Identification of the Transneuronal Homeostatic Machinery at a Central Synapse



Evan Harrell

Supervisor: Professor Gero Miesenboeck

Medical Sciences Division

University of Oxford

This report is submitted in partial fulfillment for

Doctor of Philosophy

Hilary 2014

To my family, for all their love and support.

0.1 Acknowledgements

First and foremost I'd like to thank my supervisor, Professor Gero Miesenboeck, for putting up with me and for always providing encouraging ideas and enthusiasm. Refer to Table 1 for a full list of the significant contributors to this thesis.

Name	p-value	Description
Prof GM	1.20E-08	supervisor, mentor, major creative force and financier
Harrell Family	6.04E-07	always asked "When are you going to get a real job?"
Amelie Baud	3.40E-06	journal clubs!
Andrew Lin	3.40E-06	person I go to with difficult questions, always helpful
Christoph Treiber	3.40E-06	always available for helpful discussions and advice
Clara Ferreira	3.40E-06	acted as punching bag for at least 3 years
David Owald	3.40E-06	watched some great football together and gave important advice
Diogo Pimentel	3.40E-06	always willing to help and gave important advice
Eleftheria Vrontou	3.40E-06	helpful and nice to chat with in the lab
Emmanuel Perisse	3.40E-06	many great discussions
Gaurav Das	3.40E-06	many discussions and a great cook!
Iain Mathieson	3.40E-06	journal clubs!
Jared O'Connell	3.40E-06	journal clubs! And pizza!
Jeff Donlea	3.40E-06	gave great advice and great to be around in the lab
Jessica Beevers	3.40E-06	a great decapitator and brain dissector
Michael Song	3.40E-06	an honourable servant
Ruth Brain	3.40E-06	fellow fly executioner and always available to help
Scott Waddell	3.40E-06	made sure to let me know if I was wearing too much of one colour, lots of advice
Shiko Parnas	3.40E-06	great discussions and advice
Suewei Lin	3.40E-06	helpful and knowledgeable about everything
Vincent Croset	3.40E-06	illustrator guru and fellow RNA enthusiast
Wolf Huetteroth	3.40E-06	anatomy pimp (hi wolfie wolf!)

Table 1: *Significant contributions to this thesis*

0.2 Abstract

Two different kinds of stabilising homeostatic behaviour have been observed in neurons. The first type involves the cell-autonomous maintenance of a cell-identity-based level of electrical activity. Neurons continually monitor their own electrical activity and can adjust many intracellular parameters, such as membrane ion channel densities, to keep this activity within a tight physiological range. The second type of homeostatic behaviour shares the same goal, to maintain a fixed level of electrical activity, but instead of adjusting intracellular parameters, the neuron recruits its synaptic partners to assist in maintaining a genetically prescribed activity level. This behaviour is most easily observed when a neuron is either electrically silenced by expressing an inwardly-rectifying potassium channel or rendered less sensitive to neurotransmitter through mutation of its postsynaptic receptors. Both of these perturbations result in increased synaptic drive from the presynaptic cells, either through increasing the number of neurotransmitter release sites or increasing the probability of release from single release sites.

Many genes that are instrumental in the second type of homeostatic behaviour have been identified, mainly at the neuromuscular junction in the peripheral nervous system. However, studies on transsynaptic homeostatic compensation in an intact *central nervous system* have been few and far between. Also, which, if any, of the homeostatic genes are transcriptionally regulated in the nucleus after the onset of transsynaptic homeostatic adjustment, has not been adequately addressed. This thesis has developed a system to measure transcriptome-wide gene expression levels in presynaptic circuit elements after altering the firing properties of the downstream circuit in the CNS. Many transcriptionally regulated genes have been identified and are now being tested for their potential use as reporters for transsynaptic transcriptional regulation. It might be possible to capitalise on endogenous homeostatic signalling pathways to gain genetic access to synaptically connected neurons.

0.3 Prelude

Since the beginning of the 20th century, *Drosophila melanogaster* (DM) has been an essential experimental model in the fields of genetics, developmental biology and neuroscience. First imagined to be an inexpensive teaching tool for student projects (Kohler, 1994), it is remarkable that DM eventually allowed scientists to uncover mysteries ranging from the basic principles of chromosomal organisation to the developmental processes involved in body layout and segmentation, all the way to the interaction between genes, neural circuits, and behaviour (Weiner, 1999). More recently, vast improvements in available technology have permitted *in vivo* interrogation of genetically circumscribed sets of neurons using engineered biomolecules (Luo *et al.*, 2008). Flies can now be generated expressing different transgenes in as many as three independent populations of cells (Brand & Perrimon, 1993; Lai & Lee, 2006; Potter & Luo, 2011).

Incredibly, despite such a glowing history and an undoubtedly promising future, some people still have not quite understood just how important DM is to neurobiology.

> ...Where does a lot of that earmark money end up anyway? You guys have heard some of the examples of where those dollars go. You have heard about the bridges, and you've heard about some of these pet projects that really don't make a whole lot of sense. And sometimes these dollars, they go to projects having little or nothing to do with the public good. Things like fruit fly research in Paris, France. I kid you not!...
>
> Sarah Palin, 2008

It seems that the former Governor of Alaska and candidate for the Vice Presidency Sarah Palin does not believe that understanding the fundamental processes of neuronal development and plasticity, homeostatic synaptic maintenance, and learning and memory provide any public benefit. In almost the same breath, she declared the importance of the early identification of autism-related developmental abnormalities in improving the prospects of both affected children and their parents. Ironically, neurexin was implicated as a risk factor for autism (Kim *et al.*, 2008) just before Governor Palin promised to cut fruit fly funding, and thank god she lost the subsequent election because much has been learned about neurexin since then using fruit flies as the model organism (Owald *et al.*, 2012; Knight *et al.*, 2011).

Unfortunately, Governor Palin's ignorance (definitely not limited to her knowledge of fruit fly research) about the importance of DM as an experimental model for neuroscientists is shared by the vast majority of the public and even scientists in related fields. Fellow neuroscientists and journal editors alike are very quick to label fruit fly work as low impact or not relevant to the brain function of higher beings. The large amount of time

that the average DM researcher spends arguing about whether discoveries are pertinent because of the simplicity of the DM central nervous system needs to diminish.

It is my strong belief that the major developments in the understanding of neural assemblies and perhaps even the neural correlates of consciousness will be elucidated in fruit flies (I will spare you the debate on whether flies are conscious) long before they are understood in species with larger brains. An ever-growing genetics toolbox, well-defined anatomy, and high-throughput behavioural paradigms will pave the way for huge advances in our understanding of neural circuitry and provide a foundation from which to tackle the more complex nervous systems of mammals.

Casting those lofty ambitions aside, my thesis work set out to identify the molecular machinery that allows neurons to work not only as single independent entities, but as a collective unit to maintain a necessary function. Put more simply, the aim was to look at a specific class of neuron and identify a set of genes whose expression level can be regulated based on the electrical activity of that neuron's immediate postsynaptic partners. The advanced DM genetics toolbox is unique in allowing access to anatomically divisible ensembles of pre- and postsynaptic neurons, while also granting the ability to express effector genes in one group of cells but not the other.

I intend to use this transsynaptically regulated set of genes to design reporter systems capable of retrograde circuit tracing. This kind of system would be a vast improvement on current neuroanatomical tracing methods depending either on active uptake of injected substances or neurotropic viruses. The dream is to produce a purely genetic retrograde circuit tracing paradigm by hijacking the pathways used in endogenous transsynaptic signalling. Such a system could deliver both connectivity *AND* functional information because neurons with stronger connections probably also have a stronger transcriptional impact on one other (Callaway, 2008). My one sentence goal in this work is to establish a system whereby expressing an effector gene in a single neuron of interest, it will be possible to address/visualise the main *inputs* to that neuron.

Contents

0.1 Acknowledgements . . . ii
0.2 Abstract . . . iii
0.3 Prelude . . . iv

1 Introduction **1**
1.1 History and Basics . . . 1
1.2 Neuroanatomical Tracing After Cajal . . . 4
1.2.1 Non-viral Injected Substances . . . 5
1.2.1.1 Anterograde and Bidirectional Tracing . . . 5
1.2.1.2 Retrograde Tracing . . . 6
1.2.2 Endogenously Expressed Peptides . . . 7
1.2.2.1 Wheat Germ Agglutinin . . . 8
1.2.2.2 C-terminal Fragment of Tetanus Toxin . . . 8
1.2.3 Neurotropic Viruses . . . 9
1.2.3.1 Rabies Virus . . . 9
1.2.3.2 Herpes Viruses . . . 10
1.2.4 High Resolution Methods . . . 11
1.2.5 A New Approach . . . 13
1.3 DM . . . 13
1.3.1 Genetics . . . 13
1.3.2 Behavioural Repertoire . . . 16
1.3.2.1 Olfaction . . . 16
1.3.2.2 Circadian Rhythm . . . 18
1.3.3 Anatomy . . . 20
1.3.4 The AL Circuit and the ORN/PN Synapse . . . 22
1.4 Homeostasis and Plasticity . . . 25
1.4.1 Plasticity in the Nervous System . . . 26
1.4.2 Homeostatic Activity in the Nervous System . . . 27
1.4.2.1 Intrinsic Activity Maintenance . . . 27
1.4.2.2 Transsynaptic Homeostatic Maintenance . . . 29
1.5 Transcriptome Sequencing . . . 32

1.5.1 Standard RNA-seq . . . 33
1.5.2 3' Tag DGE . . . 34
1.5.3 Bioinformatics . . . 35
1.5.3.1 Pre-mapping Read Quality Assessment . . . 35
1.5.3.2 Mapping . . . 37
1.5.3.3 Transcript Allocation . . . 37

2 Materials and Methods 39
2.1 Flies . . . 39
2.1.1 Animal Rearing . . . 39
2.1.2 Food . . . 39
2.1.3 Lines . . . 39
2.1.3.1 Commercially Available Lines . . . 39
2.1.3.2 Generated Lines . . . 42
2.2 Molecular Biology . . . 44
2.2.1 Kir2.1 and Kir2.1-nc Cloning . . . 44
2.2.2 Or-LexA Cloning . . . 46
2.3 Behavioural Monitoring . . . 46
2.3.1 Odour Valence Measurement . . . 46
2.3.2 Circadian Rhythm . . . 48
2.4 Imaging . . . 48
2.4.1 Immunohistochemistry . . . 48
2.4.2 Native Protein Imaging . . . 49
2.4.3 Confocal microscopy . . . 49
2.5 RNA Sample Preparation . . . 49
2.5.1 Third Antennal Segment Separation . . . 49
2.5.2 RNA extraction . . . 49
2.5.2.1 RNA-seq and RT-qPCR . . . 49
2.5.2.2 3' Tag DGE . . . 50
2.6 cDNA Library Generation . . . 51
2.6.1 RNA-seq and RT-qPCR . . . 51
2.6.2 3' tag DGE . . . 51
2.7 Transcriptome Sequencing . . . 51
2.7.1 Transcriptome Data Analysis . . . 51
2.8 Real Time quantitative PCR (RT-qPCR) . . . 57

3 Results 58
3.1 Early Lessons . . . 58
3.2 The DM Third Antennal Segment Transcriptome . . . 59

3.2.1 Pilot Experimental Design . . . 59
3.2.2 Purity Assessment . . . 60
3.2.2.1 Olfactory Receptor Expression . . . 60
3.2.2.2 Gustatory Receptor Expression . . . 62
3.2.2.3 Ionotropic Receptor Expression . . . 64
3.2.2.4 Johnston's Organ Comparison . . . 66
3.2.3 Circadian Gene Expression Regulation . . . 67
3.3 The Kir2.1/Kir2.1-nc System . . . 72
3.3.1 Kir2.1: Expression and Localisation . . . 72
3.3.2 Behavioural Validation . . . 73
3.3.3 Electrophysiological Validation . . . 78
3.4 Identification of the Homeostatic Machinery . . . 81
3.4.1 3′ DGE Results . . . 81
3.4.1.1 Purity Assessment . . . 82
3.4.1.2 Differential Expression . . . 84
3.4.2 RNA-seq Results . . . 92
3.4.2.1 Purity Assessment . . . 94
3.4.2.2 Differential Expression . . . 95
3.4.3 RT-qPCR and Sequencing Platform Validation . . . 99

4 Discussion **102**
4.1 Homeostatic Gene Expression Regulation . . . 102
4.1.1 Early Proliferation . . . 102
4.1.1.1 Synaptic Genes . . . 104
4.1.1.2 Homeostatic Genes . . . 106
4.1.2 Late Stress . . . 107
4.1.2.1 Synaptic and Homeostatic Genes . . . 108
4.2 Homeostasis-based Circuit Tracing Systems . . . 109
4.2.1 Genetics . . . 109
4.2.2 Incorporating Spatial Information - Intersectional Logic . . . 111
4.3 Summary of Future Work . . . 113
4.4 Conclusion . . . 114

Bibliography **115**

Chapter 1

Introduction

1.1 History and Basics

Nearly five millennia have passed since the first descriptions of brain injury were documented in the *Edwin Smith Surgical Papyrus*, a scroll most likely written by Imhotep, the grand vizier for the pharaoh Zoser (Gross, 1987). Evidently, the ancient Egyptians were aware of the brain playing a role in seizures and aphasia, and even had a notion of the laterality of brain function. Despite this, the prevailing belief at the time placed the essence of the mind in the heart, which is why embalming practices typically involved careful cardiac preservation while mindlessly discarding the brain, an act which obliges one to think almost involuntarily of the old American advertising slogan, "A mind is a terrible thing to waste."

It fell to the ancient Greeks to realise the importance of the brain. They confidently declared that the brain is the centre of the senses and Hippocrates is credited with the following astonishing hypothesis around the turn of the 4th century B.C. (Gross, 1987):

> It ought to be generally known that the source of our pleasure, merriment, laughter, and amusement, as of our grief, pain, anxiety, and tears is none other than the brain. It is specially the organ which enables us to think, see, and hear, and to distinguish the ugly and the beautiful, the bad and the good, pleasant and unpleasant ... It is the brain too which is the seat of madness and delirium, of the fears and frights which assail us ... it is there where lies the cause of insomnia and sleep-walking, of thoughts that will not come, forgotten duties, and eccentricities.
>
> Hippocrates, ≈400 B.C.

Perhaps less astonishingly, Francis Crick delivered his own self-proclaimed *Astonishing Hypothesis* in 1995 with the first line in his book (Crick, 1995).

> The Astonishing Hypothesis is that "You," your joys and your sorrows, your memories and your ambitions, your sense of identity and free will are in fact no more than the behaviour of a vast assembly of nerve cells and their associated molecules.
>
> Francis Crick, 1995

Comparing these two *hypotheses* might lead one to believe that there has not been much progress in brain science over the past two millennia. That is definitely not the case.

A few hundred years after Hippocrates' hypothesis, another Greco-Roman doctor named Galen greatly advanced the study of the brain and nervous system. He dissected and compared brains from many different species, pioneered the practice of lesioning areas of the brain to study their relevance, and began to unravel the sensory and motor nerve loops to and from the spinal cord. His findings were the reference book for medical students for practically a millennium and a half (Rocca, 1998).

It was not until the middle of the 16th century that Andreas Vesalius noticed that many of Galen's descriptions were not actually based on human anatomy. Vesalius took this opportunity to re-write the book on human brain anatomy, which he did with great aplomb in 1543 with the publication of his masterpiece *On the Fabric of the Human Body*. Even though he is remembered as perhaps the most important anatomist in the history of science, Vesalius refused to make any functional speculations based on his anatomical work. He did venture to say that the human ventricles, thought at the time to be where reason and the soul might reside, were not different from the ventricles of some other mammalian species, and therefore current physiological and cognitive models placing the seat of the human soul in the ventricles were probably incorrect (Finger, 2005).

Mindful of Vesalius' progress and on the crest of a blossoming scientific method, Rene Descartes laid down the philosophical and somewhat physiological framework of dualism by the middle of the 17th century, a concept allowing the body its own reflex responses while at the same time permitting the *soul* to intervene if required. Historians consider this to be the birth of the field of neuroscience. Dualism was the first theory to unify the ideas of man as a reflex machine and as an unpredictable being. According to Descartes, the cerebrospinal fluid (CSF) flowing out from the ventricles was the substance that the mind, more specifically the pineal gland, used to literally brainwash the body into action (Damasio, 2008). It turns out the CSF might be doing some brainwashing during sleep, which is not exactly the kind Descartes had in mind (Xie *et al.*, 2013).

While the philosophers argued, the physiologists discovered the electrical nature of the nervous system, and the neuroanatomists toiled with the cellular complexity of the neuronal trellis. Nevermind the sheer number of cells, the density of the tissue had some

believing that individual nerve cells might not even exist, a concept dubbed reticular theory (Raviola & Mazzarello, 2011). A competing view, that fundamental anatomical units called neurons make up the nervous system, was suggested by Wilhelm Waldeyer (Lopez-Munoz *et al.*, 2006).

The late 19th century was a controversial time for neuroanatomists. Enter Camillo Golgi and Santiago Ramon y Cajal. Rather ironically, Cajal used Golgi's eponymous histological staining methods to disprove reticular theory, of which Golgi was the primary supporter, and establish the neuron doctrine (Cajal, 1906; Golgi, 1906). Around the same time, Charles Sherrington started calling the connections between neurons *synapses* (Sherrington, 1897). Another ongoing debate was whether synapses were electrical or chemical. It turns out that both chemical and electrical synapses (Furshpan & Potter, 1957) exist, but this thesis will focus on the canonical chemical synapse.

A *chemical synapse*, seen in the schematic in Figure 1.1, is the connection between a presynaptic neuron and a postsynaptic cell (Cowan *et al.*, 2001). Information flows from the presynaptic neuron to the postsynaptic cell in the form of a molecule called a neurotransmitter. Bioelectric signals propagate along neuronal processes causing changes in membrane polarisation. When the presynaptic neuron is depolarised above a certain threshold, an all-or-nothing electrical event known as an action potential (or spike) is triggered, which travels down the axon of the presynaptic neuron thanks to the concerted activity of voltage-gated sodium and potassium channels (Hodgkin & Huxley, 1952). When the action potential reaches the synaptic nerve endings, voltage-gated calcium channels open and the influx of calcium causes neurotransmitter-containing vesicles to fuse to the presynaptic membrane, kick-starting a process called exocytosis (Haucke *et al.*, 2011). Exocytosis is the dumping of chemical transmitter into the synaptic cleft. The neurotransmitter freely diffuses across the cleft, where it is detected by specialised postsynaptic receptors.

Several different types of receptors can mediate different effects on the postsynaptic cell. Ionotropic receptors directly gate the influx of ions, which can depolarize or hyperpolarize the postsynaptic cell. Another type of receptor called a metabotropic receptor triggers intracellular signalling systems, usually mediated by G-protein cascades, and can also excite or inhibit firing in the postsynaptic cell (Sheng, 2001). Excitatory synapses depolarise the postsynaptic cell and inhibitory synapses hyperpolarise it. Postsynaptic neurons integrate their excitatory and inhibitory inputs, and if the excitation outweighs the inhibition by enough to cross the action potential threshold, the postsynaptic neuron fires a spike itself and passes the signal to its downstream neurons in much the same fashion (Quiroga & Panzeri, 2013).

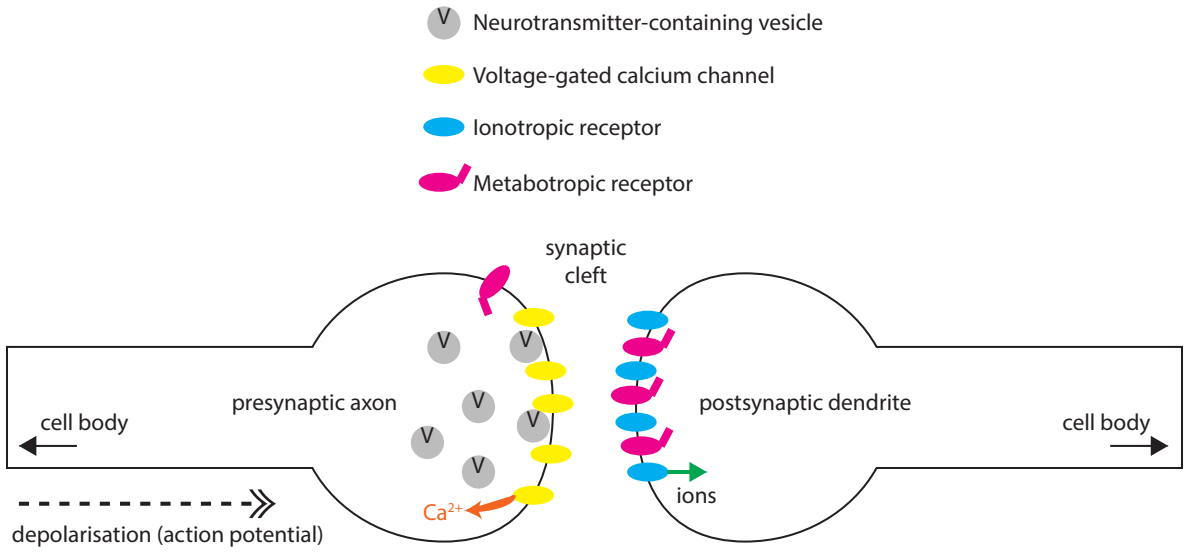
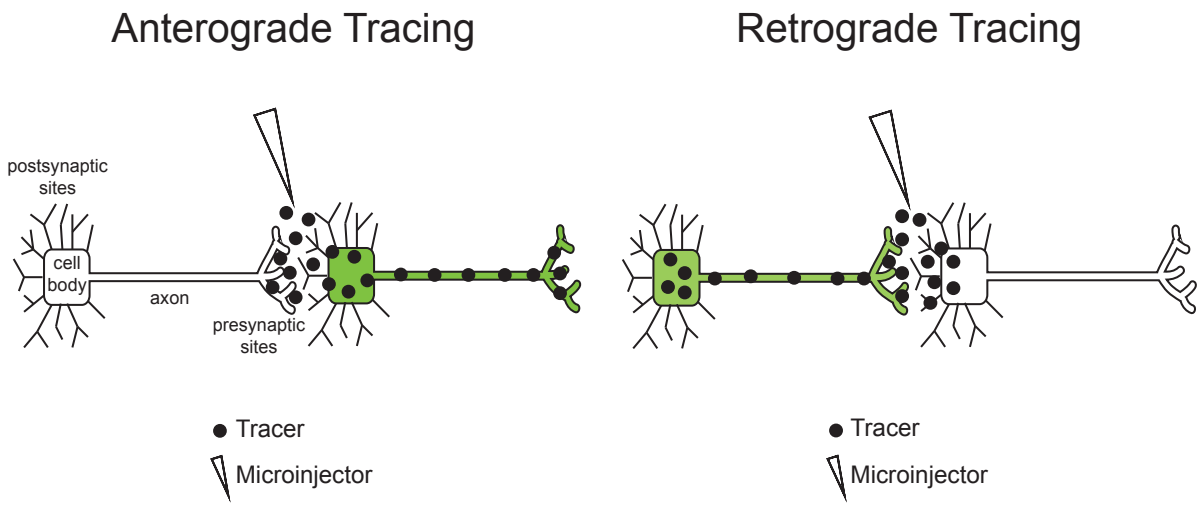


Figure 1.1: *Simplified schematic of an excitatory chemical synapse*

Obviously, team neuron, with Cajal as the figurehead, ultimately won the argument about the organisation of the nervous system. Cajal is the one who took most of the plaudits because he not only identified the fundamental units of the nervous system and the strict polarisation of nerve cells, but originated the very idea that *structure underlies function* in the nervous system (Llinas, 2003). Granted no one before Cajal saw the structure in quite so much detail, but that does not take away at all from the novelty of his conclusions. Where Vesalius refused to put forth hypotheses about function based solely on anatomy, Cajal made his career doing just that. The *Cajalian* paradigm of studying the fine neuronal structure to make deductions about function remains at the forefront of the fields of systems and molecular neuroscience alike.

