe Vernier, Serge Birman, Hiromu Tanimoto, and Thomas Preat (2012). "Slow oscillations in two pairs of dopaminergic neurons gate long-term memory formation in Drosophila". In: *Nature Neuroscience* 15.4.

Quinn, W. G., W. A. Harris, and S. Benzer (1974). "Conditioned Behavior in Drosophila melanogaster". In: *Proceedings of the National Academy of Sciences of the United States of America* 71.3, p. 708. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC388082/.

Ramya, Ravindrakumar, Chikkate Ramakrishnappa Venkatesh, and Baragur Venkatanarayanasetty Shyamala (Dec. 2024). "olf413 an octopamine biogenesis pathway gene is required for axon growth and pathfinding during embryonic nervous system development in Drosophila melanogaster". In: *BMC Research Notes* 17.1, pp. 1–10. URL: https://bmcresnotes.biomedcentral.com/articles/10.1186/s13104-024-06700-3.

Ribeiro, Carlos and Barry J. Dickson (June 2010). "Sex peptide receptor and neuronal TOR/S6K signaling modulate nutrient balancing in Drosophila". In: *Current Biology* 20.11, pp. 1000–1005. URL: https://pubmed.ncbi.nlm.nih.gov/20471268/.

Robinson, Mike J.F. and Kent C. Berridge (Feb. 2013). "Instant transformation of learned repulsion into motivational "wanting"". In: *Current Biology* 23.4, pp. 282–289. URL: https://pubmed.ncbi.nlm.nih.gov/23375893/.

Russell, Cheryl, Jan Wessnitzer, Joanna M. Young, J. Douglas Armstrong, and Barbara Webb (2011). "Dietary salt levels affect salt preference and learning in larval drosophila". In: *PLoS ONE* 6.6, p. 20100. URL: www.plosone.org.

Sabandal, John Martin, Paul Rafael Sabandal, Young Cho Kim, and Kyung An Han (May 2020). "Concerted Actions of Octopamine and Dopamine Receptors Drive Olfactory Learning". In: *Journal of Neuroscience* 40.21, pp. 4240–4250. URL:

https://www.jneurosci.org/content/40/21/4240%20https://www.jneurosci.org/content/40/21/4240.abstract.

Schäfer, Sabine and Vincent Rehder (1989). "Dopamine-like immunoreactivity in the brain and suboesophageal ganglion of the honeybee". In: *Journal of Comparative Neurology* 280.1, pp. 43–58.

Schatton, Adriana and Constance Scharff (Nov. 2017). "FoxP expression identifies a Kenyon cell subtype in the honeybee mushroom bodies linking them to fruit fly $\alpha\beta$c neurons". In: *European Journal of Neuroscience* 46.9, pp. 2534–2541. URL: https://pubmed.ncbi.nlm.nih.gov/28921711/.

Scheffer, Louis K. et al. (Sept. 2020). "A connectome and analysis of the adult drosophila central brain". In: *eLife* 9, pp. 1–74.

Schürmann, F. W., K. Elekes, and M. Geffard (May 1989). "Dopamine-like immunoreactivity in the bee brain". In: *Cell and Tissue Research* 256.2, pp. 399–410. URL: https://link.springer.com/article/10.1007/BF00218898.

Senapati, Bhagyashree, Chang Hui Tsao, Yi An Juan, Tai Hsiang Chiu, Chia Lin Wu, Scott Waddell, and Suewei Lin (Dec. 2019). "A neural mechanism for deprivation state-specific expression of relevant memories in Drosophila". In: *Nature Neuroscience* 22.12, pp. 2029–2039. URL: https://doi.org/10.1038/s41593-019-0515-z.

Shih, Meng Fu Maxwell, Fred Pejman Davis, Gilbert Lee Henry, and Josh Dubnau (Jan. 2018). "Nuclear Transcriptomes of the Seven Neuronal Cell Types That Constitute the Drosophila Mushroom Bodies". In: *G3: Genes|Genomes|Genetics* 9.1, p. 81. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC6325895/.

Simões, Patrício, Swidbert R. Ott, and Jeremy E. Niven (Aug. 2011). "Associative olfactory learning in the desert locust, Schistocerca gregaria". In: *The Journal of experimental biology* 214.Pt 15, pp. 2495–2503. URL: https://pubmed.ncbi.nlm.nih.gov/21753041/.