1.2 Neuroanatomical Tracing After Cajal

Neuroanatomical techniques have evolved a great deal in the last 50 years. Early studies like Cajal's were restricted to the use of inorganic dyes in fixed tissue, but towards the end of the 20th century, researchers discovered superior agents that living neurons can autonomously take up from the extracellular space (Malmgren *et al.*, 1978). A modern tracing experiment, as diagrammed in Figure 1.2, typically involves tracer microinjection at a location of interest in a living brain, uptake of the tracer by the neighbouring neurons, transport within the cells, and immunohistochemical detection to visualise the cells and their labelled neurites (Kobbert *et al.*, 2000). Many different kinds of substances have been used as neuroanatomical tracers, each of which has advantages and shortcomings. The substances can be divided into the following three groups: non-viral injected



(a) *Simple schematic for anterograde tracing. The tracer is microinjected, taken up, and transported away from the cell body along the axon.* **(b)** *Simple schematic for retrograde tracing. The tracer is microinjected, taken up in the axon, and transported back to the cell body.*

Figure 1.2: *Different types of neuroanatomical tracing*

substances, endogenously expressed peptides, and neurotropic viruses.

1.2.1 Non-viral Injected Substances

1.2.1.1 Anterograde and Bidirectional Tracing

In an *anterograde* tracing experiment (Figure 1.2a), the goal is to visualise the axonal projections of a particular group of neurons. Thus, the tracer is applied near the cell bodies or dendrites of the neurons of interest, taken up, transported along the axons, and then either processed enzymatically or immunohistochemically for visualisation. Most substances used for anterograde tracing also are transported retrogradely to a lesser extent, which makes them bidirectional in nature. The first anterograde tracing experiments were done using tritiated amino acids (and carbohydrates). All cells import amino acids from the surrounding environment which are incorporated into proteins and spread throughout the cytoplasm. Thus, injection of radio-labeled amino acids results in the synthesis of radio-labeled proteins, which then emit beta particles which react with silver bromide and can be visualised (Cowan *et al.*, 1972).

It was quickly realised that not only do neurons take up the radio-labeled amino acids and transport them along their axons to the nerve endings, but they also deliver them to their postsynaptic targets. This discovery was used to great effect in the mapping of the Rhesus macaque visual cortex (Wiesel *et al.*, 1974). Tritiated proline and tritiated fucose were injected into the vitreous material in the left eye. After a 3 week waiting

period, autoradiographs revealed strong labelling in the lateral geniculate nucleus (LGN) and striate cortex layer 4, both known to be important projection sites in the primary visual pathways (Wiesel *et al.*, 1974). While this has proven useful for tracing the crude anatomical layout of a neural circuit, it can also be transmitted non-synaptically by mere cell-to-cell contact. This makes it less reliable for identifying fine connectivity.

More recently, non-radioactive tracing molecules have taken centre stage as they can be visualised directly within the neuronal cytoplasm without having to worry about the management of radioactive isotopes and waste disposal. Biotinylated dextran amine (BDA) and *Phaseolus vulgaris*-lecoagglutinin (PHA-L) are the two most commonly used of such tracers (der Weel *et al.*, 1994; Reiner *et al.*, 2000). Both can be detected either histochemically or enzymatically and homogeneously fill the labeled neurons. They can even be used together and visualised on separate channels. However, neither of them travel across synapses like the radio-labeled molecules, and they also retrogradely label dendrites. Although retrograde transport is slower than anterograde (Lanciego & Wouterlood, 2011), the use of BDA and PHA-L is limited to visualising projections of the neurons near the site of microinjection.

Other substances, like neurobiotin and biocytin, have often been used in conjunction with intracellular electrophysiological recordings to visualise the full anatomical extent of a recorded neuron (Wilson *et al.*, 2004). After being placed in the recording electrode, the tracer diffuses out into the cell through the disrupted membrane and can subsequently be stained and visualised. Again, tracing is bidirectional within the neuron because the tracer fills the entire cytoplasm. These substances are less stable so need to be processed relatively quickly, and also are not actively taken up so need to be introduced into the cytoplasm through the recording electrode (Lanciego & Wouterlood, 2011). This technique has shown its power in correlating the electrophysiological properties of the many different types of interneurons around the DM antennal lobe with their respective anatomical projection patterns (Wilson *et al.*, 2004; Chou *et al.*, 2010).

1.2.1.2 Retrograde Tracing

A *retrograde* tracing experiment provides the counter approach (Figure 1.2b). The tracer is applied at the projection site, taken up by the local axon terminals, and transported back towards the neuronal cell body. The observation of retrograde morphological hypertrophy after peripheral lesions was recognised in the 19th century by Constantin von Monakow (Lanciego & Wouterlood, 2011). His observations motivated the first experimental approaches utilising retrograde axonal transport of the enzyme horse radish peroxidase (HRP). HRP was injected into either the gastrocnemius muscle or the tongue of rats and

enzymatically detected in motor neurons in the spinal cord and the hypoglossal nerve respectively several days later (Kristensson & Olsson, 1971).

Other typical agents used for retrograde tracing were developed around the same time as HRP and took the form of inorganic dyes like Fast Blue (FB) (Kuypers *et al.*, 1980) and fluoro-gold (FG) (Novikova *et al.*, 1997). They are also taken up at the injection site and have attractive spectral characteristics as well as long term stability, so are ideal for experiments involving long-range anatomical tracing. The non-toxic B fragment of the cholera toxin (CTB) has also been found to be transported retrogradely in neurons after uptake, and can provide a more stable, sensitive option than standard HRP tracing (Trojanowski *et al.*, 1982). CTB-conjugated dyes were recently used effectively to identify which medial pre-frontal cortex (mPFC) neurons project to the nucleus reuniens (NR) in the mouse (Xu & Sudhof, 2013). Two CTB-dye conjugates with different fluorophores were stereotactically injected into the NR and another thalamic nucleus (one colour in each location) in 2-day old mice. After two weeks, the fluorescent CTB could be detected in different neurons in the mPFC, almost none of which were double stained. The conclusion was that independent populations of neurons ennervate the different thalamic nuclei.

Another significant tool used for neuroanatomical tracing is fabricated nanoparticles (beads). Fluorescent beads have been engineered with special coating that allows them to be taken up by axons and transported retrogradely towards the soma (Katz *et al.*, 1984). The chemical composition and size of the beads determines the efficiency with which they are taken up by the neighbouring axons. They are highly stable, do not diffuse away from the site of application, and display no cytotoxicity (Lanciego & Wouterlood, 2011).

1.2.2 Endogenously Expressed Peptides

Both anterograde and retrograde tracing experiments with injected substances have inherent drawbacks. Local application or microinjection of the tracing agent means that any neuron in the area is free to take it up, and depending on how many different types of neurons exist in the region of injection, this can make results hard to interpret (Huh *et al.*, 2010). The solution to these specificity problems is to endogenously express the tracers in the neurons of interest by encoding them in DNA. Endogenously expressed peptides have been found that not only identify single neurons and their processes, but also travel across synapses like the radio-labeled amino acids.

Wheat germ agglutinin (WGA) and the nontoxic C-terminal fragment of tetanus toxin (TTC) are the two most commonly used endogenously expressed tracing agents (Lanciego & Wouterlood, 2011). WGA is a plant lectin and has been used as a highly sensitive tracer

in many studies in a variety of species, and most notably for this thesis in DM (Tabuchi *et al.*, 2000).

1.2.2.1 Wheat Germ Agglutinin

WGA binds to the carbohydrate moieties of synaptic glycoproteins and is imported into the neuron via endosomal transport (Huh *et al.*, 2010). In DM, although the tool was developed more than 10 years ago (Tabuchi *et al.*, 2000), not a single report has been published since then. This is most likely because it travels both anterogradely and retrogradely and crosses synapses, presenting a problem in understanding the functional layout of the circuit.

However, in mammals WGA has been used much more effectively. A WGA-cre conjugate system was recently used in an intersectional approach to limit expression of a transgene in a projection-specific manner (Xu & Sudhof, 2013). In the same report that used CTB mentioned above, a double-floxed tetanus toxin gene was virally transfected into the mPFC and at the same time a WGA-cre conjugate was transfected into a thalamic nucleus of interest. The result of this was tetanus toxin expression only in the mPFC neurons that project to the thalamic nucleus that received the WGA-cre transgene (Xu & Sudhof, 2013). WGA can be a very useful tool, but direction specificity, signal loss, and cytotoxicity present serious problems for its use purely as a synaptic tracer.

1.2.2.2 C-terminal Fragment of Tetanus Toxin

Tetanus toxin is a deadly poison that causes spastic paralysis by cleaving synaptobrevin (Schiavo *et al.*, 1992). The C-terminal fragment (TTC) of the toxin mediates the synaptic spread, but not the harmful effects. In light of this, GFP-fused TTC has been validated as an effective tracer in mammalian systems (Roux *et al.*, 2005). It binds to lipid-rich domains at the presynaptic terminal and is internalised via an endosomal mechanism that is different from the pathway responsible for CTB import (Roux *et al.*, 2005). The TTC-GFP protein has been expressed from different promoters in characterised circuits of the mouse brain and it was verified that it travels exclusively in the retrograde direction (Maskos *et al.*, 2002).

There are no published reports using TTC-GFP in DM and the system has several inherent problems. There is no amplification of the signal beyond its expression from the source. Therefore, as it emanates away from the source, there will be considerable signal loss. Cytotoxicity could be a problem as well as immune responses especially in the mammalian expression systems.

The shortcomings of all agents discussed thus far highlight the need not only for a molecule to be endogenously expressed and synaptically transmitted, but also *amplified* once transported across the synapse. The neurotropic viruses naturally fulfil this requirement.

1.2.3 Neurotropic Viruses

Neurotropic viruses, like most naturally occurring biological agents, exist because of their ability to self-amplify. Some of them have also evolved to travel unidirectionally in neurons and spread synapse-specifically. Bioengineering them to decrease pathogenicity, express marker genes, control synaptic spread, and specify neuronal susceptibility, neuroscientists have managed to harness their circuit tracing power (Callaway, 2008). The virus is injected at a specific location in the nervous system, where it is taken up by particular neurons genetically modified to become susceptible to infection.

Two types of viruses have been used, herpes viruses and rabies viruses. Because the synaptic spread of these viruses is largely based on interaction with host proteins, they have not been used at all in invertebrates and non-mammalian species. Efforts to enable viral infection of non-mammalian species with these viruses by also transgenically expressing essential host proteins have been largely disappointing (Hao *et al.*, 2008). The success of these attempts was limited to viral infection with no signs of active viral spread. In order to make them generally useful across species, the precise interacting partners vital for their propagation need to be identified. Nonetheless, they have provided very powerful methods for tracing mammalian neuronal networks.

1.2.3.1 Rabies Virus

Rabies is a negative sense, single-stranded RNA virus. With a genome that is only 12 kilobases coding for 5 genes, the rabies viral life cycle is easy to understand and control (Fields *et al.*, 2007). It is an enveloped virus, which means that the virus-specific envelope glycoproteins and the viral genetic material bud out of the host cell membrane. Rabies glycoprotein (RG) is the sole envelope protein and specifically interacts with presynaptic proteins, which is thought to be the reason rabies virus spreads exclusively in a retrograde fashion (Lafon, 2005). It has been shown to spread only across chemical synapses, and not via gap junctions or non-synaptic cell-to-cell contacts (Ugolini, 2011).

Controlled infection has been fine-tuned by stripping the rabies virus of its native glycoprotein and replacing it with the subgroup A avian sarcoma and leukosis virus (ASLV-A) glycoprotein EnvA. Coating a virus with EnvA confers a high specificity to infect only cells expressing the TVA receptor (Wickersham *et al.*, 2007). Therefore, by endogenously

expressing TVA as well as the native rabies glycoprotein (RG) in a genetically circumscribed set of cells, researchers have managed to engineer a system that can only spread across a single synapse after infecting a targeted class of cells (Wickersham *et al.*, 2007).

Different versions of this system have been designed using a related virus called the Vesicular Stomatitis Virus (VSV) (Beier *et al.*, 2013), and it has even recently been found that the original rabies virus actually spreads in the anterograde direction when it is transfected into TVA-expressing sensory neurons in the dorsal root ganglion (Zampieri *et al.*, 2014). This illustrates the problems with determining the absolute properties of viral spread. Depending on the type of neuron infected, the same virus may behave in different ways. Even with a very simple genome and well-characterised life cycle, it is clear that there is still much to be discovered about exactly how the rabies virus spreads in different contexts.

1.2.3.2 Herpes Viruses

Unlike rabies, herpes viruses have a complex life cycle, which is not surprising considering the herpes viral genome (a double stranded DNA genome) is roughly ten times as large and carries about 70 genes (Fields *et al.*, 2007). Some of the genes are transcription factors that can have different transcriptomic affects depending on the identity of the infected cell. With herpes viruses, a clear advantage is that standard DNA manipulation techniques can be used to modify the genome (Callaway, 2008). The downside is that depending on the infection phase of the virus, it can travel either in the anterograde direction or the retrograde direction (LaVail *et al.*, 2007).

It has been suggested that the H129 strain of herpes virus moves exclusively in the anterograde direction and this was tested in many characterised circuits in the mouse brain (Lo & Anderson, 2011). However, most of the evidence provided in this report was based on a lack of labelling in known inputs to infected regions, which does not necessarily mean the virus can not spread in the retrograde direction in certain circumstances or to certain classes of neurons. Put more generally, an absence of evidence is not evidence of absence. From the rabies virus story, it is clear that cell identity impacts viral spread, so testing in multiple systems is required before confidently assuming unidirectional spread.

Herpes viruses also have a varying degree of cytotoxicities and may provoke immune responses that can negatively impact experimental results. A less virulent herpes virus, called the Pseudo-rabies Virus (PRV) Bartha strain, has been used successfully as a retrograde tracer in many experimental systems (Ekstrand *et al.*, 2008). The problem with PRV is that it can infect many different types of cells, and can spread non-synaptically (Ekstrand *et al.*, 2008). It also provokes strong glial responses, which turn out to actually help the applicability of the system by stifling non-synaptic spread of

the virus (Tomishima & Enquist, 2002; Rinaman *et al.*, 1993). It is less cytotoxic than other herpes virus strains, so it does not kill the infected host cells and bud out to neighbouring, non-synaptically connected neurons, but rather travels along the infected cell's neurites and tends to spread more via synaptic contacts than other avenues (Ekstrand *et al.*, 2008).

At the current level of knowledge, the complexity of herpes virus genomes and their varying impacts on the health of different kinds of cells after infection hinder their ability to act as a general tool for studying neural circuits.

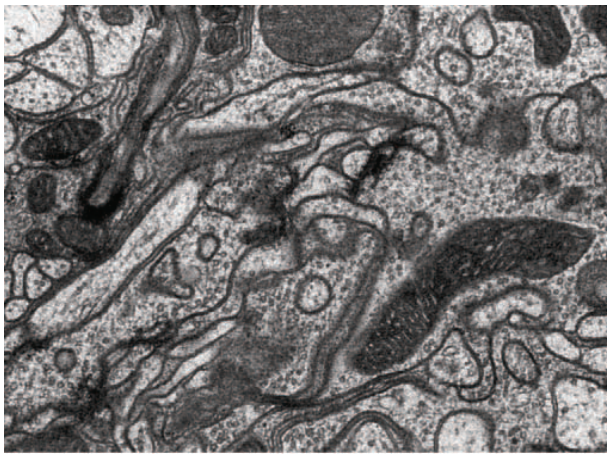
1.2.4 High Resolution Methods

A discussion of tracing techniques would not be complete without briefly mentioning the high resolution methods. Serial section electron microscopy (SSEM) is considered the gold standard in reconstructing small volumes of tissue and holistically determining all synaptic connections (Arenkiel & Ehlers, 2009). This is not a new method. Almost thirty years ago, SSEM was used in a large collaborative effort to comprehensively map the 302 neurons of the *Caenorhabditis elegans* nervous system (White *et al.*, 1986). The concept is simple. First, image the surface of an embedded piece of tissue, then slice ≈10 nanometers off the top of the tissue with a diamond knife or rotary microtome, and proceed to image the next layer. This is iterated through a necessarily small volume of tissue sample and then the images are segmented, registered, aligned, and 3D-reconstructed (Denk & Horstmann, 2004) as shown in Figure 1.3.

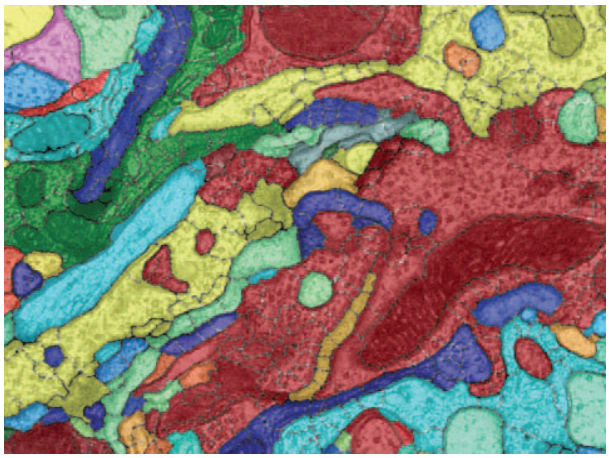
Recently, both the tissue sectioning process and reconstruction algorithms have made great strides forward (Kleinfeld *et al.*, 2011; Briggman & Denk, 2006; Hayworth *et al.*, 2006). The level of ultrastructural detail in the resulting image stack clearly depicts all synaptic connections and thus establishes a comprehensive connectome for the small piece of tissue in question. New developments include crowd sourcing of image annotation and machine learning segmentation approaches to be able to decrease the huge processing times associated with such large stacks of EM slice data (Helmstaedter *et al.*, 2013).

Unfortunately, the method is only feasible for very small sections of tissue, which means that long distance connections are lost. Although pure anatomical information is useful in its own right, SSEM struggles to reveal the strength of the synapses, whether they are excitatory or inhibitory, and which synapses are functionally relevant (Lichtman *et al.*, 2008).

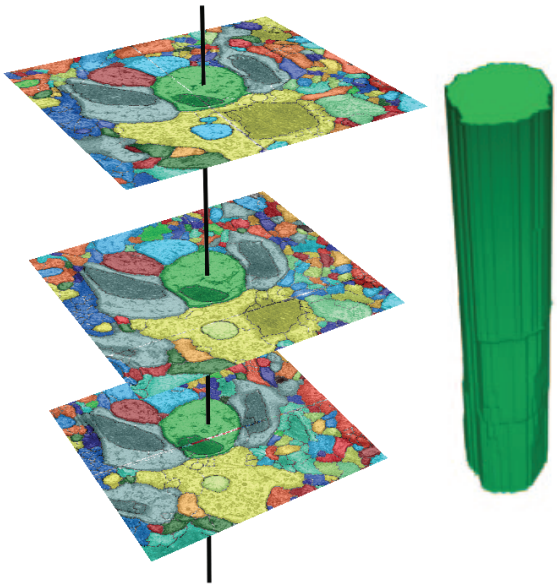
Despite this, SSEM is a very important technique and connectome projects are definitely worth their while. Very recently, an optimised SSEM protocol was used to map the repeating motif in the DM optical medulla, containing roughly 400 neurons and 9000



(a) *A representative electron micrograph taken from a thin slice of brain tissue*



(b) *A segmentation of the electron micrograph shown in Figure 1.3a*



(c) *Registration of several micrographs to render a 3D reconstruction of a neurite*

Figure 1.3: *The segmentation and registration process involved in SSEM reconstruction of a small piece of brain tissue, Reprinted with permission (Takemura* et al.*, 2013)*

connections (Takemura *et al.*, 2013). Clearly, anatomical atlases will be essential in establishing canonical neural circuit architectures, but it will also be useful to establish tracing methods with a more functional spin. This type of functional mapping has been achieved optogenetically (Katzel *et al.*, 2011). What still needs to be developed is a system capable of driving expression of an optogenetic actuator in groups of synaptically connected cells, one of the main goals of this thesis.

1.2.5 A New Approach

In designing a new approach to circuit tracing, it is important to learn from the existing paradigms, attempting to replicate their advantages while eliminating their shortcomings. It is clear from looking at the existing approaches that the best system will be (i) endogenously expressed, (ii) self-amplifying, (iii) technically undemanding and easily reproducible, (iv) able to trace both local and long distance projections, and (v) direction specific. Other factors worth thinking about are whether the information obtained should be purely anatomical or if it is possible to extract anatomical and functional information about synaptic connections simultaneously.

With these points in mind, the overarching goal of this work is to establish a circuit tracing system that hijacks endogenous homeostatic signalling systems (introduced in Section 1.4.2). The paradigm-shift takes a more hands off approach to mapping circuits and rather than introducing exogenous tracing agents, it instead disrupts the normal function of the circuit and optically monitors the homeostatic stabilising responses that are becoming more and more well-understood. A system that can label neurons that are not only genetically related, but architectonically related as well is desperately needed to enhance the already burgeoning field of optogenetics (Miesenboeck & Kevrekidis, 2005). While it might still be a long way away, this thesis takes the initial steps toward establishing such a system. To prototype and test a purely genetic tracing system, the most appropriate model organism is DM.

1.3 DM

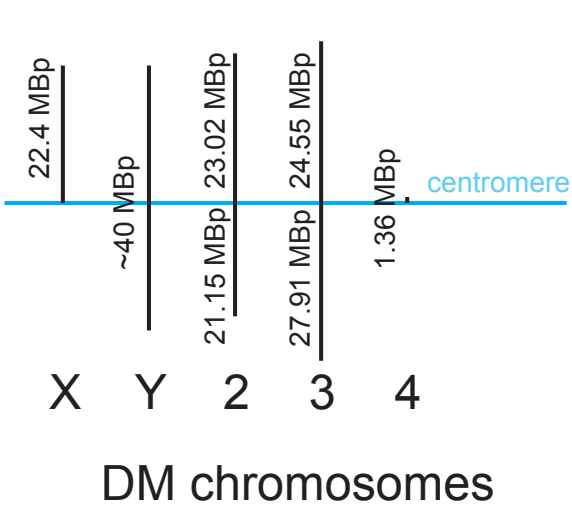
1.3.1 Genetics

What does it mean to have an "advanced genetics toolbox" and why is it important for studying neuroscience? Over the past century, many tools, exclusively available for use in DM, have been developed that make the process of generating and stabilising transgenic strains of animals trivial (Bellen *et al.*, 2010). P-element transgenesis (Rubin &

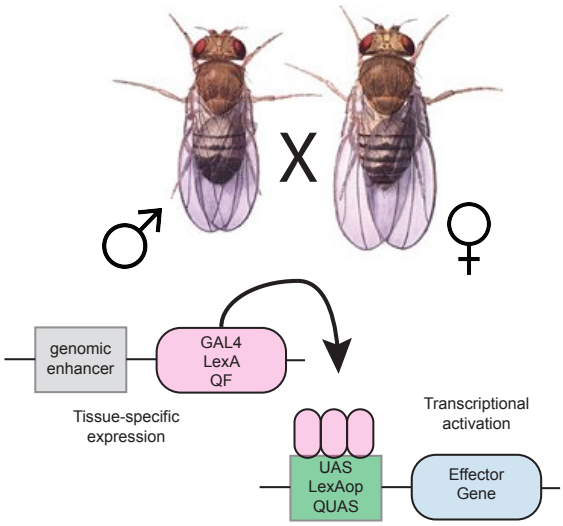
Spradling, 1982), balancer chromosomes (Muller, 1927), and a vast array of highly penetrant genetic/phenotypic markers are just a few of the classical tools. More recently, the development of the GAL4/UAS binary expression system, detailed in Figure 1.4b, reigned in the modern era for the use of DM transgenic strains for neuroscience research (Brand & Perrimon, 1993). GAL4 is a transcription factor found in budding yeast *S. cerevisiae*, and its DNA-binding domain binds to the Upstream Activating Sequence (UAS). Once bound to the UAS, the activation domain of GAL4 recruits RNA polymerase and initiates transcription of the downstream sequence.

Using both site-directed (Bischof *et al.*, 2007) and random insertional (Bier *et al.*, 1989) transgenesis techniques, GAL4 open reading frames have been placed essentially everywhere in the well-documented DM genome (Pfeiffer *et al.*, 2010), which is displayed in Figure 1.4a. Recently, transgenesis upgrades have made it possible to incorporate DNA elements as large as 100 kilobases into the DM genome using Bacterial Artificial Chromosomes (BACs) (Venken *et al.*, 2006). Transgenically inserted GAL4 is then expressed under the control of the local cis-regulatory elements (Bier *et al.*, 1989). This type of system is referred to as an *enhancer trap*. Another common practice is to clone known promoter regions and place them upstream of a GAL4 sequence which is then inserted in a known, well-insulated genomic location (Jenett *et al.*, 2012). When using this approach, it is important to capture the entire relevant promoter region to accurately recapitulate the full expression pattern of the gene of interest, which for large genes with large cis-regulatory domains can require the BACs mentioned above (Venken *et al.*, 2009).

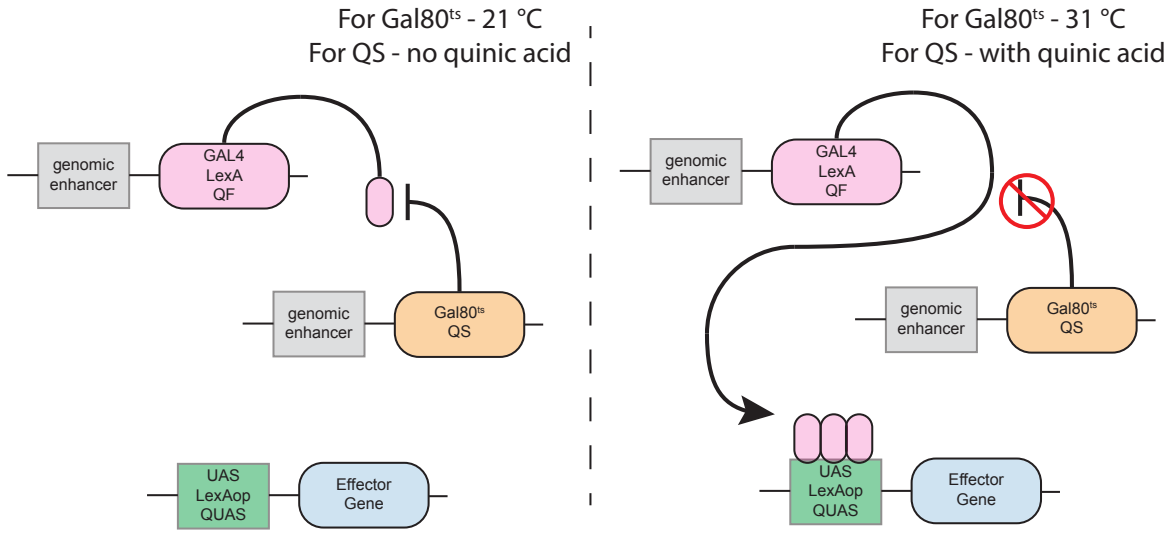
Obviously, GAL4 expression alone is only useful for labelling cells. The other half of the system has to be present in order to express an effector gene in the same tissue-specific pattern as GAL4. UAS-transgenes of all different kinds exist and can be maintained in separate transgenic fly lines from GAL4 drivers to ensure that tissue-specific expression of any gene in any pattern is only one generation of controlled breeding away. Many elaborations of this system have made it the absolute cornerstone of transgenesis in animal models. Three independent binary expression systems now exist, the GAL4/UAS system (Brand & Perrimon, 1993), the LexA/LexAop system (Lai & Lee, 2006), and the QF/QUAS system (Potter & Luo, 2011)). All of them are inducible so that expression can be activated only during time periods of interest, as shown in Figure 1.4c, using either a temperature-sensitive GAL80 ($GAL80^{ts}$) (McGuire *et al.*, 2004) or QS and quinic acid (Potter & Luo, 2011). Split-GAL4 (Luan *et al.*, 2006) lines now exist for expressing a gene only in the intersection where two independent enhancer elements are active, and the MARCM system (Lee & Luo, 1999) has also been established for expressing transgenes in single cell clones. All of these technologies have been developed as derivatives of the original GAL4 system and all provide unique, powerful genetic paradigms to interrogate the function of specific groups of cells.



(a) *The DM genome is approximately 120 MBp and made up of 4 chromosomes. The 4th chromosome is very small and does not recombine (Adams & Celniker, 2000)*



(b) *Binary expression systems make it possible to express a transgene in any available with simple controlled crossing of transgenic lines (Brand & Perrimon, 1993)*



(c) *The inducible expression systems allow expression of a gene to be temporally controlled either using temperature for the GAL4 system or ingestion of quinic acid for the QF system (Potter & Luo, 2011)*

Figure 1.4: *Genetics and binary expression systems available in DM*

1.3.2 Behavioural Repertoire

The DM genetics toolbox would not be useful at all for neuroscience if the animal did not display an interesting and complex array of behaviours. As early as 1905, Frederic Carpenter noticed that when he placed flies on the window sill in glass vials, they accumulated on the upper part of the vial facing the window (Carpenter, 1905). A little over a century later, researchers have carefully dissected the behavioural repertoire of the fruit fly and studied the neural circuitry underlying almost everything the fly can do. Courtship (Greenspan & Ferveur, 2000), olfaction (Vosshall, 2000), circadian rhythm (Konopka & Benzer, 1971), sleep (Cirelli & Bushey, 2008), learning (Busto *et al.*, 2010) and memory (Keene & Waddell, 2007) are just a few examples of the complex behaviours available for study in DM. The behaviours most relevant to this work are olfaction and circadian rhythm.

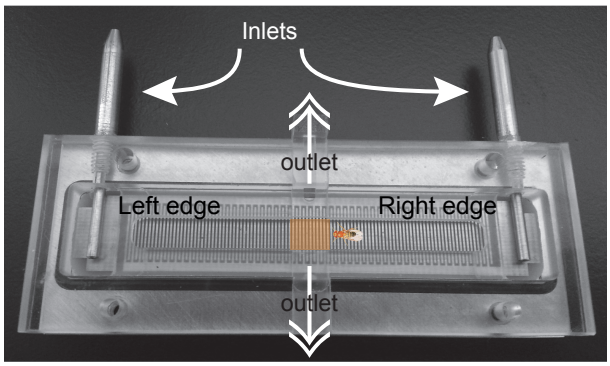
1.3.2.1 Olfaction

Olfaction, or olfactory perception, is fundamental to organismal survival, especially in DM. The fly's ability to use its sense of smell to seek out food (Beshel & Zhong, 2013), avoid danger (Suh *et al.*, 2004), and find mating partners (Kurtovic *et al.*, 2007) has been studied in great detail. Flies are capable of finely discriminating environmental volatile chemicals, associating them with concurrent salient events (Busto *et al.*, 2010), and recalling the associations (Keene & Waddell, 2007). The fundamental process of olfactory perception has mostly been investigated by monitoring either a single fly or a population of flies behavioural response when confronted with olfactory choices in a controlled environment. The neural substrates for this perception will be discussed later.

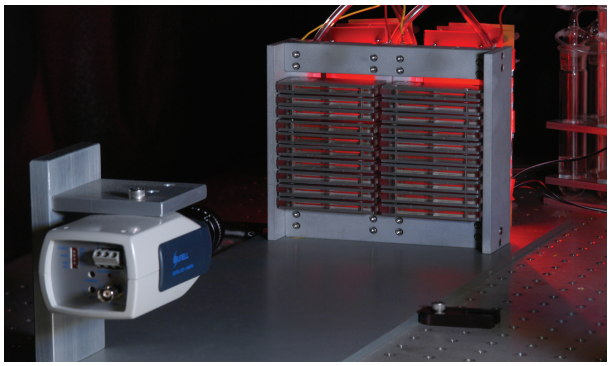
For single fly studies, flies are placed into custom-built experimental chambers like the one shown in Figure 1.5a, allowing different odours to be presented to each side of the chamber independently. The converging odour streams divide the chamber such that half of it has one olfactory quality, and the other half another. The position of the fly is continually monitored within the chamber and many chambers can be monitored simultaneously with systems like the one shown in Figure 1.5b. It is then possible to calculate an odour preference using either the amount of time the fly spends on each side of the chamber or the number of times it enters the centre of the chamber and decides to approach one odour instead of the other (Claridge-Chang *et al.*, 2009). The traces shown in Figure 1.5c illustrate the nature of the positional tracking in these systems.

Experiments measuring absolute odour valence present a pure odour diluted in solvent on one side of the chamber and only the solvent (in most cases mineral oil or water) on the other side, strictly determining whether a fly is attracted or repulsed by a particular

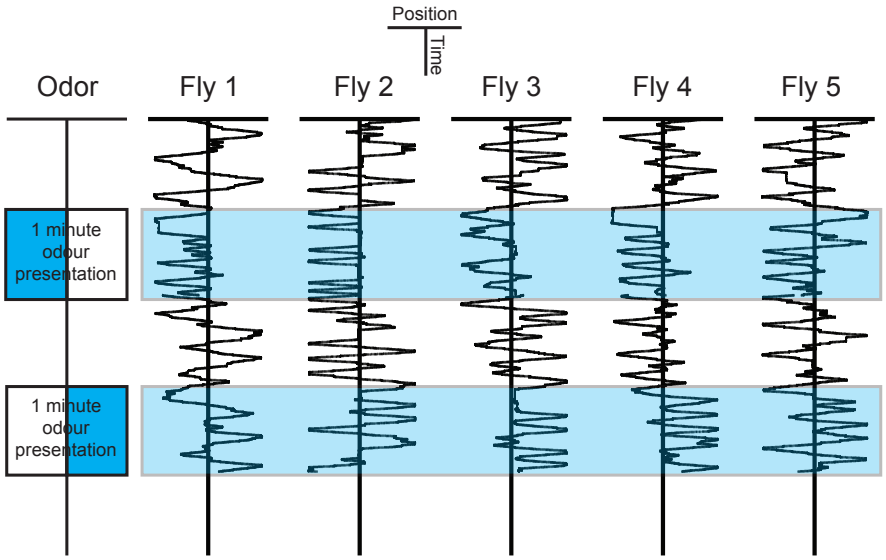
odour quality (Parnas *et al.*, 2013). Discrimination experiments pair odours to determine the fly's relative preference between them (Parnas *et al.*, 2013; DasGupta & Waddell, 2008).



(a) *A chamber used to evaluate a fly's olfactory preferences as revealed through walking behaviour (Claridge-Chang* et al.*, 2009)*



(b) *The high throughput device used to monitor many flies in real time reacting to odour presentations (Claridge-Chang* et al.*, 2009)*



(c) *Single fly walking trajectories for an experiment testing whether an odour is attractive or aversive. Odour is presented to one side of the chamber for one minute, followed by a one minute rest, and then the odour is presented on the opposite side of the chamber for one minute. The flies are attracted in this example.*

Figure 1.5: *The olfactory behaviour system used in this thesis*

An important thing to note is that different experimental setups have given vastly different attraction/repulsion scores. For example, an odour attraction assay using traps (Knaden *et al.*, 2012) instead of odour delivery chambers obtained much higher attraction scores across all odours than a single fly study (Parnas *et al.*, 2013) using the paradigm described above. Also, when the odour preference scores are based on the percentage of a population of flies occupying different sides of a T-maze after a 2 minute odour presentation, odours

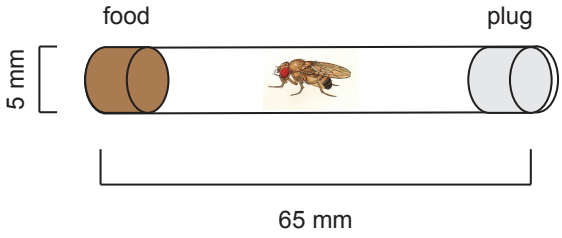
are almost always scored as aversive (Perisse *et al.*, 2013). Therefore, comparison across paradigms should be done with extreme caution.

1.3.2.2 Circadian Rhythm

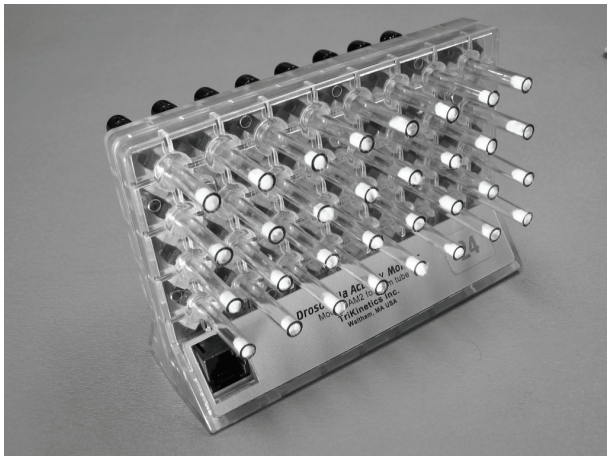
One of the earliest behaviours genetically dissected in flies was the clear daily activity pattern comprised of a bout of locomotor activity (walking) just after lights on, followed by a steady but lower amount of activity in the afternoon, and then a bout of activity just after lights off, as shown in Figure 1.6c (Menegazzi *et al.*, 2012). This daily rhythm persists even when the flies are put in constant darkness, showing the existence of an internal molecular clock (Menegazzi *et al.*, 2012). Many years ago, the Benzer group found several mutants that either cycled through this pattern too slow, too fast, or not at all and then mapped the genetic mutations in these three different types of flies all to the period (*per*) gene (Konopka & Benzer, 1971). Since then, this simple behaviour has been dissected in every which way.

Many DM genes have been shown to be under direct circadian regulation (Ceriani *et al.*, 2002; Claridge-Chang *et al.*, 2001; Lin *et al.*, 2002; McDonald & Rosbash, 2001; Ueda *et al.*, 2002; Hughes *et al.*, 2012) and the neural circuits involved are also well-characterised. The core molecular clock involves concerted cycling of several interacting transcription factors in a negative feedback loop. The basic mechanism of negative feedback is that after the *per* gene is translated and dimerises with another clock component called timeless (*tim*), it enters the nucleus and represses further *per* transcription until the other clock genes (cryptochrome *cry* and doubletime *dt*) clear the per/tim protein complex from the nucleus (Price *et al.*, 1998; Emery *et al.*, 1998). Then the process begins again.

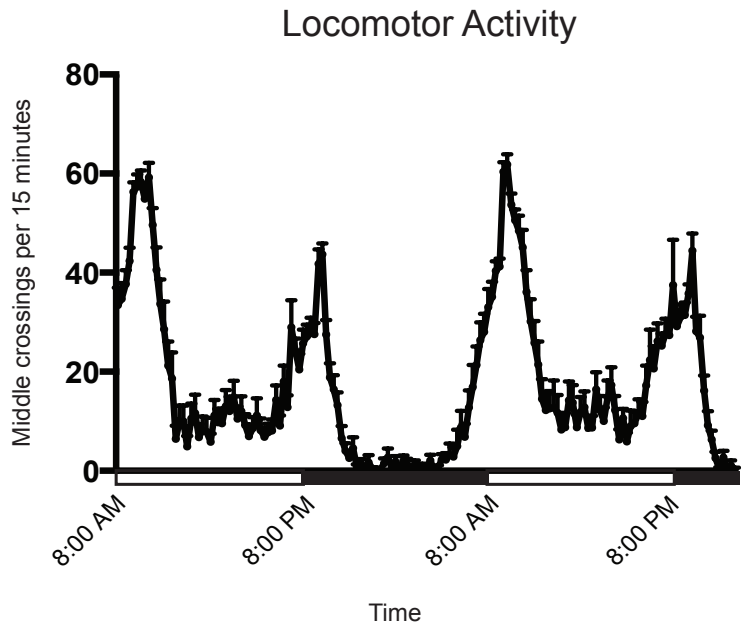
The great detail with which the circadian cycle is understood is in large part due to the elegant simplicity of the behavioural monitoring apparatus. Flies are placed in thin walled glass tubes with food on one end and a plug on the other as shown in Figure 1.6a, which are then threaded through a manifold that fires an infrared beam directed through the midpoint of the tube and counts each time a fly crosses the beam (Pfeiffenberger *et al.*, 2010). The manifold or *monitor* is pictured in Figure 1.6b. Beam crossings are tallied every minute, and when the fly is active, the beam crossings per minute increase. This setup is easily multiplexed so that hundreds of single flies can be monitored simultaneously. Everything from sleep (Donlea *et al.*, 2011) to the effects of social experience (Ganguly-Fitzgerald *et al.*, 2006) can been studied using this technology.



(a) *A glass tube used to monitor a fly's walking activity over several days*



(b) *The device used to monitor many flies at once over several days*



(c) *Daily activity patterns of wild type flies in a 12:12 LD cycle with lights on at 8:00 AM and lights off at 8:00 PM, middle crossings counted in 15 minute bins and displayed with standard errors for n=28 flies*

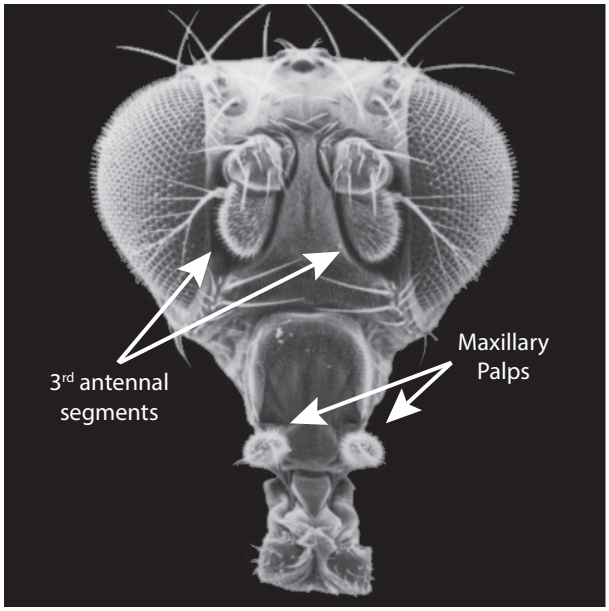
Figure 1.6: *The high-throughput activity monitoring system to record circadian locomotory rhythms*

1.3.3 Anatomy

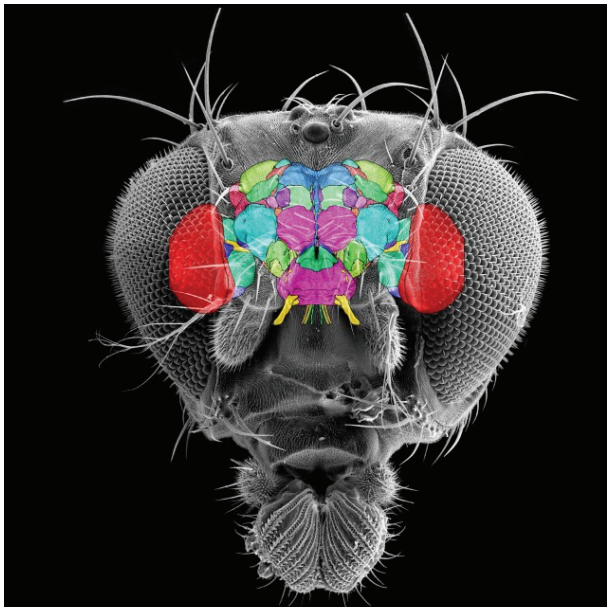
The well-characterised anatomy of the DM adult central nervous system (CNS) shown in Figure 1.7b is the third pillar of the DM model system. This thesis focuses on the supraesophageal ganglion of the adult brain, which is a paired bilateral structure subdivided into three fused neuromeres: the protocerebrum, the deuterocerebrum, and the ventral tritocerebrum (Power, 1943). The region of particular interest is the antennal lobe (AL), highlighted in Figure 1.7c which resides in the deuterocerebrum and has many similarities with the mammalian olfactory bulb (Stocker *et al.*, 1990). The AL neuropils are the first relay station for incoming olfactory information, and they contain the axon terminals of the olfactory receptor neurons (ORNs), shown in Figure 1.7c, the dendritic endings of the olfactory projection neurons (PNs), shown in Figure 1.7d, and the many neurites of both excitatory (eLNs) and inhibitory (iLNs) local neurons (Stocker *et al.*, 1990; Shang *et al.*, 2007). Evidence also exists for both peptidergic and aminergic inputs to the ALs coming from different parts of the protocerebrum (Carlsson *et al.*, 2010; Dacks *et al.*, 2009).

Beginning at the periphery, volatile chemicals are detected by the bipolar, cholinergic ORNs, whose somata are housed in the third antennal segments and maxillary palps, shown in Figure 1.7a (Vosshall & Stocker, 2007). The dendrites of the ORNs interleave themselves into small hair-like structures called sensilla that are easily visible using a regular dissection microscope (Laissue & Vosshall, 2008). Importantly, the anatomical separability of both the third antennal segments and the maxillary palps allows easy isolation of these structures from the rest of the animal by careful dissection. The third antennal segment is the predominant olfactory organ containing approximately 1200 neurons as well as other support cells (Vosshall & Stocker, 2007). ORNs make up about 30% of the third antennal segment tissue (Vosshall & Stocker, 2007). The population of ORNs is made up of many different cellular classes, with each ORN class expressing a different combination of odourant receptors (ORs) (Vosshall *et al.*, 1999). The canonical ORN classes express the co-receptor Or83b (also called Orco) in addition to one or a few of the 50-60 known ORs, resulting in a unique chemical receptive field (Hallem & Carlson, 2006).

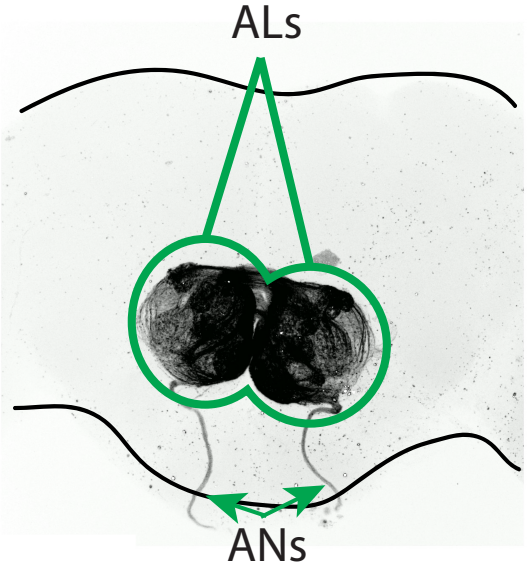
All of the ORN axons converge at the proximal side of the third antennal segment to form the antennal nerve (AN), the largest nerve fibre in the entire DM central brain (Power, 1943). The AN is enclosed in a nucleated envelope which is continuous with that of the central brain, which also contains the axons of other neurons housed in the Johnson's Organ (JO) in the second antennal segment (Power, 1943). The ORN axons run within the AN until they reach the ventrolateral side of the AL, where they flare out bilaterally just in front of the ≈50 AL glomeruli (Power, 1943). Finally, ORNs that are of the same



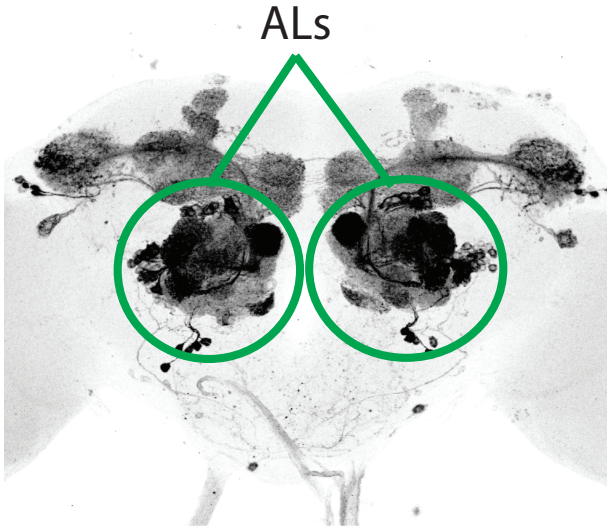
(a) *External structures of the olfactory system*



(b) *The adult brain as it rests inside the head*



(c) *The axonal projections of the ORNs (Or83b-LexA driven LexAop-myr-mCherry)*



(d) *Canonical PNs expressing GFP (GH146-GAL4 driven UAS-mCD8-GFP)*

Figure 1.7: *The DM adult brain and olfactory system, AL - antennal lobe, AN - antennal nerve*

class express the same OR set (there are 10-100 neurons of each ORN class), and all ORNs expressing the same OR set terminate in the same AL glomerulus (Vosshall *et al.*, 1999).