Sinakevitch, Irina T., Adrian N. Smith, Fernando Locatelli, Ramon Huerta, Maxim Bazhenov, and Brian H. Smith (Oct. 2013). "Apis mellifera octopamine receptor 1 (AmOA1) expression in antennal lobe networks of the honey bee (Apis mellifera) and fruit fly (Drosophila melanogaster)". In: *Frontiers in Systems Neuroscience* 7.OCT. URL: https://pubmed.ncbi.nlm.nih.gov/24187534/.

Squair, Jordan W., Matthieu Gautier, Claudia Kathe, Mark A. Anderson, Nicholas D. James, Thomas H. Hutson, Rémi Hudelle, Taha Qaiser, Kaya J.E. Matson, Quentin Barraud, Ariel J. Levine, Gioele La Manno, Michael A. Skinnider, and Grégoire Courtine (Dec. 2021). "Confronting false discoveries in single-cell differential expression". In: *Nature Communications* 12.1, p. 5692. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC8479118/.

Stabler, Daniel, Mushtaq Al-Esawy, Jennifer A. Chennells, Giorgia Perri, Alexandria Robinson, and Geraldine A. Wright (Feb. 2021). "Regulation of dietary intake of protein and lipid by nurse-age adult worker honeybees". In: *The Journal of Experimental Biology* 224.3, jeb230615. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC7888720/.

Strausfeld, Nicholas J. (Aug. 2002). "Organization of the honey bee mushroom body: Representation of the calyx within the vertical and gamma lobes". In: *Journal of Comparative Neurology* 450.1, pp. 4–33. URL: /doi/pdf/10.1002/cne.10285%20https://onlinelibrary.wiley.com/doi/abs/10.1002/cne.10285%20https://onlinelibrary.wiley.com/doi/10.1002/cne.10285.

Strausfeld, Nicholas J., Uwe Homburg, and Peter Kloppenberg (Aug. 2000). "Parallel organization in honey bee mushroom bodies by peptidergic kenyon cells". In: *Journal of Comparative Neurology* 424.1, pp. 179–195. URL: /doi/pdf/10.1002/1096-9861%2820000814%29424%3A1%3C179%3A%3AAID-CNE13%3E3.0.CO%3B2-K%20https://onlinelibrary.wiley.com/doi/abs/10.1002/1096-9861%2820000814%29424%3A1%3C179%3A%3AAID-CNE13%3E3.0.CO%3B2-K%20https://onlinelibrary.wiley.com/doi/10.1002/1096-9861(20000814)424:1%3C179::AID-CNE13%3E3.0.CO;2-K.

Strube-Bloss, Martin Fritz, Martin Paul Nawrot, and Randolf Menzel (Feb. 2011). "Mushroom Body Output Neurons Encode Odor–Reward Associations". In: *Journal of Neuroscience* 31.8, pp. 3129–3140. URL: https://www.jneurosci.org/content/31/8/3129%20https://www.jneurosci.org/content/31/8/3129.abstract.

Suenami, Shota, Satoyo Oya, Hiroki Kohno, and Takeo Kubo (Oct. 2018). "Kenyon cell subtypes/populations in the Honeybee Mushroom Bodies: Possible function based on their gene expression profiles, differentiation, possible evolution, and application of genome editing". In: *Frontiers in Psychology* 9.OCT, p. 376105. URL: www.frontiersin.org.

Suenami, Shota, Rajib Kumar Paul, Hideaki Takeuchi, Genta Okude, Tomoko Fujiyuki, Kenichi Shirai, and Takeo Kubo (June 2016). "Analysis of the Differentiation of Kenyon Cell Subtypes Using Three Mushroom Body-Preferential Genes during Metamorphosis in the Honeybee (Apis mellifera L.)" In: *PLOS ONE* 11.6, e0157841. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0157841.

Szyszka, Paul, Alexander Galkin, and Randolf Menzel (June 2008). "Associative and non-associative plasticity in Kenyon cells of the honeybee mushroom body". In: *Frontiers in Systems Neuroscience* 2.JUN. URL: https://pubmed.ncbi.nlm.nih.gov/18958247/%20https://pubmed.ncbi.nlm.nih.gov/18958247/?dopt=Abstract.

Takeda, K. (July 1961). "Classical conditioned response in the honey bee". In: *Journal of Insect Physiology* 6.3, pp. 168–179. URL: https://www.sciencedirect.com/science/article/abs/pii/0022191061900609.

Takeuchi, Hideaki, Eriko Kage, Miyuki Sawata, Azusa Kamikouchi, Kazuaki Ohashi, Maya Ohara, Tomoko Fujiyuki, Takekazu Kunieda, Kazuhisa Sekimizu, Shunji Natori, and Takeo Kubo (2001). "Identification of a novel gene, Mblk-1, that encodes a putative transcription factor expressed preferentially in the large-type Kenyon cells of the honeybee brain". In: *Insect Molecular Biology* 10.5, pp. 487–494. URL: https://pubmed.ncbi.nlm.nih.gov/11881813/.