The somata of the secondary PNs and the various eLN and iLN cells reside in three separate clusters on the periphery of the AL; the anterodorsal cluster, the lateral cluster, and the ventral cluster (Lai *et al.*, 2008). The anterodorsal cluster contains ≈75 uniglomerular, ipsilaterally projecting, cholinergic PNs which send axons through what is known as the inner antennocerebral tract (iACT) to the higher areas of the protocerebrum called the lateral horn (LH) and the mushroom body (MB) (Lai *et al.*, 2008). These are considered to be canonical PNs. The lateral cluster is a diverse population of ≈200 cells containing eLNs (cholinergic) and iLNs (GABAergic and glutamatergic), canonical PNs (≈35 cells in this cluster), atypical multiglomerular PNs and even neurons sending projections to other parts of the brain outside of the AL (Lai *et al.*, 2008). The ventral cluster is comprised of ≈50 PNs that project uniformly through the middle antennocerebral tract (mACT) (Lai *et al.*, 2008). These are both canonical and atypical PNs that can be either cholinergic or GABAergic. The mACT bypasses the MB and projects directly to the LH (Power, 1943).

Within the well-characterised structure of the AL, there are many useful GAL4 driver lines that give access to subsets of the different neuronal classes. Or83b-GAL4 and Or83b-LexA lines, which were used in Figure 1.7c, grant access to the vast majority of ORNs (Larsson *et al.*, 2004). Individual OR drivers have also been created for every characterised ORN class (Couto *et al.*, 2005). Between GH146-GAL4, pictured in Figure 1.7d, Mz19-GAL4, NP3062-GAL4, and R81E09-GAL4 it is possible to access both large and small subsets of canonical PNs (Brand & Perrimon, 1993; Manseau *et al.*, 1997; Motojiro & Ito, 2000; Jenett *et al.*, 2012). There are also lines granting access to iLNs (GH298-GAL4), eLNs (kras-GAL4 or c305-GAL4), and even non-canonical inhibitory PNs from the ventral AL cluster (Mz699-GAL4) (Brand & Perrimon, 1993; Dubnau *et al.*, 2003; Manseau *et al.*, 1997). The DM genetic systems described earlier have shown their power by allowing detailed functional dissection of the flow of olfactory information and the role of different neuronal classes in the AL.

1.3.4 The AL Circuit and the ORN/PN Synapse

The propagation of olfactory information at both the primary (ORN) and secondary (PN) levels is very well understood in DM, as documented in Figure 1.8. At the ORN level, huge data sets exist detailing the electrophysiological responses of about half of the ORN types to a comprehensive set of chemical stimuli (Hallem & Carlson, 2006). This data shows that the interaction between ORNs and environmental volatile chemicals is

complex, with most ORN types responding to several chemicals and most chemicals activating several ORN types. ORNs can be broadly tuned to many molecules, narrowly tuned to a single molecule, or anywhere in between. Different ORN types have different levels of basal activity, with volatile chemicals able to have both excitatory and inhibitory effects. Generally, increasing the concentration of the chemical ligand increases the effect on the basal firing rate (Wilson, 2013). Higher odour concentrations bring about broader responses across ORN types. The temporal dynamics of ORN firing is also complex and varied depending on ORN type, ligand, and concentration (Wilson, 2013). Ephaptic non-synaptic effects have recently been demonstrated in different classes of ORNs sharing the same sensillium, so mixtures of odours can already interfere with each other's representations at the primary stage of olfactory processing (Su *et al.*, 2012).

ORNs transmit their sensory signals through excitatory, cholinergic synapses to a complex network of PNs, iLNs, and eLNs (Wilson, 2013). Activated iLNs directly inhibit the presynaptic ORNs (Olsen & Wilson, 2008), as well as their neighbouring PNs and eLNs. The lateral inhibition provided by the iLNs makes PN responses transient and also increases the dynamic range of odour that PNs can detect (Wilson & Laurent, 2005). PNs are thus more broadly tuned than ORNs and it takes much higher odour concentrations to saturate their firing (Wilson & Laurent, 2005). The eLNs spread excitation within the AL by activating PNs via direct electrical connections (gap junctions) and at the same time inhibit PN activity indirectly by activating iLNs through chemical synapses (Yaksi & Wilson, 2010). PNs show highly correlated activity within a glomerulus because of gap junctions (Kazama & Wilson, 2009), and also form chemical synapses onto both eLNs and iLNs. Canonical, uniglomerular PNs receive direct input only from a single class of ORN. ORNs are the primary source of excitatory input to the circuitry of the AL. When ORN axons are cut, the basal activity in the circuit is abolished (Kazama & Wilson, 2009).

Looking specifically at the ORN/PN synapse, a few characteristics stand out. First, larger PN dendritic arbors draw larger *unitary* excitatory postsynaptic currents (uEPSCs), *unitary* meaning the result of a single presynaptic spike (Kazama & Wilson, 2008). Detailed quantal analysis revealed that this is because larger glomeruli have a correspondingly larger number of ORN presynaptic release sites (Kazama & Wilson, 2008). This has also been supported by direct synaptic counting in the AL (Mosca, 2013). The larger currents, however, result in the same level of membrane depolarisation in PNs with different size dendritic arbors because increasing the surface area of the dendritic arbor decreases the input resistance (Kazama & Wilson, 2008), and thus the neuron requires more current to obtain a certain level of depolarisation. Decreasing the input resistance transgenically by overexpressing the Kir2.1 potassium channel in PNs postsynaptic to a specific glomerulus also resulted in an increased uEPSC (Kazama & Wilson, 2008). This suggests it is not

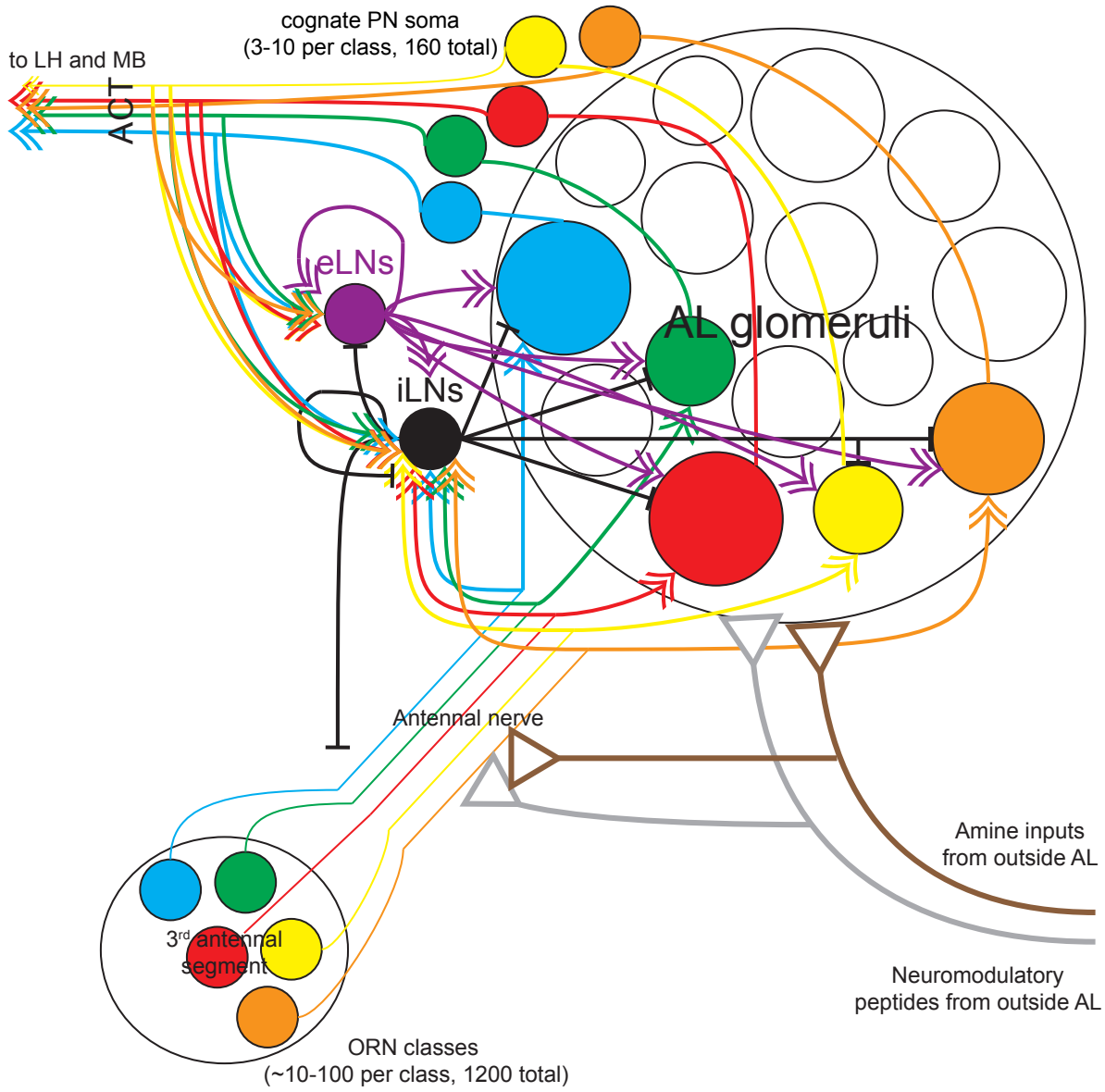


Figure 1.8: *A full, quasi-anatomical circuit diagram for the antennal lobe circuit, cognate ORN/PN pairs are colour coded, double arrows mean excitatory synapses, flat lines mean inhibition, and triangles mean neuromodulatory synapses, abbreviations - ACT - antennocerebral tract, LH - lateral horn, MB - mushroom bodies*

the size that matters, but the intrinsic properties of the PNs that are matched by their presynaptic inputs. Another report showed that potassium channel mediated silencing in PNs also results in an increased number of PN release sites in the mushroom body calyx, the third level of olfactory information processing (Kremer *et al.*, 2010). This suggests there is homeostatic plasticity at many levels in the olfactory circuit.

Other studies have shown that chronic ORN activation by raising flies in odour-concentrated environments can have an effect on the size of AL glomeruli (Sachse *et al.*, 2007). Along with the glomerular volume changes, the odour-raised flies are behaviourally desensitised to the conditioned odour. It seems this change is attributable to increased iLN responses, and not a product of ORN output adjustments (Sachse *et al.*, 2007). Also, in developmental terms it has been shown that activity is not required for ORNs and PNs to form the delicate AL glomerular structure (Jefferis *et al.*, 2001), but genetic ablation of certain cell-adhesion molecules can cause ORN-PN glomerular mismatching (Hong *et al.*, 2012).

Taken together, these data suggest that at the ORN/PN synapse after developmental circuit formation, homeostatic cooperation between PNs and ORNs sets the quantal levels in each glomerulus. The amount of synaptic drive provided by ORN inputs (presynaptic terminals) seems to be precisely set to the input needs of the postsynaptic PNs in a glomerulus-specific manner (Kazama & Wilson, 2008; Mosca, 2013). Homeostatic systems like this have been widely reported in neural circuits and provide a way to relate neurons that might not be genetically similar (Miesenboeck & Kevrekidis, 2005), but are connected. The ORN/PN synapse is thus an ideal place to study homeostatic interactions between pre- and postsynaptic neurons because the presynaptic cells are anatomically separable (they are in the third antennal segment), and the postsynaptic cells are genetically accessible (with GH146-GAL4).

1.4 Homeostasis and Plasticity

Claude Bernard originated the concept of homeostasis when he used the phrase *milieu intérieur* to describe a systemic form of regulation precisely maintaining certain important parameters within the internal environment of the body (Cooper, 2008). An important distinction, especially when talking specifically about the nervous system, is to clearly separate the *homeostatic processes* from what is known as *plasticity*. Homeostatic processes maintain *stability*. Plastic processes usually follow competitive or correlation-based rules and promote *change*, for example the strengthening or weakening of a synaptic connection concomitant with learning (Turrigiano & Nelson, 2004). In a healthy nervous system, these two processes are in constant flux and often at odds with each other. It might be helpful to picture a tug-of-war with plastic forces pulling the system in one

direction and homeostatic processes pulling it in the other, both keeping the other in check to ensure whatever net changes get consolidated are constructive for the system as a whole.

1.4.1 Plasticity in the Nervous System

The best way to introduce the homeostatic processes in the nervous system is to first establish precisely what they are not. When any neuroscientist thinks about plasticity, the first terms that come to mind are long term potentiation (LTP) and long term depression (LTD). These processes have received a lot of attention because of their purported role in adaptive behaviours, learning, and memory (Bliss & Collingridge, 1993). LTP and LTD are manifested as changes in a postsynaptic neuron's sensitivity to neurotransmitter presentation at specific synapses (Bliss & Lomo, 1973). Most of the LTP and LTD research has focused on postsynaptic modifications as a result of correlation-based firing rules (Luscher *et al.*, 2000). The known mechanisms are focused on the composition and physical properties of postsynaptic receptors, whether having to do with phosphorylation states, membrane localisation and receptor density, or subunit composition (Luscher *et al.*, 2000). Coincidence detectors like N-methyl-D-aspartate (NMDA) type glutamate receptors report correlated firing of both presynaptic and postsynaptic neurons, and gate their channels only when both neurons fire in a particular fashion. The mechanism for this is that the receptors depend on both the binding of neurotransmitter AND depolarisation in order to gate their channels (Mg^{2+} blocks the channel if there is no depolarisation) (Luscher *et al.*, 2000). Large influxes of calcium flow through the open channel pore into the postsynaptic cell and initiate the strengthening or weakening of synaptic connections through many different biochemical pathways (Glanzman, 2010).

Alternatively, plasticity can also be present on the presynaptic side. Kandel's Nobel prize-winning work in *Aplysia* showed that non-associative synaptic depression and facilitation are attributable to respective decreases or increases in presynaptic transmitter release (Castellucci *et al.*, 1970). For facilitation, external neuromodulatory neurons are recruited, and the binding of serotonin to G protein-coupled receptors (GPCRs) in the presynaptic membrane of the sensory neuron results in cyclic adenosine monophosphate (cAMP) production and protein kinase A (PKA) activation, eventually leading to increased vesicular release (Castellucci *et al.*, 1970). Other pathways like the protein kinase C (PKC) pathway have also been shown to be involved (Royer *et al.*, 2000). Ultimately, facilitative plasticity in the presynaptic neuron, dependent on heterosynaptic input, results in an increase in the number of release sites, enhanced global excitability, as well as increased mobilisation of synaptic vesicles in the presynaptic neuron, leading to the adaptive sensitisation of the synapse (Glanzman, 2010).

Where both of these plasticity examples overlap is that they are derived from correlated firing. For LTP or LTD, presynaptic and postsynaptic neurons must fire in a certain order to strengthen or weaken the synapse. For sensitisation, the neuromodulatory neuron must fire along with the presynaptic neuron to strengthen the synapse. In contrast, homeostatic behaviour, instead of being based on the correlated firing of two or more neurons, comes from the intrinsic programming of neurons to exhibit certain defined levels of activity based on cellular identity. Exactly what neurons do to maintain their genetically programmed physiological characteristics is the interest of homeostatic research.

1.4.2 Homeostatic Activity in the Nervous System

1.4.2.1 Intrinsic Activity Maintenance

The most straightforward examples of homeostatic compensatory behaviour come from studies on cultured cortical networks. After a few days of development inside a dish, the networks, made up of excitatory pyramidal neurons and inhibitory interneurons, display spontaneous activity that can be modulated pharmacologically for defined periods of time by washing in toxins and then washing them out (Turrigiano & Nelson, 2004). When global firing activity is pharmacologically blocked by treatment with tetrodotoxin (TTX) for 48 hours and the toxin is subsequently washed out, a pronounced increase in pyramidal neuron miniature excitatory postsynaptic current (mEPSC) amplitudes is observed (Turrigiano *et al.*, 1998). The mEPSCs are scaled to approximately twice what they were before the treatment (Turrigiano *et al.*, 1998). This experiment is illustrated in Figure 1.9a. The converse is also true. When firing activity is elevated by bicuculline treatment to block $GABA_A$ receptor mediated inhibition for 48 hours, the mEPSC amplitudes fall to approximately 70% of the before treatment values (Turrigiano *et al.*, 1998). Global excitability is therefore under bidirectional homeostatic control.

A similar homeostatic behaviour has been observed when transfecting Kir2.1, an inwardly-rectifying potassium channel, into a single cell within a cultured network. A homeostatic response was not detected when transfecting Kir2.1 at a developmentally early stage before synapse formation (Burrone *et al.*, 2002). In this case, neural silencing with Kir2.1 actually decreased the number of synapses that formed, the mEPSC frequency, as well as the evoked postsynaptic current amplitudes (Burrone *et al.*, 2002). However, transfection of Kir2.1 *after* synapses were allowed to form resulted in a pronounced increase in the mEPSC amplitude, synapse density, and synapse size, successfully returning firing rates to wild type levels (Burrone *et al.*, 2002). Based on this result, it is clear that homeostatic synaptic scaling is (i) sensitive to developmental timing and (ii) dependent on the existence of synaptic sites. Additionally, it clearly means that the activity of single neurons regulates homeostatic maintenance, not just network activity levels.

In another experimental system, the light-gated cation channel channelrhodopsin-2 (ChR2) was used to drive optogenetic excitation of pyramidal neurons in organotypic hippocampal slices from rodents (Goold & Nicoll, 2010). Overexcitation of the neurons led to compensatory postsynaptic depression of α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor and NMDA receptor responses through different CaM kinase pathways (Goold & Nicoll, 2010). Thus, receptor sensitivities can be cell-autonomously modulated by intrinsic changes in firing rate. In primary cultured DM neurons, synaptic blockade with the neurotoxin curare results in an increase in expression of the nicotinic acetylcholine receptor $D\alpha7$, which in turn causes an upregulation in both transcription and translation of Shal K^+ channels to stabilise synaptic potentials (Ping & Tsunoda, 2012).

Excitability levels are thus tightly regulated in many different types of neurons in many different species. Taking everything together, it is clear that neurons monitor their own intrinsic activity. These intrinsic homeostatic systems must be built on fundamental principles of feedback that allow neurons to maintain identity-specific levels of activity in an unstable environment. From control theory, the design features of such a system should include (i) a set-point or target state to maintain, (ii) activity sensors, (iii) error signals, and (iv) regulatory feedback (Davis, 2006). A block diagram with each of these components is shown in Figure 1.9c. To date, only some of the molecular players in these functional modules have been identified. Difficulties in understanding the underlying processes are evident in that it is not known exactly how neurons select a target activity level; how this choice is related to cellular identity; how the sensor signal is fed back and processed by the system; and what molecular programs the neuron executes to preserve its target excitability state.

What can be said about this process is that intrinsic homeostatic maintenance has a vast parameter space. In exploring this parameter space, it has been shown that drastically different ion channel densities can actually reconstitute very similar electrophysiological characteristics in given neuronal subtypes (Marder & Goaillard, 2006). Even neurons of the same type with identical firing behaviour taken from paired ganglionic hemispheres of the same animal can show considerable variation in individual ion channel conductances and expression levels (Schulz *et al.*, 2006). These results highlight two very important ideas: (i) it is not a target level of transcription of a particular gene that neurons are trying to maintain, but instead a global functional state; and (ii) there are many different parameter sets capable of producing the same functional output (Marder & Goaillard, 2006).

Thus, it seems neurons, rather than tightly regulating expression levels of single genes, are probably keeping a multitude of biophysical parameters within broad, acceptable functional ranges. When the excitability of a neuron becomes inadequate because of

environmental changes, one or many biochemical pathways can be regulated to alter a neuron's level of excitability. It follows that there are many solutions to the same problem, and the nearest solution for one neuron in parameter space might not be the nearest solution for a neighbouring neuron of the same cell type. It is then clear that monitoring expression levels of single genes is not appropriate for searching out the molecular players in intrinsic homeostatic maintenance. Measuring the expression levels of all genes is more appropriate.

1.4.2.2 Transsynaptic Homeostatic Maintenance

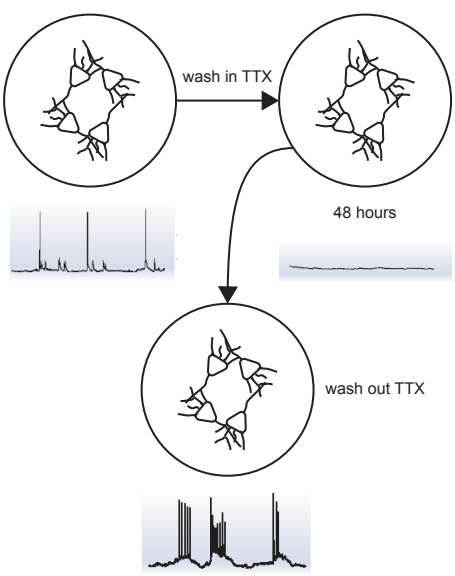
Zooming out to the circuit level, just as individual neurons monitor and stabilise their own intrinsic levels of activity, it has also been shown that synaptically-connected neurons help each other achieve this stability (Davis, 2006). Presumably, neurons do this by sending messages to their presynaptic partners in the form of retrograde signals. Upon receiving the messages, the presynaptic neurons can activate programs to increase or decrease synaptic drive, thus assisting the postsynaptic neuron achieve a desired level of activity. This type of transsynaptic compensation, often called *synaptic homeostatic compensation*, is another well-established property of neuronal networks (Davis, 2006).

Transsynaptic homeostatic signalling has mostly been studied at the neuromuscular junction (NMJ). In the mouse, mutants with reduced levels of postsynaptic acetylcholine receptors (AChRs) in their muscles compensate for this by increasing the number of quanta released per motorneuron impulse (Sandrock *et al.*, 1997). This reconstitutes wild type amplitudes for evoked postsynaptic potentials (EPSPs) despite lower levels of receptor expression. This same phenomenon has also been observed at the DM third instar larval NMJ (Petersen *et al.*, 1997), which is glutamatergic instead of cholinergic. The experiments have gone further at the DM third instar larval NMJ, verifying that receptor mutants (Petersen *et al.*, 1997), acute pharmacological receptor blockade (Frank *et al.*, 2006), and even inducible expression of Kir2.1 in postsynaptic muscle (Paradis *et al.*, 2001) all result in homeostatic increases in the presynaptic release probability. The release probability has been shown to be modulated by changes in calcium influx per presynaptic spike (Zhao *et al.*, 2011) and changes in the size of the release ready synaptic vesicle pool (Weyhersmuller *et al.*, 2011). This kind of homeostatic behaviour is represented in Figure 1.9b. Additionally, presynaptic silencing (Crawford *et al.*, 2012) and altered vesicular filling (De Gois *et al.*, 2005) have been suggested to play roles in homeostatic modulation of presynaptic relsease.

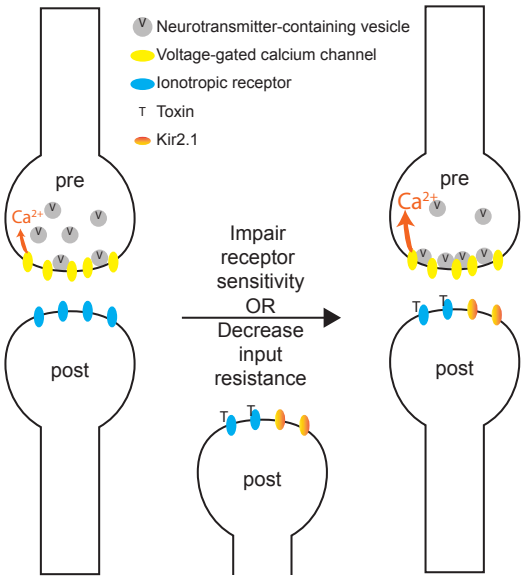
Many single genes have been identified that confer homeostatic competence to the system, but none have quite pinpointed which molecules directly mediate the retrograde signal.

The lone example of a retrograde signalling pathway exerting a direct impact on presynaptic release is the Ephrin receptor pathway, which acts in concert with ephexin and Cdc42 on presynaptic calcium channels (Frank *et al.*, 2009). Examples of single genes necessary for synaptic homeostatic competence at the DM larval NMJ are the $Ca_v2.1$-type calcium channel *cacophony* (Frank *et al.*, 2006), the RhoGAP *crossveinless-c* (Pilgram *et al.*, 2011), the transcription factor *gooseberry* (Marie *et al.*, 2010), the bone morphogenetic protein (BMP) signalling cascade (Goold & Davis, 2007), the GTPase Rab3 and its accessory proteins (Muller *et al.*, 2011, 2012), and the ENaC channels ppk11 and ppk16 (Younger *et al.*, 2013). However, very few of these genes, besides ppk11 and ppk16, have been shown to be *transcriptionally* regulated as a result of the postsynaptic perturbation.

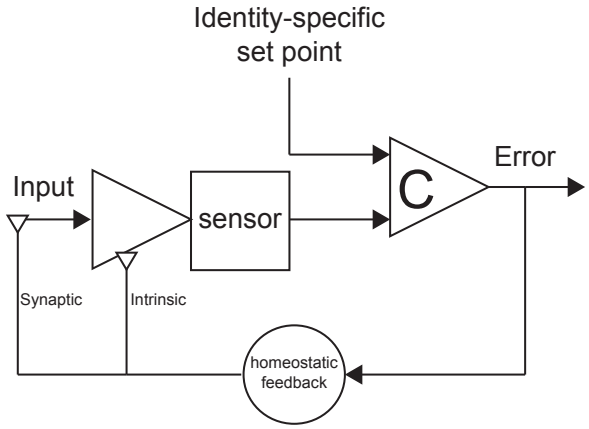
As mentioned in Section 1.3.4, synaptic homeostatic behaviour has also been observed at a central synapse in DM (Kazama & Wilson, 2008), providing a unique opportunity to study synaptic homeostatic processes between two neurons *in vivo*. At the ORN/PN synapse, PNs with larger dendritic arbors draw higher uEPSCs (Kazama & Wilson, 2008). These higher currents are due to homeostatic matching of the presynaptic neuron release characteristics to the postsynaptic demand for synaptic drive, manifested in an increase in the number of presynaptic release sites (Kazama & Wilson, 2008; Mosca, 2013). If this is indeed happening, what molecules mediate the retrograde signal to inform the presynaptic neurons how many presynaptic sites are necessary? What genetic modules in the presynaptic neuron are responding to the signal in order to provide the suitable levels of synaptic drive? Presently, the genetic tools in DM are adequate to answer some of these questions, and that is one of the main goals of this thesis. The experimental system will involve inducible expression of a neuronal silencer in the postsynaptic PNs after synapse formation, and monitor *transcriptome*-wide gene expression in the presynaptic ORNs. An immediate application of understanding which transcriptional modules might be activated in presynaptic neurons is a new kind of retrograde tracing system. If postsynaptic neurons can regulate presynaptic transcription, this signalling system can be exploited to drive expression of simple reporters (eGFP), optical activity reporters (GCaMP), or optogenetic actuators (ChR2).



(a) *Schematic demonstrating that neurons increase excitability after relief from prolonged activity blockade*



(b) *Schematic demonstrating how presynaptic neurons increase synaptic drive in response to either the blockade of postsynaptic receptors or postsynaptic electrical silencing with Kir2.1 by altering calcium influxes and the size of the release ready vesicle pool*



(c) *A block diagram showing the main components of neuronal homeostatic systems, homeostatic feedback can both alter a cell's intrinsic excitability and change the release properties of presynaptic partners, open triangles at the end of lines signify a change in the properties*

Figure 1.9: *Homeostatic systems in neurons*

1.5 Transcriptome Sequencing

Not so long ago, many scientists believed that a fully sequenced genome would provide the answers to almost all of their questions about health and disease. A few years later, it was realised that the genome is only the tip of the iceberg. The dynamic state of the genome, including the structural state of the DNA, the epigenetics of the cell, and transcriptional and posttranscriptional RNA modifications, has come into focus (Wang *et al.*, 2009). At the forefront of this research in genome dynamics, transcriptomics studies have yielded huge amounts of information and demonstrated just how complex nature can be in genome elaboration. In most experimental organisms in which the genome is sequenced, huge catalogs of coding and non-coding RNA species, small RNAs, microRNAs, promoter-associated RNAs, and other transcribed genetic elements have been constructed and are now publicly accessible (Consortium, 2004).

How exactly the transcriptome is measured has changed drastically in the past decade. The first generation serial analysis of gene expression (SAGE) relied on standard reverse transcription and cloning techniques, followed by Sanger sequencing of the cloned cDNA fragments (Velculescu *et al.*, 1995). These low-throughput methods were very time consuming and laborious. Next, methods were introduced that relied mostly on hybridisation-based microarray technologies (Clark *et al.*, 2002). Fluorescent cDNA's were created and hybridised to chips engineered with many different prefabricated oligonucleotide sequences arranged in a two dimensional matrix. The fluorescence readout of each spot on the chip was used to ascertain a relative expression level. Although microarray experiments did provide informative data, the limited dynamic range, inability to detect novel transcripts, and difficulties comparing results across different experiments meant new methods were required (Wang *et al.*, 2009).

The discovery and improvement of next-generation sequencing (NGS) technologies rectified many of those problems. The continually falling costs of high-throughput, massively parallel sequencing approaches have effectively rendered microarrays obsolete. Increasingly longer read lengths, lower error rates and the ability to sequence from both ends of a single molecule (paired-end reads) have greatly improved the quality and reproducibility of NGS technologies for transcriptomics (Trapnell *et al.*, 2013). Crucially for this thesis, special amplification-based technologies have also been developed for sequencing from tiny starting amounts of total RNA in a reproducible fashion (Ramskold *et al.*, 2012). Several studies have now shown that it is possible to extract total RNA from less than a few hundred cells and carry out robust transcriptomic and genomic analyses based on data provided by NGS pipelines (Perrat *et al.*, 2013; Hughes *et al.*, 2012; Sasagawa *et al.*, 2013).

In DM specifically, the developmental transcriptome at almost every stage (Graveley *et al.*, 2011), tissue-specific transcriptomes (Hughes *et al.*, 2012), and even transcript expression level changes based on several different physiological states (starvation (Harbison *et al.*, 2005), sleep deprivation (Zhou *et al.*, 2012), circadian clock (Hughes *et al.*, 2012)) have all been published. The time has come to use this burgeoning technology to answer more directed physiological questions in specific cell populations. The success of NGS approaches combined with the powerful genetics toolbox in DM make it an ideal system to screen gene expression in a discrete population of cells.

Different types of transcriptome sequencing experiments can provide different levels of detail into the transcriptional landscape. Standard RNA sequencing (RNA-seq) and 3′ tag digital gene expression (3′ tag DGE) are the two types of experiment carried out in this thesis and will be described in more detail.

1.5.1 Standard RNA-seq

There are many platforms available for standard RNA-seq, and most are built on similar principles. Briefly, RNA-seq is carried out by generating a library of cDNA fragments from a starting sample of RNA using one of many reverse transcription (RT) approaches (Wang *et al.*, 2009). After the RT, cDNA is fragmented either enzymatically or mechanically and size selected (Harbers & Kahl, 2012). Adaptors are ligated to both sides of the selected cDNA fragments allowing them to be manipulated and attached to the sequencing device, and either directly sequenced in a high throughput manner or amplified first and then sequenced for cases with very low amounts of starting material (Wang *et al.*, 2009). The cDNA molecules can either be sequenced from one side (single-end sequencing) or both sides (paired-end sequencing). Paired-end reads are very helpful in eliminating mapping ambiguities at later stages in data processing. Reads can be anywhere between 50-1000 bp depending on the NGS platform used (Harbers & Kahl, 2012). In this thesis, all reads were 100 bp long and sequenced using Illumina HiSeq2000 technology.

The raw data, comprised of sequences and their corresponding quality scores, are trimmed of low quality sequences and known contaminants (adapters and primers), and then mapped to a reference genome and/or reference transcripts using one of the many available software suites. The end product is a comprehensive transcriptional map including both the structure of the transcripts and the expression levels of each gene in the form of a read count number (which can be normalised) to the particular feature/gene. If there is no reference genome or reference transcriptome available, the sequencing reads can be assembled *de novo* to create a novel transcriptional map based only on the collected data. In this case, paired-end reads and a very high level of coverage are essential (Wang *et al.*, 2009).

Once the reads have been assembled into transcripts and assigned, the resulting high density transcriptome maps can be compared across samples to look for differential splicing, changes in expression, changes in transcriptional start sites (TSS), and many other quantitative differences (Trapnell *et al.*, 2013). Ideally, due to the inherent variation in these experiments even across biological replicates, it is recommended to include at the very least 3 biological replicates of each condition. More replicates are definitely a good idea if the questions being asked pertain to lowly expressed genes or relatively small differences in expression levels (Trapnell *et al.*, 2013).

Ultimately, RNA-seq differential expression results should be confirmed with other independent approaches to make sure they are not an artifact of the library generation process. When looking genome-wide, even if taking stringent false positive thresholds, it is still likely to see a high number of false positives due to the sheer number of tests being carried out simultaneously. Taking a p-value cutoff of 0.05 and carrying out 16000 simultaneous tests, one would expect to see 800 false positives based on chance alone. Benjamini-Hochberg correction is employed by most software modules to adjust significance thresholds and minimise false positives.

Validation of differential expression can be done with real time quantitative PCR (RT-qPCR), microarrays, or different sequencing platforms. When done correctly with several housekeeping genes, RT-qPCR is a good tool to validate single hits obtained in RNA-seq experiments (Ling & Salvaterra, 2011). Unfortunately, most published RT-qPCR experiments only use single housekeeping genes and really should not be trusted, especially if the abundance of the housekeeping gene is not close to the candidates being tested. When attempting to validate a large number of candidate genes, RT-qPCR can become very time-consuming and expensive. Microarrays and 3′ tag DGE can do better at validating large numbers of genes in a time and cost-efficient manner (Trapnell *et al.*, 2013).

1.5.2 3' Tag DGE

Depending on the type of research question being asked, it is very important to select an appropriate sequencing approach. To ask purely about differential expression without any concern for the structure of the transcripts, tag sequencing approaches are applicable (Harbers & Kahl, 2012). In the 3′ tag DGE system, RNA is ultrasonically fragmented and the 3′ ends of the fragments are separated using oligo-dT chromatography. Some other protocols do not fragment the RNA ultrasonically, but rather wait until the cDNA has been reverse-transcribed before enzymatically cleaving the transcripts with a simple 4 base cutter (Asmann *et al.*, 2009). The enzymatic protocol results in much shorter tags (≈20 bp), whereas ultrasonication can generate much longer tags (≈100 bp) (Zhang *et al.*, 2013). Most of the resulting reads then map to the 3′ UTR or final exon of their

respective genes. By counting unique tags and comparing across samples, it is possible to detect changes in gene level expression even for very lowly expressed genes, but transcript information at the individual isoform level is lost. Also, genes with non-unique tags at their 3′ end are difficult to assay using these methods. However, increasing the tag length alleviates this issue (Zhang *et al.*, 2013). It has been clearly demonstrated that 3′ tag DGE has much higher sensitivity and reproducibility than microarrays, and expression quantification correlates very well with RT-qPCR across many orders of magnitude (Asmann *et al.*, 2009; Zhang *et al.*, 2013).

So far, 3′ tag DGE has received much less attention than standard RNA-seq. The current focus of most transcriptomics studies is to catalog the transcriptome and document newly discovered splicing events (Consortium, 2004), as well as to come to grips with all the different ways the genome can be altered post-transcriptionally, but 3′ tag DGE exclusively studies differential gene expression. If expression differences between conditions are not significant, RNA-seq can be still be regarded as useful data if it detects new isoforms, different splice junctions, or undocumented transcriptional start sites. The 3′ tag DGE data would be useless if no expression differences were detected.

The well-documented agreement between all of these different ways to measure the transcriptome justifies their continued use as experimental tools. Based on the nature of the methods, it is clear that for sensitivity, discovery of new genes and splicing events, reproducibility, and the ability to compare many experiments done at different times, sequencing based methods are far and above the best option. That being said, the size and complexity of the resulting datasets need to be handled by skilled bioinformaticians who are clued into the software available for analysis. The only way to plan and carry out successful NGS experiments is to set up beforehand adequate computing facilities capable of parallel processing of large amounts of data in a timely fashion.

1.5.3 Bioinformatics

The bioinformatics requirements for NGS data are substantial, thus restricting many labs in their use of these cutting edge techniques. There is a large amount of available freeware, but comparisons between the different software modules are generally inadequate, biased or both. Starting with raw reads and quality scores, the various software options explored in this thesis will be detailed here following the work flow in Figure 1.10.

1.5.3.1 Pre-mapping Read Quality Assessment

At this stage, it is important to trim low quality base calls and remove potential contaminants like overrepresented adaptor sequences, unless the read mapper used is quality

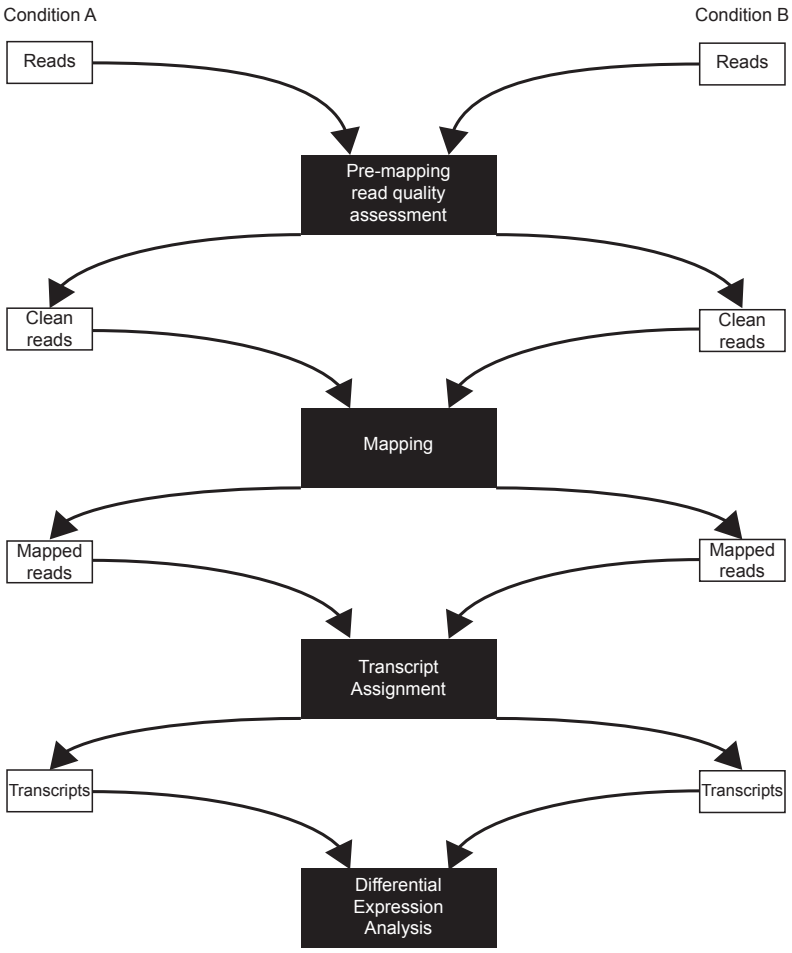


Figure 1.10: *The full work flow for the bioinformatics analysis involved in RNA-seq*

aware (*bowtie2* will be introduced later and is quality aware). Either way, cleaning up the low quality base calls and foreign sequences speed up the mapping step significantly. A free software package called *FastQC* provides a good way to visualise the quality of the raw data and it runs several important tests to flag any oddities or amplification artifacts (Andrews, 2012). A java-based software called *Trimmomatic* provides a powerful and simple read trimmer capable of removing adapters, primers, or any undesired sequence and works on both single-end and paired-end libraries (Bolger, 2011). The *CLC workbench* provides a much more extensive way to both visualise and correct read anomalies, but it is not free (Harbers & Kahl, 2012).

1.5.3.2 Mapping

If the experimental model organism has a reference genome sequence available, mapping the reads is much faster because *de novo* assembly is not required (Trapnell *et al.*, 2012). As this is the case in DM, it narrows the options down to a few read mappers. The most widely used mapper, *bowtie2*, is part of the *tuxedo* software suite which also includes *tophat* for RNA-seq splice junction mapping, *cufflinks* for transcript isoform assembly and quantification, and *cummeRbund* for visualising RNA-seq analysis (Trapnell *et al.*, 2012). The advantages of using *bowtie2* to map reads over the competitors are that it is faster, takes up a smaller memory footprint, and takes into account the quality score of each base (Langmead & Salzberg, 2012). *Tophat* uses *bowtie2* to map short reads to a reference sequence while at the same time detecting splice junctions within the reads (Trapnell *et al.*, 2012). A single RNA-seq library consisting of ≈15 million paired-end reads takes approximately 5 hours to map using this tool on a regular desktop computer. *Stampy* is another read mapper that handles insertions and deletions and intrinsic genomic variation better than *bowtie2*, but at the cost of significantly longer processing times (≈ 24 hours for the same task) and increased memory requirements (Lunter & Goodson, 2011).

1.5.3.3 Transcript Allocation

Once the reads are mapped, they need to be assembled into transcripts. Again, there are publicly available sets of reference transcripts for some experimental organisms. These are based on the conglomeration of many published studies and curated by large bioinformatics centres around the world. DM has one of the best annotated genomes, and using the references can greatly expedite the transcript calling process.

The main issues involved in detecting and quantifying transcripts are (i) assigning reads to specific isoforms, which is difficult sometimes because of the similarity between exonic gene sequences and many isoforms share some but not all exons (count uncertainty),

and (ii) accounting for variability in measurements across biological replicates (count overdispersion) (Trapnell *et al.*, 2013). The majority of the available software addresses one of these problems but ignores the other.

The Bioconductor package *DEseq* uses a negative binomial distribution to model overdispersion based on raw read counts, but does not attempt to address ambiguities in isoform mapping (Anders & Huber, 2010). Another available software package also within the Bioconductor realm, *edgeR* (Robinson *et al.*, 2010), uses an overdispersed Poisson model that is moderated across genes using empirical Bayesian methods. Both of these platforms choose to ignore count uncertainty, and instead only tackle overdispersion. They both require raw read counts as input and can carry out differential expression tests across many conditions. Because they are specifically built to monitor gene-level differential expression, they are more sensitive than the competitors in detecting differential expression at the gene level.

Another software option, *cuffdiff2*, juggles both count uncertainty and count overdispersion issues by modeling how measurement variability depends both on expression levels and the splicing structure of genes (Trapnell *et al.*, 2013). Uncertainty in the fragment counts between isoforms is modeled with a beta distribution and overdispersion is simultaneously modeled with a negative binomial, resulting in a beta negative binomial distribution reflecting both sources of variability (Trapnell *et al.*, 2013). The result is that *cuffdiff2* seems to be much more stringent in declaring statistical significance in differential expression at the gene level. Additionally, *cuffdiff2* can only test two conditions at a time.

All of these differential expression detection approaches operate on precise assumptions about the underlying distributions of read counts, and can be susceptible to outliers and other anomalies that are very common in transcriptomics studies (Li & Tibshirani, 2013). In light of this, permutation-based non-parametric methods are available that do not depend on distributional assumptions (Li & Tibshirani, 2013). These methods are the most appropriate when the number of biological replicates of each condition exceeds 8 (Anders & Huber, 2010). A software package called *SAMseq* exists that has been adapted from microarray analysis modules that can carry out this permutation analysis to detect differentially expressed genes by generating an empirical distribution and directly calculating the corresponding p-values as well as correcting for multiple testing (Li & Tibshirani, 2013).

To be absolutely sure about genes called as differentially expressed, running many software packages and taking an intersection is the safest and most thorough approach to assure the genes are not artifacts of the specific statistical methods being employed.

Chapter 2

Materials and Methods

2.1 Flies

2.1.1 Animal Rearing

Fly stocks were kept for long term storage in standard conditions (SC) at 21°C, 65% humidity on a constant 12:12 light/dark (LD) regimen.

For experiments involving $GAL80^{ts}$, crosses were set up and kept in SC. Eclosing flies were collected every day on CO_2 pads, and then put back in SC until they were moved to the induction incubator. The induction incubator was kept at 31°C, 70% humidity on the same 12:12 LD schedule with an external source of moisture (a beaker of water). All LD cycles were synchronised to minimise environmental variation.

2.1.2 Food

For long term health and maintenance, individual stocks were kept in rich cornmeal and molasses based food with brewer's yeast (brown food). Small scale crosses were also kept on brown food. Large scale crosses were set up in food without molasses (yellow food). Although molasses increases the number of eggs, it also makes the food soupier and thus many flies get stuck and die.

2.1.3 Lines

2.1.3.1 Commercially Available Lines

Drivers

y^1 w^{1118}; $P\{GawB\}GH146$ (Brand & Perrimon, 1993) (BDSC #30026)

For driving gene expression in postsynaptic circuit elements, the GH146-GAL4 driver line

was selected because it expresses cleanly in large subsets of canonical PNs from all three AL cell clusters. It drives expression in ≈45 of the 70 anterodorsal PNs, ≈41 lateral PNs , and ≈7 of the 50 ventral PNs (Figure 2.1). In total, GH146-GAL4 covers ≈100/160, or ≈63% of the documented PNs. Further, only looking at canonical *excitatory* PNs, GH146-GAL4 drives in ≈100/120, only missing out on a few in the anterodorsal lineage. Importantly, it does not express in any ORNs or LNs, but it does express in ≈6 cells in the LH and ≈3 cells ventral to the subesophageal ganglion. The central brain expression pattern is shown in Figure 2.1. GH146-GAL4 also expresses in a subset of ≈30 glia per hemisphere that are born near the AL and migrate along the antennal nerve towards the third antennal segment (Sen *et al.*, 2005). The role these glia play and the extent to which they enter the third antennal segment are poorly characterised.

y^1 w^*; $P\{pdf\text{-}GAL4.P2.4\}2$ (Helfrich-Forster *et al.*, 2000) (BDSC #6900)
This driver line was used to validate the function of the Kir2.1/Kir2.1-nc system. It drives in a small population of neurons important for setting the central circadian molecular clock (Helfrich-Forster *et al.*, 2000). It is well-documented that disrupting the electrical activity of pdf^+ neurons using this line disturbs the circadian locomotory behaviour of the fly (Nitabach *et al.*, 2002).

Driver Repression

$P\{tubP\text{-}GAL80^{ts}\}Sxl^9, w^*/FM7c$ (McGuire *et al.*, 2004) (BDSC #7016)
This was used to repress GAL4 mediated expression in several experiments. Any time $GAL80^{ts}$ was used on the X chromosome, it was this particular insertion. It has been recombined and/or rebalanced in many experiments.

w^*; $P\{tubP\text{-}GAL80^{ts}\}20$; $TM2/TM6B$ (McGuire *et al.*, 2004) (BDSC #7019)
This was used to repress GAL4 mediated expression in several experiments. Any time $GAL80^{ts}$ was used on the second chromosome, it was this particular insertion. It has been recombined and/or rebalanced in many experiments.

w^*; sna^{Sco}/CyO; $P\{tubP\text{-}GAL80^{ts}\}7$ (McGuire *et al.*, 2004) (BDSC #7018)
This was used to repress GAL4 mediated expression in several experiments. Any time $GAL80^{ts}$ was used on the third chromosome, it was this particular insertion. It has been recombined and/or rebalanced in many experiments.