Talay, Mustafa, Ethan B. Richman, Nathaniel J. Snell, Griffin G. Hartmann, John D. Fisher, Altar Sorkaç, Juan F. Santoyo, Cambria Chou-Freed, Nived Nair, Mark Johnson, John R. Szymanski, and Gilad Barnea (Nov. 2017). "Transsynaptic Mapping of Second-Order Taste Neurons in Flies by trans-Tango". In: *Neuron* 96.4, pp. 783–795.

Tanaka, Nobuaki K., Hiromu Tanimoto, and Kei Ito (June 2008). "Neuronal assemblies of the Drosophila mushroom body". In: *Journal of Comparative Neurology* 508.5, pp. 711–755. URL: https://pubmed.ncbi.nlm.nih.gov/18395827/.

Tedjakumala, Stevanus R., Jacques Rouquette, Marie Laure Boizeau, Karen A. Mesce, Lucie Hotier, Isabelle Massou, and Martin Giurfa (July 2017). "A tyrosine-hydroxylase characterization of dopaminergic neurons in the honey bee brain". In: *Frontiers in Systems Neuroscience* 11, p. 263462. URL: www.frontiersin.org.

Tempel, B. L., N. Bonini, D. R. Dawson, and W. G. Quinn (1983). "Reward learning in normal and mutant Drosophila". In: *Proceedings of the National Academy of Sciences of the United States of America* 80.5 I, pp. 1482–1486. URL: https://pubmed.ncbi.nlm.nih.gov/6572401/.

Tully, T, T Preat, S C Boynton, and M Del Vecchio' (1994). "Genetic Dissection of Consolidated Memory in Drosophila". In: *Cell* 79, pp. 35–47.

Tully, T and W G. Quinn (Mar. 1985). "Classical conditioning and retention in normal and mutant Drosophila melanogaster". In: *Journal of Comparative Physiology A* 157.2, pp. 263–277. URL: https://pubmed.ncbi.nlm.nih.gov/3939242/.

Vergoz, Vanina, Edith Roussel, Jean Christophe Sandoz, and Martin Giurfa (Mar. 2007). "Aversive Learning in Honeybees Revealed by the Olfactory Conditioning of the Sting Extension Reflex". In: *PLOS ONE* 2.3, e288. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0000288.

Vieira, Amanda Rodrigues, Nayara Salles, Marco Borges, and Theo Mota (May 2018). "Visual discrimination transfer and modulation by biogenic amines in honeybees". In: *Journal of Experimental Biology* 221.9. URL: https://dx.doi.org/10.1242/jeb.178830.

Vogt, Katrin, Yoshinori Aso, Toshihide Hige, Stephan Knapek, Toshiharu Ichinose, Anja B. Friedrich, Glenn C. Turner, Gerald M. Rubin, and Hiromu Tanimoto (Apr. 2016). "Direct neural pathways convey distinct visual information to drosophila mushroom bodies". In: *eLife* 5.APRIL2016.

Waddell, Scott (June 2013). "Reinforcement signaling in Drosophila; dopamine does it all after all". In: *Current opinion in neurobiology* 23.3, 10.1016/j.conb.2013.01.005. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC3887340/.

Walker, Samuel James, Veró Nica, María Corrales-Carvajal, and Carlos Ribeiro (2015). "Postmating Circuitry Modulates Salt Taste Processing to Increase Reproductive Output in Drosophila Article Postmating Circuitry Modulates Salt Taste Processing to Increase Reproductive Output in Drosophila". In: *CURBIO* 25, pp. 2621–2630. URL: http://dx.doi.org/10.1016/j.cub.2015.08.043http://dx.doi.org/10.1016/j.cub.2015.08.043.

Witthöft, W. (Jan. 1967). "Absolute anzahl und verteilung der zellen im him der honigbiene". In: *Zeitschrift für Morphologie der Tiere* 61.1, pp. 160–184. URL: https://link.springer.com/article/10.1007/BF00298776.

Wright, Geraldine A., Julie A. Mustard, Sonya M. Kottcamp, and Brian H. Smith (Nov. 2007). "Olfactory memory formation and the influence of reward pathway during appetitive learning by honey bees". In: *Journal of Experimental Biology* 210.22, pp. 4024–4033. URL: https://dx.doi.org/10.1242/jeb.006585.