Overexpression Lines

w^*; $P\{UAS\text{-}mCD8::GFP\}$; $Sb/TM3, Ser$ (Lee & Luo, 1999) (BDSC #5137)
Any pictures in this thesis with mCD8::GFP as the reporter were done using this line.

w^*; $P\{UAS\text{-}TeTxLC\}G2$ (Keller *et al.*, 2002) (BDSC #28838)
This was used in the behavioural experiments to see how odour preferences change when

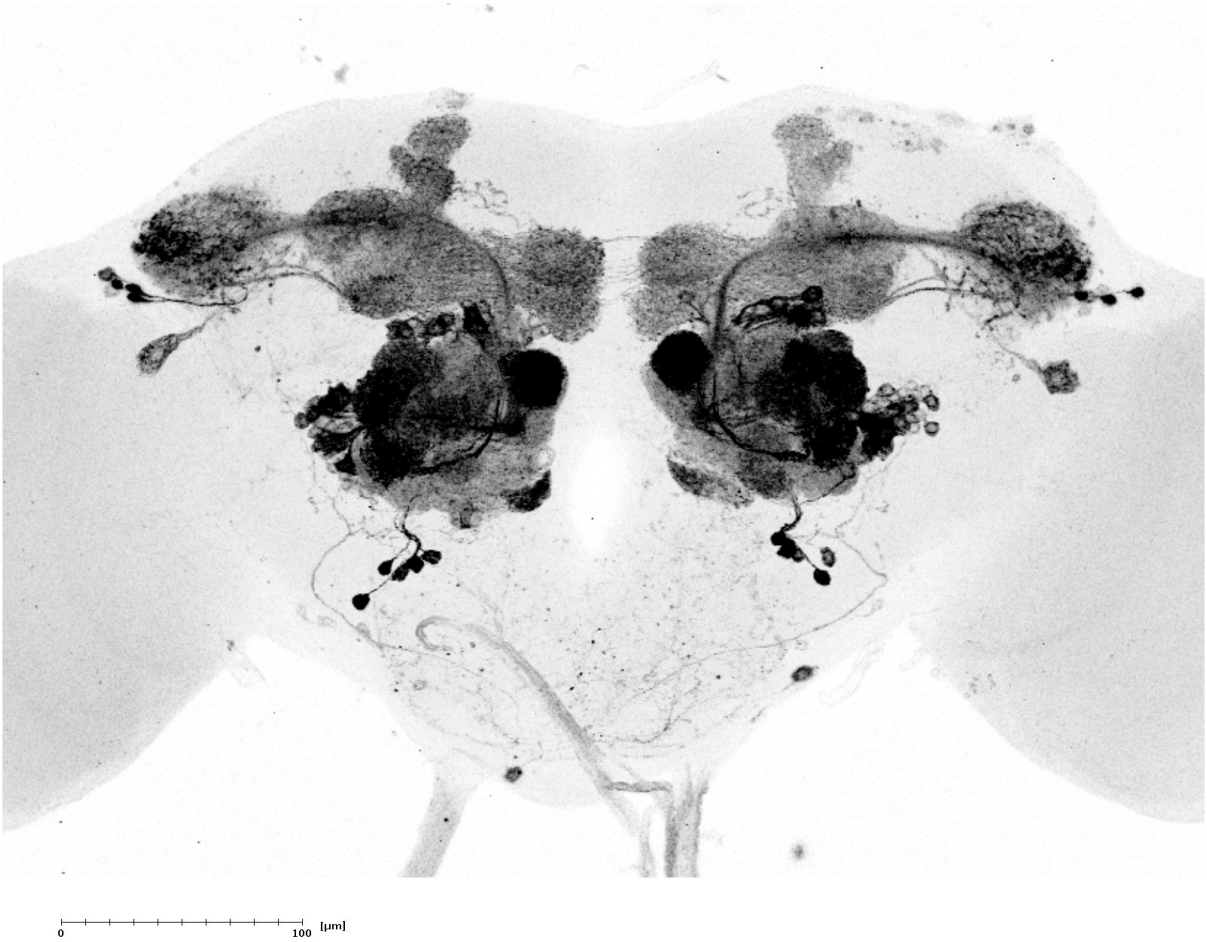


Figure 2.1: *GH146-GAL4 driven UAS-mCD8::eGFP*

the $GH146^{+}$ cells are blocked with a tetanus toxin light chain that cleaves synaptobrevin (Keller *et al.*, 2002).

w^{*}; $P\{UAS\text{-}TeTxLC_{inactive}\}A2$ (Keller *et al.*, 2002) (BDSC #28840)
This was used as a control for the behavioural experiments that overexpress an inactive form of the tetanus toxin light chain (Keller *et al.*, 2002).

w^{*}; $P\{UAS\text{-}shibire^{ts}\}/TM3, Ser$ (Kitamoto, 2001) (gift from Toshihiro Kitamoto)
This was used in the behavioural experiments to see how odour valence changes when $GH146^{+}$ cells are blocked with $shibire^{ts}$, a temperature-sensitive dynamin mutant.

w^{*}; $P\{UAS\text{-}eGFP::Kir2.1_{human}\}/CyO$ (Baines *et al.*, 2001) (BDSC #6596)
This line is a published Kir2.1 that was used as a positive control when validating the DM codon-optimised system (Baines *et al.*, 2001).

w^{*}; $P\{LexAop\text{-}myr::mCherry\}/TM3, Sb$ (Diegelmann *et al.*, 2008)

This was used to visualise the expression pattern of the Or83b-LexA driver shown in Figure 1.7c.

Injection Lines

y^1 $M\{vas\text{-}int.B\}ZH\text{-}2A$ w^*; $P\{CaryP\}attp2$ (BDSC #25710 and #36313)

This line was constructed from the two available BDSC lines in order to be able to use w^* as a marker for ΦC31 injections targeted to the attp2 insertion site on the third chromosome. The endogenously expressed transposase greatly increased the efficiency of integration.

Balancer Lines

w^*; sp/CyO; $MKRS/TM6B, Tb$ (BDSC #3703)

This line was repeatedly used to double balance all stocks that needed it. The TM6B balancer was always used on the third chromosome because it is a better balancer than MKRS.

Wild Type

Canton-S (BDSC #1)

This line was used as a wild type control for behavioural experiments with $shibire^{ts}$.

w^{1118} (BDSC #5905)

This line was used as a wild type white-eyed strain to outcross injected flies.

2.1.3.2 Generated Lines

In-house Injections

w^*; sp/CyO; $P\{UAS\text{-}eGFP::Kir2.1_{dm}\}attp2/TM6B, Tb$

w^*; sp/CyO; $P\{UAS\text{-}eGFP::Kir2.1_{dm}\text{-}nc\}attp2/TM6B, Tb$

Flies were placed on grape juice agar plates with a dab of yeast paste in cages for 24 hours at 22°C prior to embryo preparation. During the morning on the day of injection, the flies were flipped to new plates every hour to get them laying and accustomed to being flipped. During injections, plates were flipped every 20 minutes to confine embryos to the appropriate stage.

Immediately after flipping, embryos of the genotype *y vas-int.B w; attp2* were dechorionated with 50% bleach for 2 minutes on the grape juice agar plates with light shaking. They were then washed thoroughly with distilled water, poured into a fine mesh sieve, and washed again thoroughly to remove any traces of bleach. The dechorionated embryos were then lined up on a piece of agar stained with blue dye for easier visualisation. They were transferred to a glass injection slide with double-sided tape, desiccated for 7

minutes in a box with cobalt chloride, and covered with a drop of Polysciences Halocarbon 200 oil (catalog # 9002-83-9). Vaseline was used to restrict the oil from spreading too thinly. The embryos were then serially injected using the Eppendorf Femtojet Express with ≈150 ng/μL plasmid DNA containing a small amount of green dye to ensure successful microinjection. For these injections, pJFRC2-10XUAS-IVS-eGFP::Kir2.1 and pJFRC2-10XUAS-IVS-eGFP::Kir2.1-nc were the plasmids used, which will be described in Section 2.2.

Injected embryos were kept at 18°C until larval hatching and then transferred one by one to brown food and placed in SC. Eclosing adult males were singly crossed to w^{1118} virgin females. Red-eyed males were selected in the next generation and double balanced by repeated crossing to $w^*;\ sp/CyO;\ MKRS/TM6B, Tb$ virgin females.

BestGene Injections

$w^*;\ P\{UAS\text{-}Kir2.1_{dm}\}attp2/TM3, Sb$
$w^*;\ P\{UAS\text{-}Kir2.1_{dm}\text{-}nc\}attp2/TM3, Sb$
$w^*;\ P\{UAS\text{-}Kir2.1_{dm}\}attp40/CyO$
$w^*;\ P\{UAS\text{-}Kir2.1_{dm}\text{-}nc\}attp40/CyO$
$w^*;\ P\{LexAop\text{-}Kir2.1_{dm}\}attp2/TM3, Sb$
$w^*;\ P\{LexAop\text{-}Kir2.1_{dm}\text{-}nc\}attp2/TM3, Sb$
$w^*;\ P\{LexAop\text{-}Kir2.1_{dm}\}attp40/CyO$
$w^*;\ P\{LexAop\text{-}Kir2.1_{dm}\text{-}nc\}attp40/CyO$
$w^*;\ P\{QUAS\text{-}Kir2.1_{dm}\}attp2/TM3, Sb$
$w^*;\ P\{QUAS\text{-}Kir2.1_{dm}\text{-}nc\}attp2/TM3, Sb$
$w^*;\ P\{QUAS\text{-}Kir2.1_{dm}\}attp40/CyO$
$w^*;\ P\{QUAS\text{-}Kir2.1_{dm}\text{-}nc\}attp40/CyO$

These lines were created to express Kir2.1/Kir2.1-nc without any eGFP-tag using GAL4, LexA, and QF drivers. Some future experimental approaches might require the GFP channel to be available (GCaMP, synaptophluorin imaging) and these lines will facilitate those approaches.

$w^*;\ P\{Or7a\text{-}nlsLexA::p65\}attp40/CyO$
$w^*;\ P\{Or67a\text{-}nlsLexA::p65\}attp40/CyO$
$w^*;\ P\{Or82a\text{-}nlsLexA::p65\}attp40/CyO$

These lines were created to provide some flexibility in driving different elements under LexA control. Up until these were made, almost all sparse PN and ORN drivers were only available through the GAL4 system. These will allow the use of GAL4 drivers for PNs and LexA drivers for the presynaptic ORNs in the same animal.

2.2 Molecular Biology

2.2.1 Kir2.1 and Kir2.1-nc Cloning

A DM codon-optimised version of the human Kir2.1 gene was synthesised at MWG Eurofins with the following sequence:

ATGGGTAGTGTTCGGACGAATCGCTACTCTATCGTGTCGTCGGAGGAAGAT
GGCATGAAATTGGCAACCATGGCGGTGGCTAACGGCTTCGGTAACGGTAAA
TCGAAAGTTCACACCCGCCAACAGTGTCGTTCTCGATTTGTGAAGAAGGAC
GGACATTGCAACGTACAGTTCATCAATGTGGGCGAAAAGGGACAACGCTAT
CTAGCGGACATATTTACAACGTGTGTGGATATCAGGTGGCGTTGGATGTTG
GTGATCTTCTGCCTCGCTTTCGTGCTGAGCTGGCTGTTCTTCGGCTGTGTA
TTCTGGCTCATCGCACTGCTGCATGGAGATCTTGATGCCTCCAAAGAGGGC
AAGGCTTGCGTAAGCGAAGTGAACTCCTTTACTGCCGCGTTTCTGTTTTCC
ATCGAGACTCAGACCACGATAGGCTACGGATTCCGATGCGTTACCGACGAG
TGCCCCATAGCCGTGTTTATGGTGGTGTTCCAGAGTATTGTCGGTTGCATC
ATTGATGCCTTCATTATCGGCGCAGTCATGGCCAAGATGGCCAAGCCCAAG
AAGAGGAACGAGACACTCGTGTTCTCACACAACGCCGTTATTGCCATGCGC
GACGGCAAACTGTGCCTCATGTGGCGCGTCGGAAACTTGCGGAAATCGCAC
CTAGTGGAAGCTCATGTCCGCGCGCAGCTCCTGAAGTCCCGGATAACATCG
GAAGGCGAGTACATCCCGCTGGATCAGATCGACATAAATGTGGGCTTTGAT
AGCGGAATTGACCGCATCTTCTTGGTGAGCCCCATCACCATCGTTCACGAG
ATCGACGAAGATAGCCCATTGTACGATCTGTCCAAACAGGATATCGACAAT
GCCGACTTCGAGATTGTCGTCATTCTGGAGGGTATGGTGGAAGCCACTGCC
ATGACAACGCAATGTCGAAGCTCTTATCTTGCAAATGAGATACTGTGGGGA
CACAGGTATGAGCCGGTCCTGTTTGAGGAGAAGCACTACTACAAGGTTGAT
TACTCGCGTTTCCATAAGACGTATGAGGTGCCCAACACTCCCCTGTGCTCAG
CTCGCGATCTTGCGGAAAAGAAGTACATTCTGTCCAACGCCAACAGCTTTT
GCTATGAGAATGAAGTAGCACTAACCTCCAAGGAAGAGGATGACTCCGAGA
ATGGCGTTCCAGAGAGCACCAGTACGGACACCCCTCCGGATATTGACTTGC
ACAATCAGGCCTCAGTCCCTCTGGAGCCACGTCCGCTGAGAAGAGAGAGTG
AGATTTGATGA

The Kir2.1 DM codon-optimised sequence was combined with an N-terminal DM codon-optimised GFP (Addgene plasmid # 26224) using hybrid PCR with the primers in Table 2.1.

This N-terminal fusion was TOPO cloned into pENTR/D-TOPO (Life Technologies product # K2400-20). Site-directed mutagenesis was carried out in pENTR/D-TOPO using the primer 5'-GACCACGATAGGCTACAGCTTCCGATGCGTTACCG-3′ and the

Primer Name	Primer Sequence
eGFP-5'	CACCATCAACATGTCCAAAGGTGAAGAACTGT
eGFP-GGGS-3′	CTCTACAAGGGCGGTGGATCCGGTAGTGTTCGGACGAAT CGCTACTCTAT
GGGS-Kir2.1-5'	AACACTACCGGATCCACCGCCCTTGTAGAGCTCATCCAT GCCGTGTGTAA
Kir2.1-3′	CTCGAGTCATCAAATCTCACTCTCTCTTCT

Table 2.1: *Primers for eGFP::Kir2.1 hybrid PCR*

QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent product # 210518), which resulted in the mutation of amino acid 146 from a glycine to a serine (G146S). Kir2.1-G146S (referred to in this thesis as Kir2.1-nc) has been shown to abolish the flow of potassium through the channel pore without disturbing the membrane localisation of the protein (Haruna *et al.*, 2007). It thus creates a non-conducting potassium channel. Other similar mutations that have been used to accomplish this same goal actually disrupt trafficking and membrane localisation and would not have been suitable for transcriptomics (Nitabach *et al.*, 2002).

The destination vector was prepared using Addgene plasmid # 26214 (pJFRC2-10XUAS-IVS-mCD8::GFP). pJFRC2-10XUAS-IVS-mCD8::GFP was digested with XhoI and XbaI to remove the mCD8::GFP. The vector backbone was then blunted with the DNA Polymerase I large (Klenow) fragment (New England Biolabs product # M0210S). Gateway conversion cassette B (Life Technologies catalog # 11828-029) was then blunt-end ligated into the vector to create pJFRC2-10XUAS-IVS-GWcassetteB.

Both eGFP::Kir2.1 and eGFP::Kir2.1-nc were then recombined into the destination vector pJFRC2-10XUAS-IVS-GWcassetteB using the standard LR reaction protocols in the Gateway pipeline to give pJFRC2-10XUAS-IVS-eGFP::Kir2.1 and pJFRC2-10XUAS-IVS-eGFP::Kir2.1-nc.

For the occasion when it will be beneficial to express Kir2.1 or Kir2.1-nc without an eGFP tag, the Kir2.1 and Kir2.1-nc sequences were later PCR amplified with primers 5 '-CACCATCAACATGGGTAGTGTTCGGACGAATCGCTAC-3′ and 5'-CTCGAGTC ATCAAATCTCACTCTCTCTTCT-3′ and then TOPO cloned into pENTR/D-TOPO. The same destination vector as before was used for the UAS lines (pJFRC2-10XUAS-IVS-GWcassetteB) a destination vector called PMO384 was used for the LexAop2 lines, and pQUAS-attB-GW was used for QUAS lines (Hong *et al.*, 2012). PMO384 has 13 LexAop2 sites, gypsy insulator elements, an attB recombination site, an SV40 transcription termination sequence, and is Gateway competent. pQUAS-attB-GW is the same as pUAS-attB-GW (Tea *et al.*, 2010) with the UAS site swapped for a QUAS site (Hong *et al.*, 2012).

2.2.2 Or-LexA Cloning

For the future of the project, it will be important to be able to independently label specific presynaptic and postsynaptic neurons. The current tools are mostly available only as GAL4 drivers, and therefore it will be necessary to generate LexA and/or QF lines to push the project forward. With this in mind, three Or-LexA lines were generated. These Or-LexA drivers were chosen based on the ability to address their cognate PNs with GAL4 drivers, thus granting access to known cognate sets of presynaptic and postsynaptic neurons.

Genomic DNA was extracted from the existing GAL4 lines for all 3 of these drivers and PCR was done from the genomic DNA using the original primers used to make the GAL4 drivers listed in the Table 2.2 (Couto *et al.*, 2005).

Primer Name	Primer Sequence
Or7a-5'	CACCACCGCATCCCGATCAAGACACAC
Or7a-3′	TGATGGACTTTTGACGCCTGGGAATA
Or67a-5'	CACCCTTCCGCAAATGAATGGAAAACATCA
Or67a-3′	TTCAATATGGTTTGAATAAACACTGATTGATACAGG
Or82a-5'	CACCCTTAGCATTGAGTCCTCTGTTACAAGTGTTGTTATT
Or82a-3′	GACCCCAGTTTCTAGAACATGAAAGGATTG

Table 2.2: *Primers for Or-LexA promoter cloning*

These were then all TOPO cloned into pENTR/D-TOPO and Gateway recombined into Addgene plasmid # 26230 (pBPnlsLexA::p65Uw) resulting in three plasmids that were sent to BestGene for injection.
(i) pBP-Or7a-nlsLexA::p65Uw
(ii) pBP-Or67a-nlsLexA::p65Uw
(iii) pBP-Or82a-nlsLexA::p65Uw

They were all incorporated at the attp40 landing site on the second chromosome.

2.3 Behavioural Monitoring

2.3.1 Odour Valence Measurement

The *Multiplex*, which was shown in Figure 1.5b, is a high-throughput behavioural machine allowing simultaneous presentation of two odours to 20 single flies (Claridge-Chang *et al.*, 2009). It is also capable of delivering electric shock punishment to train the flies, but that functionality was not used in this work. A mass-flow-controller-monitored carrier

stream of air is pushed through the system which provides ambient wind levels in the chambers of ≈1.5 kilometres per hour. There are electronically controlled valves that when opened, direct a portion of the carrier stream through odour vials which contain odours diluted in solvent. With incoming airflow, the odours bubble into the air stream and are presented to each fly in each chamber simultaneously. The three odours used in this thesis were methyl salicylate (Sigma Aldrich catalog # M6752, diluted to 10% in mineral oil), ethyl butyrate (Sigma Aldrich catalog # 75563, diluted to 1% in mineral oil), and geranyl acetate (Sigma Aldrich catalog # 173495, diluted by 10% in mineral oil). Three different lines of odour valence experiments were done with the *Multiplex* in this thesis.

shibire^{ts} Experiments

Flies were raised in SC until they were 5 days old. They were then placed in the chambers shown in Figure 1.5a, preheated at 31°C without air flow for 20 minutes, then placed for 15 minutes on clean carrier stream air flow in the Multiplex at 31°C to allow them to acclimatise. They were then presented an odour on one side of the chamber and plain mineral oil on the other side of the chamber for one minute, followed by a one minute break, and then the mineral oil and odour were presented on the opposite sides for one minute. Four odours in total were presented to single groups of 20 flies and they were always presented in the same order. The first odour gave inconsistent results and was thrown out.

Kir2.1 Experiments

Flies were raised in SC until it was time for $GAL80^{ts}$-mediated induction of Kir2.1 expression. For the induction, they were placed at 31°C in food for the desired number of hours (12, 24, 48, or 96). The day of the experiment, they were placed in the chambers shown in Figure 1.5a, preheated at 31°C without air flow for 20 minutes, and then placed on carrier stream air flow at 31°C for 15 minutes to allow them to acclimatise. They were then presented an odour on one side of the chamber and plain mineral oil on the other side of the chamber for one minute, followed by a one minute break, and then the mineral oil and odour were presented on the opposite sides for one minute. Four odours in total were presented and they were always presented in the same order. The first odour gave inconsistent results and was thrown out.

TeTxLC Experiments

Flies were treated exactly as described in the Kir2.1 section above except these flies were not tested with 96 hours of induction.

Valence Calculation

The flies odour valences were calculated in two independent ways. The *decision* metric takes into account each time a single fly enters the middle of the chamber designated

the *decision zone*. For these experiments the decision zone was set to be a symmetric area around the centre of the chamber including 20% of the full chamber length on either side. Each time the fly enters and leaves this area, it is counted as a choice either to approach the odour or avoid the odour (approach the mineral oil). Flies making less than 4 decisions per minute were discarded. The total number of approach decisions minus the number of avoid decisions divided by the total number of decisions gives an attraction/aversion score.

Another metric, the *time* metric, simply calculates how much time the fly spends on each side of the chamber when the odour is being presented by taking total time on the odour side of the chamber minus the total time on the mineral oil side divided by the length of the experiment (2 minutes). The different metrics generally support the same conclusion unless a fly is not very active, in which case it is discarded anyway (at least 4 decisions must be made). The decision metric will be reported in this thesis.

2.3.2 Circadian Rhythm

Three day old female flies were placed into glass tubes with food on one end and a plug on the other as shown in Figure 1.6a. They were placed in an incubator at 31°C in 24 hour dark (DD) conditions and monitored for 5-7 days to observe their locomotory behaviour.

2.4 Imaging

2.4.1 Immunohistochemistry

Five day old female flies were anaesthetised on ice and dissected in PBS. Within 20 minutes of dissection, brains were transferred to ice cold PBS-formaldehyde 4% and fixed for 20-30 minutes at room temperature. After fixing, brains were immediately washed three times with ice cold PBT 0.1%, and then washed an additional 3 times for 20 minutes each also in ice cold PBT 0.1%. Brains were then blocked for one hour at room temperature in PBT 0.1% with 5% NGS, and left in primary antibody for 20 hours at 4°C in PBT 0.1% with 5% NGS. Brains were washed 6 times the next day at room temperature in PBT 0.1%, and then left overnight at 4°C in secondary antibody with PBT 0.1% and 5% NGS. Brains were then mounted on glass slides, cleared with Vectashield overnight at room temperature, and imaged the next day.

2.4.2 Native Protein Imaging

Female flies aged 4-6 days were anaesthetized on ice and dissected in PBS. Immediately after dissection, brains were transferred to ice cold PBS-formaldehyde 4% and fixed for 1-2 hours at room temperature. After fixing, brains were immediately washed three times with ice cold PBT 0.1%, and then washed an additional 3 times for 20 minutes each also in ice cold PBT 0.1%. Brains were mounted on glass slides, cleared with Vectashield, and left at room temperature overnight for imaging the next day.

2.4.3 Confocal microscopy

For whole brain confocal stacks, 1-1.5 μm sections were taken on a Leica TCS SP5 microscope using a water immersion 25X objective. When possible, stacks were z-compensated by increasing the PMT gain as the scan moved deeper into the tissue.

2.5 RNA Sample Preparation

2.5.1 Third Antennal Segment Separation

Flies were sorted into groups of 20-30 when they first eclosed and kept in precisely controlled temperature conditions until decapitation. They were then manually decapitated on a CO_2 pad using a surgical scalpel. The heads were transferred with a paint brush ≈12 at a time to petri dishes on dry ice. Once a petri dish contained ≈50 heads, it was sealed with parafilm and placed at -80°C until the day of RNA extraction.

On the day of extraction, the sealed petri dishes containing heads were removed from the -80°C on dry ice, dipped in liquid nitrogen for 60 seconds, vortexed at full strength for 60 seconds and then unsealed and placed on a dry-ice-chilled glass stand under a standard dissection microscope. Individual third antennal segments could then easily be picked with fine forceps and placed directly into 100 microlitres (μL) of TRIzol in a glass well. The third antennal segments were then transferred en masse to a 1 mL dounce homogeniser (VWR product # 432-1270) to begin the total RNA extraction.

2.5.2 RNA extraction

2.5.2.1 RNA-seq and RT-qPCR

The dounce homogenisers were kept coarse by regularly grinding glass beads within them into sand, with extra care taken not to grind too much which results in the smoothing

of the inside of the glass. This is very important because if they are allowed to become too smooth, the third antennal segments do not homogenise fully. The homogenisers are always cleaned with either 50% bleach or hydrogen peroxide (or both) before and after use to prevent contamination.

The third antennal segments were transferred into homogenisers in 100 μL of TRIzol and ground until they were visually confirmed to be fully disrupted under a standard dissection scope. That means there are no longer any visible intact third antennal segments. The homogeniser was then topped up with 900 μL of TRIzol, and the resulting 1 mL of homogenate was transferred into 1.5 mL RNase-free microcentrifuge tubes. The tubes were left at room temperature for 3-5 minutes before 225 μL of chloroform was added to each. They were then closed, inverted 4-6 times, briefly spun down, and then kept at room temperature for another 3 minutes. The tubes were then spun for 15 minutes at 13,000 rpm (≈15000XG) at 4°C. Next, 400 μL of the upper aqueous phase was placed in a fresh, pre-chilled 1.5 mL RNase-free microcentrifuge tube, being very careful not to disturb the interphase and organic phase below. An additional 400 μL of 70% ethanol (RNase-free) was then added to the tube and mixed by pipetting up and down. All 800 μL was then added to a cold RNeasy minelute column (Qiagen catalog # 74004). Standard protocols for the RNeasy kit were used along with DNaseI on-column digests to eliminate genomic DNA contamination. The column was then double eluted with 15 μL of RNase-free water. Samples were split into three 5 μL aliquots, snap frozen in liquid nitrogen, and then placed in the -80°C freezer for long term storage.

2.5.2.2 3' Tag DGE

Everything was done exactly the same way as the RNA-seq/RT-qPCR samples until the precipitation step. Instead of 400 μL of 70% ethanol, 400 μL of pure isopropanol was added. The tubes were mixed by pipetting up and down and left for 15 minutes at room temperature. The tubes were then spun at 13,000 rpm (≈15000XG) at 4°C for 20 minutes. The supernatant was pipetted away, then the pellets were washed 3 times with 70% ethanol (RNase-free), and spun down for 10 minutes at 13,000 rpm (≈15000XG) at 4°C after each wash. The pellets were then dried for 10 minutes at room temperature and resuspended in 5 μL of RNase-free water. Finally, they were snap frozen in liquid nitrogen, and placed in the -80°C freezer for long term storage.

2.6 cDNA Library Generation

2.6.1 RNA-seq and RT-qPCR

The mRNA was oligo-dT separated from total RNA (≈1 ng starting material) and amplified using the SMARTer Ultra Low RNA Kit for Illumina Sequencing (Clontech catalog # 634936) with 14 cycles of amplification. After cDNA fragmentation, the sequencing libraries were prepared using the NEBNext Ultra DNA Library Prep Kit for Illumina (NEB catalog # E7370L) with an additional 15 cycles of amplification.

2.6.2 3' tag DGE

This particular protocol was carried out at MWG Eurofins Genomics and certain details were not made available because of their proprietary nature. The total RNA was first ultrasonically fragmented. Poly(A)-tailed RNA fragments were then isolated using oligo-dT chromatography. An RNA adapter was next ligated to the 5'-ends of the isolated RNA fragments. First-strand cDNA synthesis was then performed using an oligo(dT)-adapter primer and reverse transcriptase. The resulting cDNA was PCR-amplified using a high fidelity DNA polymerase. Each final cDNA library was purified, size selected, quantified, and quality-checked by capillary electrophoresis. Sequencing was carried out on the HiSeq2000 platform (Illumina) using the 1x100 bp run mode according to the manufacturer's instructions.

2.7 Transcriptome Sequencing

In all cases, cDNA libraries were sequenced on Illumina HiSeq2000 machines multiplexed with 6 samples per lane, resulting in 12-18 million paired-end reads or 15-30 million single-end reads. The reads were delivered in fastq file format for analysis.

2.7.1 Transcriptome Data Analysis

Raw reads (both paired-end reads and single-end reads) in this work were all 100 bp in length. Fastq files, containing reads and quality scores, were first run through the FastQC software package (Andrews, 2012). FastQC runs a number of tests, the most useful of which is a look at highly repeated sequences present in the raw reads. Any 50 bp sequences occurring in more than 200 of the first 200000 reads in the fastq file were examined to determine their source. Any of these highly abundant sequences that were determined to be from a primer or amplification artifact (does not map to annotated

DM chromosomes) were then eliminated using the Trimmomatic software (Bolger, 2011). Reads were also quality trimmed by scanning each read with a 4-base sliding window and cutting when the average quality dropped below 15. Trimmed reads shorter than 25 bp were discarded because they are unlikely to map uniquely.

The output of Trimmomatic thus included intact paired-end reads and two unpaired reads files containing the orphaned reads from each end (when one read is thrown out and the other end is not) for paired-end libraries, and just trimmed reads for single-end libraries. The orphaned paired-end read files for each library were concatenated.

Next, the reads were mapped to the genome using *tophat2* and the *Ensembl* DM genome release 5.74. They were then assigned to annotated transcripts according to the transcript file from the Berkeley Drosophila Genome Project (BDGP) release 5.74 using *cufflinks*, and merged into an experiment-wide gtf file using *cuffmerge*. The gtf file was then used directly to detect differential expression with *cuffdiff2*, or indirectly to produce raw read count files suitable for analysis in *DESeq*, *DESeq2*, *edgeR*, and *SAMseq*. The raw read count files were generated using *htseq*.

After differential expression analysis, an R-based software package called *topGO* was used to carry out GO analyses on the sets of genes determined to be differentially expressed. For some of the data, pathway analysis was carried out using the R-based software package called *GAGE*. All differential expression analysis methods will be described even though the results section will focus on the *DESeq2*-based analysis. A comparison between the theory of the different methods will clear up why certain approaches were used instead of others.

Cuffdiff2

The *cuffdiff2* software attempts to measure differential expression at the *isoform* level. Because many sequencing reads can be uniquely mapped to a certain exonic location but then not assigned to a particular isoform of the corresponding gene (many isoforms share sets of exons), *cuffdiff2* models the true abundance of a transcript at the isoform level using a beta distribution (Trapnell *et al.*, 2013). It simultaneously also models the variance between biological replicates for the same isoform (dispersion) with a negative binomial distribution like most of the other differential expression detection modules.

Generally, the attempt to assign reads to single isoforms leads to a much more conservative differential expression analysis. Many true positives are potentially missed using this tool, and it also seems to struggle with outliers. The *cuffdiff2* software is really only useful for calling isoform switching events, while the other methods presented are better suited for gene-level differential expression detection. For the purposes of this work, isoform switching is not relevant, so *cuffdiff2* analysis will be excluded from the data section.

Despite its issues, *cuffdiff2* does identify some of the same differentially expressed genes as the other methods, but mainly only the ones with very high levels of significance.

DESeq

The *DESeq* package for R, part of the Bioconductor software suite, uses a much simpler approach than *cuffdiff2* to carry out statistical tests for differentially expressed genes (Anders & Huber, 2010). *DESeq* takes as input raw read counts in table format where rows correspond to genes and columns to samples. Each sample is then assigned *metadata* stipulating which condition the sample belongs to and which samples are biological replicates (Anders & Huber, 2010).

First, the size factor (sequencing depth) of each library is calculated as the median (across all genes) of the ratio of observed read counts for a gene in that library compared to the geometric mean of the observed counts for that gene in all libraries (a pseudo-reference sample based on all sequenced libraries) (Anders & Huber, 2010). These size factors provide normalisation to enable comparison of libraries sequenced to different depths (Anders & Huber, 2010).

Other statistical approaches have calculated RPKM (reads per kilobase of exon per million reads mapped), FPKM (fragments per kilobase of exon per million reads mapped), or TPM (transcripts per million) to normalise across libraries (*cuffdiff2*) or use a quantile normalisation method (*edgeR*) (Trapnell *et al.*, 2013; Robinson *et al.*, 2010). It seems that the different normalisation techniques do have serious implications for calculating differential expression. The *DESeq* method for calculating the size factors is not as sensitive to outliers or dependent on expression level as some of the other methods (Anders & Huber, 2010).

Estimating the normalised mean read count number for a particular gene under a single experimental condition is rather straightforward, just divide each measured raw count by the size factor for the library and take the average across the biological replicates of a specified condition (Anders & Huber, 2010). Here, it is possible to tell the software to ignore the known conditions of the samples, referred to as the *blind* approach. The *blind* approach uses all gene count data from all libraries to estimate the mean (and later the variance), and can be desirable in certain situations without high numbers of replicates or with very high variation present in the data (Anders & Huber, 2010). The default method takes a *per condition* approach which uses only gene counts from biological replicates to estimate the mean for each gene in each condition.

Estimating the variance is slightly more complicated. As RNA-seq typically involves less than 10 biological replicates, raw variance estimates using traditional calculations are very noisy. *DESeq* uses a generalised linear model (GLM) of the gamma family to fit a local regression on the mean/variance graph across all genes (Anders & Huber, 2010).

This allows the incorporation of information from genes expressed at similar levels in estimating their variance, thus sharing information between genes related by expression level.

Many times in the results section, the regularised log (*rLog*) transformation is used to compare gene expression levels and assess the quality and reproducibility of the sequencing data. This transformation takes the blind approach for estimating the mean and rescales the data onto a Log_2 scale by fitting a model with terms for each sample with associated coefficients estimated from the data. The resulting transformation takes the form $log_2(q_{ij}) = x_j\beta_i$. q_{ij} is proportional to the mean expression level of gene i in sample j. x_j contains the row in the gene expression table, and β_i is the vector of coefficients for gene i. This method is very close to the normalisation method described above, so it was considered the best transformation to use when assessing the quality of the data before differential expression testing.

Once the mean and variance models are fitted for each gene, testing for differential expression is carried out by computing the joint probability of observing the normalised average read counts for each gene in each condition under the null hypothesis that there is no difference in gene expression level between conditions (Anders & Huber, 2010). These joint probabilities are calculated treating each averaged, normalised gene count number in each condition as a sum of NB-distributed random variables. After adding some shot noise to the variance term and adjusting for biases, numerical calculation of p-values can be carried out and adjusted for multiple testing using Benjamini-Hochberg (Anders & Huber, 2010).

DESeq is very simple to use and fast computationally and does well controlling type-1 errors (false positives). It is more sensitive than *edgeR* for highly expressed genes and is generally less conservative than *cuffdiff2*.

edgeR

The *edgeR* package, also part of the Bioconductor software suite, provides a different approach for modeling count overdispersion sticking to the same underlying negative binomial distribution used in *DESeq* (Robinson *et al.*, 2010). Basically, rather than taking a single gene-based estimate for the variance or exclusively sharing information between genes expressed at similar levels like *DESeq*, *edgeR* uses empirical Bayes (EB) prior distributions to squeeze single gene-based variance estimates using either the information given by *all* other genes or the variance of genes expressed at similar levels (like *DESeq*) (Robinson *et al.*, 2010). The decision of which to use is governed by the degree to which the gene-wise variance conforms to the experiment-wide variance patterns and is implemented using a weighted conditional likelihood approach. In order to transform the count libraries of varying sizes to be independent and identically distributed (iid), *edgeR*

employs a quantile adjustment based on the geometric mean of the library sizes (called quantile adjusted conditional maximal likelihood or qCML) (Robinson & Smyth, 2008). This is analogous to the *size factors* from *DESeq*.

Once the iid *pseudodata* is generated, *edgeR* decides how much it can shrink the single gene-based variances by comparing them with both the experiment-wide dispersion structure and the expression-level based dispersion estimates. The result is the *shrinking* of single gene-based variances to a certain degree depending on how their variance compares with the common variance of other genes measured in the experiment (Robinson & Smyth, 2007). The effect of this is to improve the inferential power for genes with lower counts, and part of this method is incorporated by *DESeq2* to improve on the *DESeq* framework. The downside of this method is that it sacrifices some inferential power for genes with higher counts because read count variance increases with expression level (Anders & Huber, 2010). This increase in variance with expression level causes highly expressed genes to deviate more from the common dispersion estimates which are mostly based on genes with low to average expression levels. If the variance does not follow the common estimates, the shrinkage technique does not have much of an effect (Robinson & Smyth, 2007).

DESeq2

DESeq2 provides a more statistically robust and outlier insensitive version of *DESeq* but does not allow as much user control in the process (Love *et al.*, 2014). The size factors are calculated using the same median-of-ratios method that was described for *DESeq*. The gene-wise dispersions are then estimated individually and regressed against the normalised gene-wise means using a GLM of the gamma family. This regression is then used to generate a prior distribution for the dispersions of each gene. Like *DESeq*, this prior distribution takes into account the variation seen for genes expressed at similar levels. The final dispersion estimate is then calculated and provides a *squeezed* version of the gene-wise estimate by taking into account the prior information. The final estimate is called the maximum a posteriori (MAP) value. Outliers with gene-wise dispersion estimates more than 2 standard deviations away from the GLM-fitted dispersion curve are not squeezed (Love *et al.*, 2014).

A new feature in *DESeq2* is that the log fold changes (LFCs) for differential expression are also estimated and shrunk in a similar way to the dispersions. The priors are chosen from zero-centered normal distributions, and outliers with very large fold changes (usually the lowly expressed genes) are adjusted using quantile matching. The incorporation of this prior allows the the MAP estimates to squeeze LFCs for genes with either low expression values or high levels of dispersion (Love *et al.*, 2014).

The other attractive new features in *DESeq2* are a built-in independent filtering system which removes very lowly expressed genes from the analysis to avoid wasting statistical tests on low-count data and an outlier detection scheme using Cook's distance to eliminate outliers altogether from highly-replicated experiments or to ignore genes with outliers in experiments with few replicates (Love *et al.*, 2014). Because of its robustness and ease of use, *DESeq2* was chosen as the primary differential expression tool used to call gene regulation in this thesis.

SAMseq

All of the methods discussed so far for calling differential expression are heavily reliant on data following certain distributional assumptions (in most cases negative binomial distributions). In some cases, these assumptions are violated and the software will call genes differentially expressed that are not or, much worse, miss genes that are differentially expressed (Li & Tibshirani, 2013). Non-parametric methods, like *SAMseq*, have been developed that rely on rank-based test statistics and are free of assumptions about underlying distributions (Li & Tibshirani, 2013). These methods are based on permutation tests to empirically determine null distributions of the test statistics, and therefore are only applicable with a high number of biological replicates (Anders & Huber, 2010).

The first issue to be addressed when using non-parametric methods is that they can only be applied when the different samples have equal sequencing depths (Li *et al.*, 2012). *SAMseq* handles this by resampling every library to have equal sequencing depth (new depth equal to the geometric mean of sequencing depths across all libraries). This introduces some extra variation into the data, but it is averaged out by repeating the process several times and taking the average number of read counts for each gene in each sample as a final measure (Li & Tibshirani, 2013).

After computing the Wilcoxon test statistics for each gene, a permutation plug-in method is used to estimate the false discovery rates (FDRs) (Li & Tibshirani, 2013). Briefly, the same method is used to recalculate the Wilcoxon statistics with a randomly permuted read count table as many times as the user feels is necessary. The FDRs are then estimated based on a comparison between the Wilcoxon statistics with the correct read table versus the permuted read count tables (Li & Tibshirani, 2013). This method at the very least matches the performance of the parametric methods when the data adheres to the distributional assumptions (Li & Tibshirani, 2013). When outliers are present, this non-parametric method outperforms all other available analysis tools, given there are enough replicates in the experiment to use it (Li & Tibshirani, 2013).

Generally Applicable Gene Set Enrichment (GAGE) Analysis

Another piece of software used in this thesis is called the *GAGE* package and is available through Bioconductor in the R environment (Luo *et al.*, 2009). *GAGE* is capable of doing

differential expression analysis by itself, but the statistical methods are much weaker than the stand-alone packages already detailed. The advantage of *GAGE* is that it can look at more than just individual genes. Differentially expressed genes called by the other analysis tools were fed into *GAGE*, which uses the fold change measurements as test statistics to then carry out two-sample t-tests across annotated KEGG pathways (Luo *et al.*, 2009). The main interest in using *GAGE* was to look for retrograde signalling pathways that could be tied to homeostatic processes.

topGO

For GO analysis, the R package *topGO* was used to compare differentially expressed genes called by *DESeq2* with all genes having a normalised expression level greater than 10, the so-called the gene universe (Alexa *et al.*, 2006). *topGO* carries out the Fisher's exact test, which calculates p-values using the hypergeometric distribution. Many of the tests were on very small sets of differentially expressed genes so the p-values were not corrected for multiple testing.

2.8 Real Time quantitative PCR (RT-qPCR)

All RT-qPCR experiments in this thesis were done on the Roche Light Cycler 480 II machine. Single gene assays were designed with hydrolysis probes from the Roche Universal Probe Library (UPL). Briefly, the UPL contains 180 short hydrolysis probes that bind many places in the genome. The specificity is provided by designing primers around the chosen probe. For each gene of interest, two assays were designed and standard curves were run to ensur efficient amplification. For relative quantification between samples, technical triplicates were used for each gene in each sample and at least 3 housekeeping genes were run in the same plate as the candidate genes to normalise expression values (Vandesompele *et al.*, 2002).

Chapter 3

Results

This chapter will be split into four sections. The first section will quickly describe a few guidelines learned from early experimental approaches. The second section will validate the third antennal segment RNA preparation both in terms of purity and its ability to detect gene expression regulation, while highlighting some circadian-regulated genes and pathways. In the third section, the Kir2.1/Kir2.1-nc experimental system will be validated. The fourth and final section will identify gene expression regulation in third antennal segment transcriptomes after varying lengths of induction of Kir2.1/Kir2.1-nc expression in the $GH146^{+}$ cell population. To reiterate, the goal of this work was to electrically silence the postsynaptic population of $GH146^{+}$ cells and identify transcriptionally regulated genes in the third antennal segment.

3.1 Early Lessons

Preliminary experiments to study gene expression on the scale of the entire transcriptome provided some important insight into the experimental design process, and the experience gained from all of these early studies boils down to three useful guidelines.

(1) Firstly, and most importantly, is that in order to make informative gene expression comparisons between pooled groups of animals, the groups need to be as isogenic as possible. Further, if the test group of animals is overexpressing a transgene, the control group needs to also overexpress a similar transgene in the same cells and even in the same subcellular compartment. In early experiments not included in this thesis, a control group of animals isogenic to the experimental group in every way except for the insertion of a UAS-effector transgene was used (the control was just the attp2 insertion site chromosome). When it came time to analyse the data, it was difficult to determine whether the differentially expressed genes were due to the specific overexpression of the transgene or just due to overexpression of any transgene.

(2) When a company providing NGS services says the data will be ready in 4-8 weeks, expect the data in 12 weeks.

(3) Finally, when sequencing highly amplified cDNA libraries from ultra low amounts of input RNA, the purity of the RNA sample material is of the utmost importance. Contaminating tissue, genomic DNA, and even ambient pollen can have detrimental impacts on the experiment. Thus, extreme care needs to be taken when isolating the tissue and extracting the total RNA. Every effort should be taken before sequencing to make sure the tissue comes exclusively from the desired organ. This can be done by using simple reverse transcription PCR (rt-PCR) to look for genes that are expected to be expressed in the tissue and to also detect any likely sources of contamination.

3.2 The DM Third Antennal Segment Transcriptome

3.2.1 Pilot Experimental Design

The all-important first step towards studying gene expression at the whole-transcriptome level is to develop a high-purity and easily repeatable sample preparation protocol. With very small tissues that are difficult to isolate, this step can be excruciating. This section will describe a pilot experiment done to validate the tissue preparation protocol chosen and at the same time reveal novel circadian-regulated genes in the DM third antennal segment. It has long been known that there is a *peripheral* circadian clock operating in the third antennal segment that cycles odour detection sensitivities and reaches a peak sensitivity in the middle of the night (Krishnan *et al.*, 1999). To validate a sample preparation protocol, it is important to establish both reproducibility and purity, and to get a feel for the precise level of differential expression that can be detected with the material at hand.

The experimental set up was simply three biological replicates of 5-day old male Canton-S flies that were decapitated and frozen between 12:00 PM and 3:00 PM (the day or *D* group) and three biological replicates of 5-day old male Canton-S flies that were decapitated and frozen between 12:00 AM and 3:00 AM (the night or *N* group). The protocols for isolation of third antennal segments and extraction of total RNA were as described in the Methods section for $3'$ DGE studies. The resulting 6 cDNA libraries were sequenced on one lane of an Illumina HiSeq2000 machine to produce 100 bp single-end read libraries and the mapping statistics can be found in Table 3.1.

ID	Total Reads	Mapped Reads	Uniquely Mapped	AV R^2
D1	20,857,543	14,732,179(70.63%)	13,342,850(63.97%)	0.9897
D2	19,466,527	14,408,098(74.01%)	13,373,355(68.70%)	0.9931
D3	18,908,595	14,656,324(77.51%)	13,440,037(71.08%)	0.9919
N1	22,905,480	18,406,329(80.36%)	17,187,773(75.04%)	0.9955
N2	21,095,058	16,864,414(79.94%)	15,564,220(73.78%)	0.9963
N3	23,547,296	18,429,070(70.63%)	16,918,247(71.85%)	0.9963
AV	21,130,083	16,249,402(76.79%)	14,971,080(70.74%)	0.9938

Table 3.1: *D1, D2, and D3 are biological replicates of the day group and N1, N2, and N3 are biological replicates of the night group, the Average R^2 is calculated based on the transformed rLog values, AV - average*

3.2.2 Purity Assessment

To assess the RNA preparation purity, many different families of positive and negative control genes were pulled out of the dataset. Luckily, in the DM third antennal segment, there are many gene families that have been comprehensively studied and thus provide ample information to directly assess the purity of the material.

3.2.2.1 Olfactory Receptor Expression

As mentioned in the Introduction, the OR gene family is very well characterised in the third antennal segment and maxillary palp (Laissue & Vosshall, 2008), providing a useful set of genes to assess the purity of third antennal segment preparations by looking at which ORs are detected and which are not. Also, due to the hand-picked nature of the preparation, the maxillary palps are a likely source of contamination because they also detach from the head when it is frozen and shaken. Palp contamination is acceptable because the palps contain mostly ORNs, which are precisely the cells of interest. Nonetheless, a lack of palp ORs in the sequencing data is a good indicator that the picking of third antennal segments was clean.

Most of the OR expression patterns are published and were originally observed by RNA *in situ* hybridisation (Vosshall *et al.*, 1999). Subsequently, the OR gene promoters were cloned to make GAL4 drivers, some of which drive exclusively in a single ORN class and some of which drive ectopically in many classes of ORN (Couto *et al.*, 2005). The rLog-transformed expression levels of the OR gene family can be seen in Figure 3.1 divided into three subsets: those known to be expressed in the third antennal segment, those known to be expressed in the maxillary palp, and the rest (most are either larval ORs or the exact location of expression is unknown). Each subset is sorted by average rLog expression level across the six sequenced cDNA libraries.

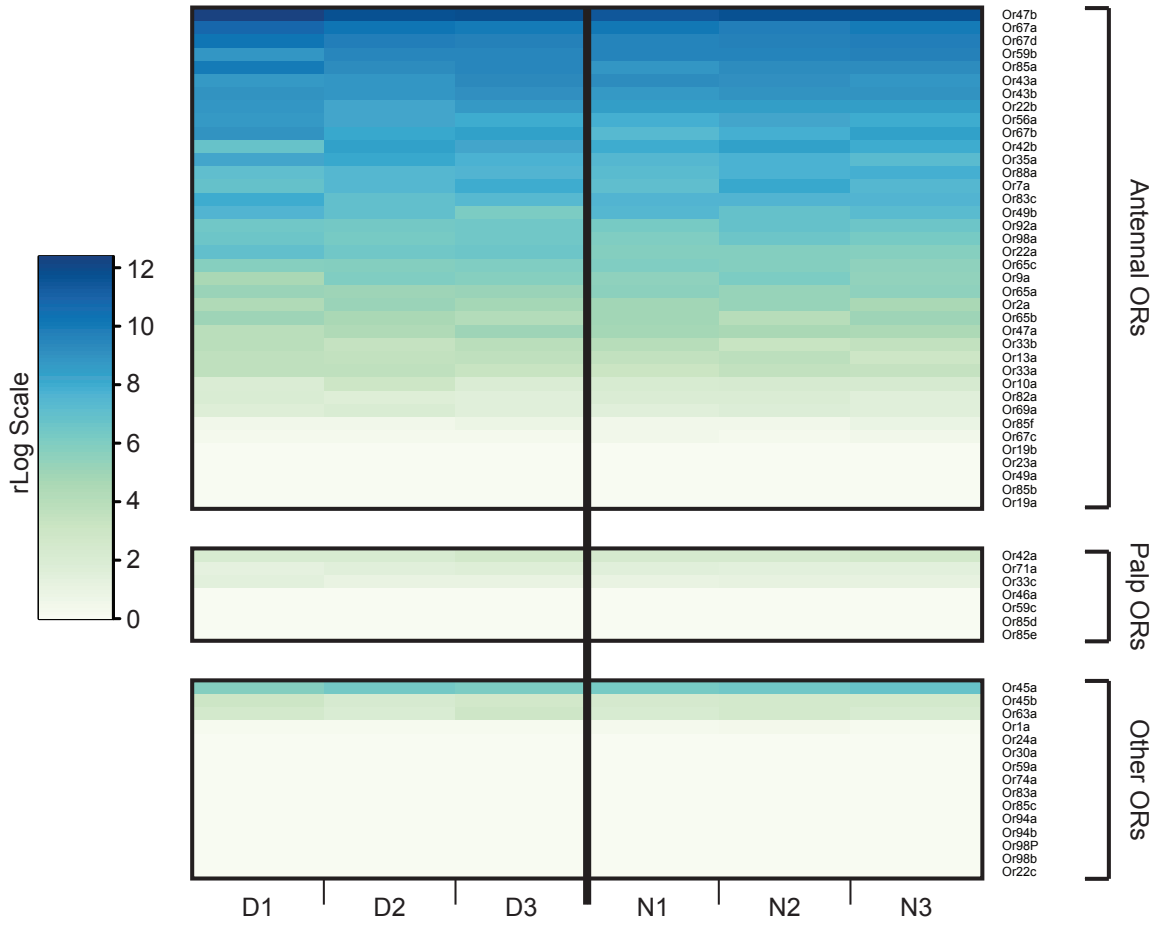


Figure 3.1: *Heatmap of expression rLog values across sequencing libraries for OR genes, annotated at right with expected expression tissue based on literature review. Or83b has not been included in the list because its expression values were so high that it obscured the scale*

The data visualised this way clearly shows strong detection of known third antennal segment ORs. For aesthetic reasons, the Or83b co-receptor was left out of this heatmap because its high expression strength obscured the rLog scale.