Wright, Geraldine A., Julie A. Mustard, Nicola K. Simcock, Alexandra A.R. Ross-Taylor, Lewis D. McNicholas, Alexandra Popescu, and Frédéric Marion-Poll (Dec. 2010). "Parallel Reinforcement Pathways for Conditioned Food Aversions in the Honeybee". In: *Current Biology* 20.24, p. 2234. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC3011020/.

Yamagata, Nobuhiro, Toshiharu Ichinose, Yoshinori Aso, Pierre Yves Plaçais, Anja B. Friedrich, Richard J. Sima, Thomas Preat, Gerald M. Rubin, and Hiromu Tanimoto (Jan. 2015). "Distinct dopamine neurons mediate reward signals for short- and long-term memories". In: *Proceedings of the National Academy of Sciences of the United States of America* 112.2, pp. 578–583. URL: https://www.pnas.org/content/112/2/578%20https://www.pnas.org/content/112/2/578.abstract.

Yang, Shiyi, Sean E. Corbett, Yusuke Koga, Zhe Wang, W. Evan Johnson, Masanao Yajima, and Joshua D. Campbell (Mar. 2020). "Decontamination of ambient RNA in single-cell RNA-seq with DecontX". In: *Genome Biology* 21.1. URL: https://pubmed.ncbi.nlm.nih.gov/32138770/.

Yin, J. C.P., J. S. Wallach, M. Del Vecchio, E. L. Wilder, H. Zhou, W. G. Quinn, and T. Tully (Oct. 1994). "Induction of a dominant negative CREB transgene specifically blocks long-term memory in Drosophila". In: *Cell* 79.1, pp. 49–58. URL: https://pubmed.ncbi.nlm.nih.gov/7923376/.

Young, Fletcher J., Amaia Alcalde Anton, Lina Melo-Flórez, Antoine Couto, Jessica Foley, Monica Monllor, W. Owen McMillan, and Stephen H. Montgomery (Feb. 2024). "Enhanced long-term memory and increased mushroom body plasticity in Heliconius butterflies". In: *iScience* 27.2, p. 108949. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC10864207/.

Zhang, Wenxin, Liangliang Wang, Yinjiao Zhao, Yufei Wang, Chaoyang Chen, Yu Hu, Yuanxiang Zhu, Hao Sun, Ying Cheng, Qinmiao Sun, Jian Zhang, and Dahua Chen (July 2022). "Single-cell transcriptomic analysis of honeybee brains identifies vitellogenin as caste differentiation-related factor". In: *iScience* 25.7, p. 104643. URL: http://www.cell.com/article/S2589004222009154/fulltext%20http://www.cell.com/article/S2589004222009154/abstract%20https://www.cell.com/iscience/abstract/S2589-0042(22)00915-4.

Zhang, Y V., Jinfei Ni, and Craig Montell (June 2013). "The molecular basis for attractive salt-taste coding in drosophila". In: *Science* 340.6138, pp. 1334–1338. URL: http://science.sciencemag.org/.

Zhang, Yameng, Allan Hermann Pool, Tongtong Wang, Lu Liu, Elin Kang, Bei Zhang, Liang Ding, Kirsten Frieda, Richard Palmiter, and Yuki Oka (Dec. 2023a). "Neural control of sodium consumption and taste valence". In: *Cell* 186.26, p. 5751. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC10761003/.

— (Dec. 2023b). "Parallel neural pathways control sodium consumption and taste valence". In: *Cell* 186.26, pp. 5751–5765. URL: https://www.sciencedirect.com/science/article/pii/S0092867423011716#fig3.

Zheng, Grace X.Y., Jessica M. Terry, Phillip Belgrader, Paul Ryvkin, Zachary W. Bent, Ryan Wilson, Solongo B. Ziraldo, Tobias D. Wheeler, Geoff P. McDermott, Junjie Zhu, Mark T. Gregory, Joe Shuga, Luz Montesclaros, Jason G. Underwood, Donald A. Masquelier, Stefanie Y. Nishimura, Michael Schnall-Levin, Paul W. Wyatt, Christopher M. Hindson, Rajiv Bharadwaj, Alexander Wong, Kevin D. Ness, Lan W. Beppu, H. Joachim Deeg, Christopher McFarland, Keith R. Loeb, William J. Valente, Nolan G. Ericson, Emily A. Stevens, Jerald P. Radich, Tarjei S. Mikkelsen, Benjamin J. Hindson, and Jason H. Bielas (Jan. 2017). "Massively parallel digital transcriptional profiling of single cells". In: *Nature Communications* 8, p. 14049. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC5241818/.