One evident problem with the data is that some published third antennal segment ORs were not detected. Five ORs fall into this category. Or19a and Or19b are reported to be co-expressed in trichoid sensillar ORNs ennervating glomerulus DC1, Or23a is reported to be expressed in trichoid sensilla and ennervates glomerulus DA3, Or49a is reported to be expressed in basiconic sensilla and ennervates glomerulus DL4, and Or85b is reported as only *tentatively* expressed in the third antennal segment (Laissue & Vosshall, 2008).

Ignoring Or85b because of the uncertainty about its expression pattern, the four genes that have been confidently reported in third antennal segments are probably expressed at

such low levels that the current sequencing method lacks the sensitivity to detect them. It could be that there are very few of the particular ORNs that express these ORs or that the actual expression level per neuron is very low. It is interesting that the majority of them are found in trichoid sensilla, which are supposed to be specialised for pheromone sensing neurons (Laissue & Vosshall, 2008). Later in this thesis, the high-amplification RNA-seq data will confirm that all of these ORs except Or23a are in fact present in the third antennal segment. The main takeaway here is that lowly expressed genes might not be detected in this preparation.

Looking at the other parts of the heatmap, of the seven ORs confirmed to be expressed in the maxillary palp, very low levels of three of them were detected and the remaining four were not observed. The three lowly expressed palp ORs were detected across all replicates in every sequenced third antennal segment library in this thesis, suggesting they could in fact be present in the third antennal segment. Despite these three lowly expressed palp ORs, the data suggests there is very little if any contamination of the third antennal segment tissue with maxillary palps.

Finally, from the other ORs, Or45a, Or45b, and Or63a were all detected across all sequencing libraries. This suggests that although they are thought to be larval ORs, expression actually persists into adulthood (Mathew *et al.*, 2013). Or45a and Or45b have known ligands but the ligand for Or63a is unknown (Mathew *et al.*, 2013). It does appear that Or63a is differentially expressed between European and African fly populations, along with a few other ORs which are pheromone detectors (Catalan *et al.*, 2012). This suggests that the ligand for Or63a might be an unknown pheromone.

3.2.2.2 Gustatory Receptor Expression

Slightly after the discovery of the OR family and the publication of the DM genome sequence (Adams & Celniker, 2000), an entirely different family of sensory G-protein coupled receptors (GPCRs) called the gustatory receptors (GRs) was identified in the fly genome (Clyne *et al.*, 2000). The GRs, like the ORs, are well studied and a handful have been reported to be expressed in the third antennal segment (Clyne *et al.*, 2000). In fact, some GRs have been shown to actually function as ORs (they react with volatile chemicals and are not necessarily contact taste receptors as the name *GR* implies) (Suh *et al.*, 2004). In total, the GR family is made up of ≈65 genes (Scott *et al.*, 2001).

The GRs provide another control set of genes to assay the purity of the RNA preparation protocol. Apart from the five GRs that are documented to be expressed in the third antennal segment, most have been found in the taste organs on the labellum of the fly and on the legs (Clyne *et al.*, 2000). Unfortunately, once the genes were seen to be expressed in the labellum, the reports focused on these organs and do not look (or do

not mention) whether the genes are expressed elsewhere. For this work, it means that genes not documented to be expressed in the third antennal segment might in fact be there and just have not been studied in that location. However, it does appear from the expression data displayed in Figure 3.2 that most GRs are not expressed in the third antennal segment, at least at the detection threshold in this preparation.

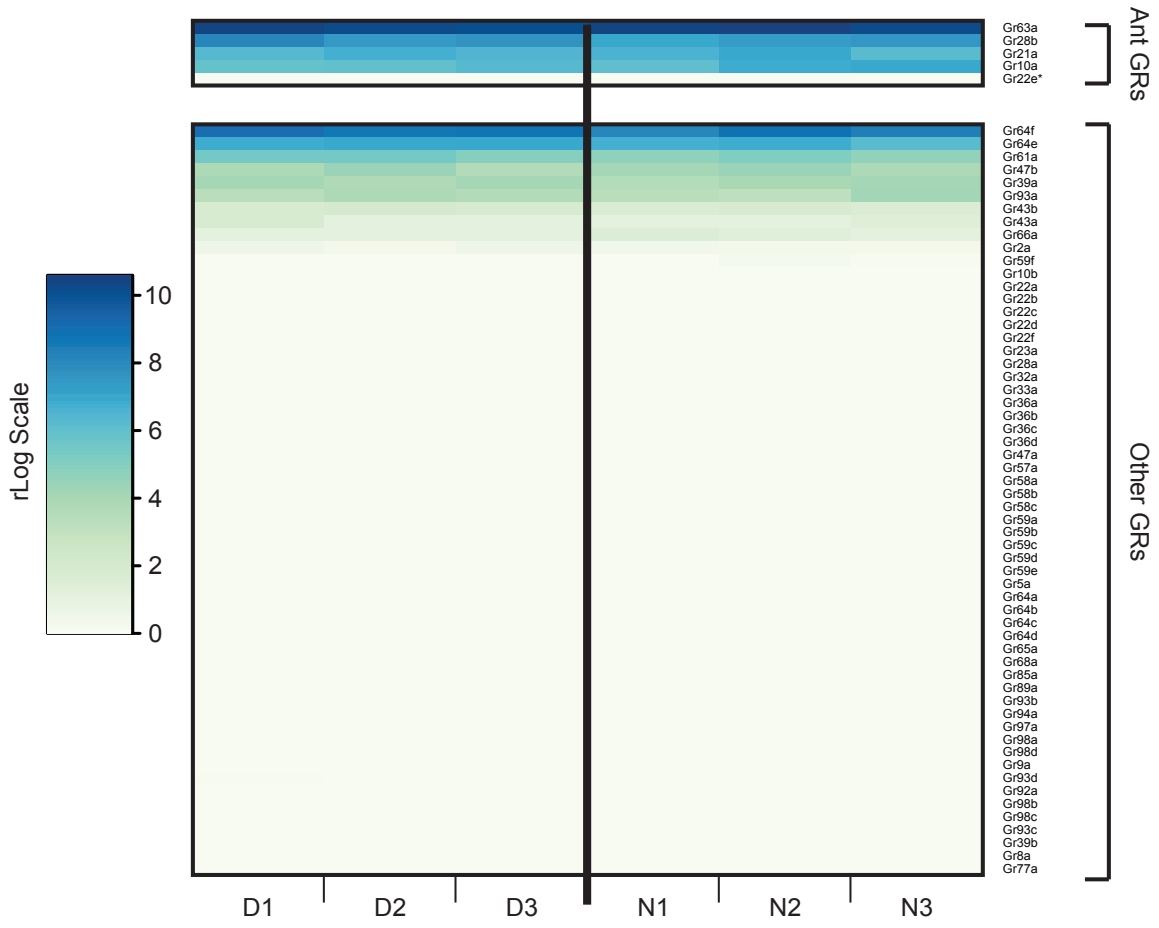


Figure 3.2: *Heatmap of expression rLog Values across sequencing libraries for GR genes, annotated at right with expected expression tissue based on literature review, * - detected with GAL4 driver but not in situ*

From the GR genes that were expected to be found, all genes except for Gr22e were detected in every sequenced library. The single study claiming Gr22e is expressed in the third antennal segment was based on a promoter-GAL4 transgenic fly that probably does not express exactly like the endogenous gene (Dunipace *et al.*, 2001). It is confirmed later that even high-amplification RNA-seq fails to detect any Gr22e sequencing reads. As for the known antennal GRs that were observed, Gr28b is known to be expressed in the arista (Ni *et al.*, 2013), attached to the third antennal segment, while Gr63a and Gr21a have been shown to function in concert as a CO_2 receptor in the third anten-

nal segment (Suh *et al.*, 2004). Thus, all of the true positive GRs were present in the sequencing data.

An interesting finding was that several GRs that have not been reported to be expressed in the third antennal segment were observed. These GRs sense sugars (Gr61a, Gr64e, Gr64f) and bitters (Gr39a, Gr93a) in the labellum (Miyamoto *et al.*, 2013) (Weiss *et al.*, 2011). The detection of these GRs in the third antennal segment is unlikely to be due to labellar contamination in the preparations because several other documented labellar GRs were *not* detected. Another novel GR in the third antennal segment, Gr47b, was also detected, but this receptor has not been characterised except for the fact that it is expressed in the larval terminal organ (TO) (Kwon *et al.*, 2011).

The presence of taste receptors in the third antennal segment, especially sugar and bitter detectors, points towards an interesting and novel view of the interaction between taste and smell in DM. If there are taste neurons in the third antennal segment, where they project and whether they have an effect on olfactory coding are important questions, especially in how that could play a role in appetitive conditioning in the T-maze. It could be that when flies are given sugar while being presented an odour, the gustatory representation in the third antennal segment can alter the olfactory quality of the odour. This means that the odour would be perceived differently than when it is presented in the absence of a gustatory stimulus (sugar). Later, the animals are usually asked to retrieve their memory by presenting only the associated odour and not the sugar (Krashes *et al.*, 2009). It could be that providing only part of the olfactory/gustatory cue at retrieval explains why immediate memory scores are lower for appetitive memory in the t-maze than they are for aversive conditioning (Perisse *et al.*, 2013).

3.2.2.3 Ionotropic Receptor Expression

It was not long after the GRs that a third group of sensory receptors, the Ionotropic Receptors (IRs), was discovered (Benton *et al.*, 2009). The IRs are evolutionarily related to the ionotropic glutamate receptor (iGluR) family, which is prevalent in both vertebrate and invertebrate nervous systems (Croset *et al.*, 2010). Despite the homology, IRs in the fly have lost their glutamate sensitivity and display high sequence divergence, and thus probably detect unique ligands giving the fly yet another ensemble of sensory transducers. There are ≈61 IRs found in the fly genome, and they can be found in a wide range of tissues (Benton *et al.*, 2009). Most of them have been studied in the third antennal segment and labellum, so they qualify as another interesting gene set to extract from the sequencing data.

As for the known third antennal segment IRs, Figure 3.3 shows that they were all detected except for Ir75d. Ir75d is reportedly expressed in sensory neurons ennervating three

different coeloconic sensilla (ac1, ac2, and ac4) in the third antennal segment (Benton *et al.*, 2009). The high-amplification RNA-seq libraries presented later detected Ir75d in all replicates, so again this gene falls below the detection threshold for this preparation.

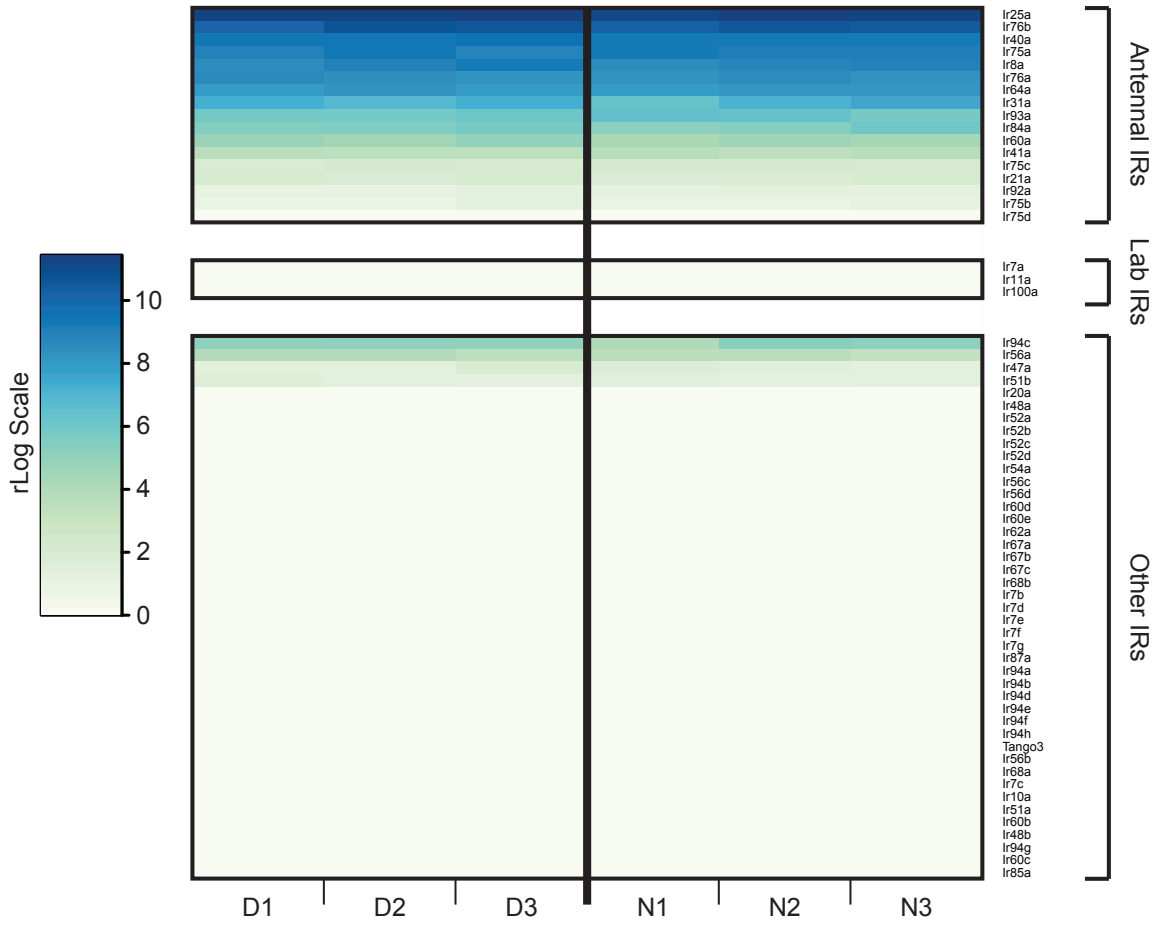


Figure 3.3: *Heatmap of expression rLog Values across sequencing libraries for IR genes, annotated at right with expected expression location based on literature review*

As for the negative controls, the three known labellar IRs were all not detected in this dataset, which would suggest that no labellar contamination is present, in agreement with the GR analysis (Benton *et al.*, 2009). Further, four IRs that have not been localised to the third antennal segment were seen in the data. Ir94c, Ir56a, Ir47a, and Ir51b were all observed at low levels of expression across all sequenced libraries. None of these have been studied explicitly.

The full complement of OR, GR, and IR genes present in the third antennal segment paints a picture of a versatile olfactory organ that not only detects odour, but also contains taste receptors, temperature sensors, and even light detectors. Perhaps referring to the third antennal segment as the *nose* of the fly is selling the organ a little bit short.

3.2.2.4 Johnston's Organ Comparison

As a final quality check, it is possible to compare the third antennal segment transcriptome to published accounts of the second antennal segment transcriptome, or more specifically to a set of genes known to be enriched in the Johnston's Organ (JO), which is housed in the second antennal segment (Senthilan *et al.*, 2012). Unfortunately, the JO data was measured with microarrays instead of deep sequencing, so the comparison has to be done on a qualitative and not a quantitative level. Nonetheless, because second antennal segments resemble third antennal segments and are physically attached to them before freezing and vortexing, it is a relevant comparison to check the purity of the preparation.

The JO is the anatomical equivalent of the ear for DM (Senthilan *et al.*, 2012). The experimental design for the JO study was to compare single gene mutant parental line control flies with a double mutant progeny line that loses the JO entirely. Using a subtraction method, they identified 100 genes that are enriched in the JO and thus should be expressed at a relatively strong level in the second antennal segment (Senthilan *et al.*, 2012).

Taking these 100 genes, Figure 3.4 shows their expression levels in the third antennal segment, which in the left heatmap are sorted by mean expression level across replicates. In the heatmap on the right, the third antennal segment expression levels have simply been re-sorted based on the levels of expression detected in the second antennal segment microarrays. If there was a large amount of second antennal segment contamination, it would be expected to see a higher correlation between these two heatmaps. Looking on the right in Figure 3.4, it is clear that many genes were not detected at all and the genes that were detected show a low correlation in expression values, thus supporting the notion that the manual third antennal segment separation can distinguish third antennal segments from second antennal segments.

After looking at several sets of genes, it is clear that the third antennal segment RNA sample preparation successfully separates and purifies adequate starting material to carry out further differential expression analysis with little to no contamination from maxillary palps, labellum, or second antennal segments. The different sensory receptor gene families that exist in DM give third antennal segments a clear and identifiable transcriptomic character, which provides a straightforward way to evaluate the quality and purity of the preparations. In and of itself, this data is a very useful resource for groups interested exclusively in the machinery of sensory reception via the third antennal segment.

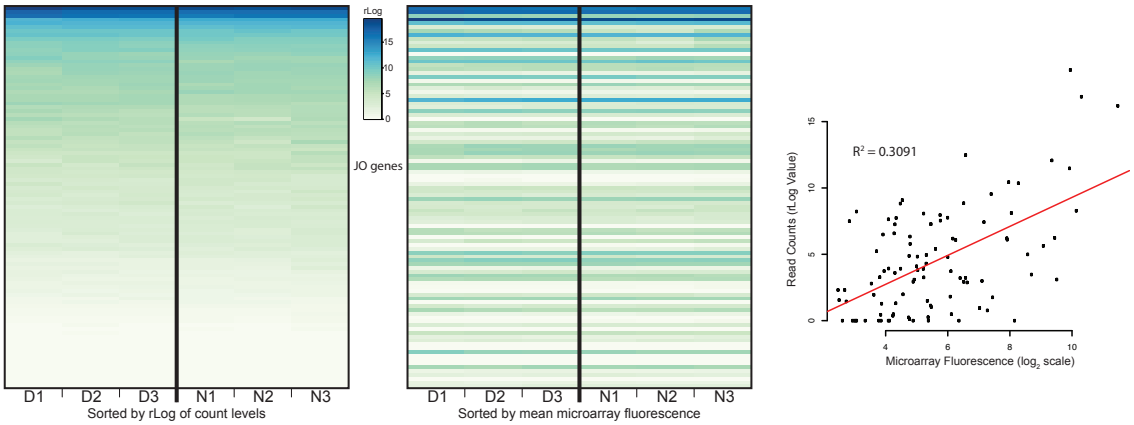


Figure 3.4: *Two heatmaps showing expression of the 100 JO enriched genes sorted by expression value in third antennal segments on the left and sorted by JO expression from the microarray data on the right, on the far right is a scatterplot of the same data showing the low correlation. Many genes are not detected at all in third antennal segments.*

3.2.3 Circadian Gene Expression Regulation

To verify that the preparation and sequencing can identify differentially expressed genes in the third antennal segment, the D and N sequencing libraries were next compared using the many differential expression analysis tools. *Cuffdiff2*, *DESeq*, *edgeR*, and *DESeq2* were all tested in this process and *DESeq2* provided the most outlier-insensitive, robust method. Before presenting the *DESeq2* analysis, it is important to note that this experiment was not meant to be an exhaustive study of circadian cycling in the third antennal segment. It was only hoped to prove that true differential expression can be detected with the RNA extraction and sequencing methods at hand, and that the canonical circadian clock genes (*per*, *tim*, *vri*, etc.) would be observed in the results. To keep the costs of the pilot experiment low, it was decided only to measure the extremities of the circadian time course. A more thorough approach would have been to take material at several time points and use the dedicated statistical tools to detect rhythmicity of expression (Hughes *et al.*, 2012).

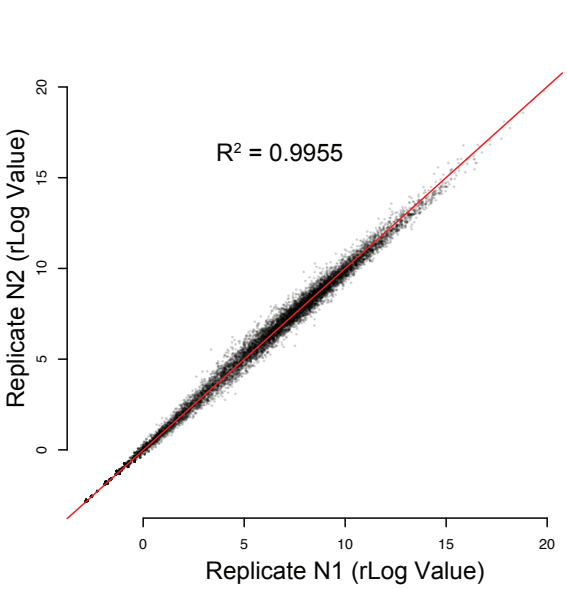
There have been at least 6 reports published since 2001 measuring circadian gene expression cycling in fly heads or brains at the whole-transcriptome level (Ceriani *et al.*, 2002; Claridge-Chang *et al.*, 2001; Lin *et al.*, 2002; McDonald & Rosbash, 2001; Ueda *et al.*, 2002; Hughes *et al.*, 2012). The first 5 reports all used microarrays to study cycling transcription in whole heads, and showed a remarkable variability in the number and identity of cycling genes discovered. This was largely due to differences in RNA preparation protocols and inadequacies of the early statistical methods. Meta-analyses were next carried out on the datasets which standardised analysis approaches, and unified the evidence for which transcripts actually cycle (Keegan *et al.*, 2007; Wijnen *et al.*, 2006). The two meta-analyses are therefore much more useful as a comparison tool than any of

the individual datasets. More recently, a report was published measuring the circadian transcriptome in dissected adult brains using RNA-seq with whole-transcriptome amplification of low amounts of total RNA (Hughes *et al.*, 2012). There is thus ample data available to use as a benchmark for this experiment.

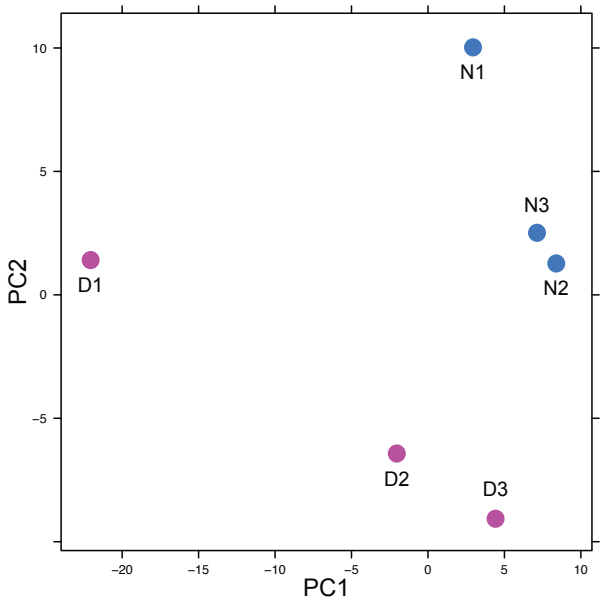
Differential expression analysis between D and N third antennal segment cDNA libraries using the *DESeq2* analysis software is presented in Figure 3.5. Biological replicates in this dataset show a very high correlation (Figure 3.5a and Table 3.1) and the D and N libraries cluster based on their first two principal components (Figure 3.5b). The differential expression analysis identified almost all of the known circadian clock genes including vri, clk, per, and timeless with FDR-adjusted p-values < 0.01 (Table 3.5c). In total, 47 differentially expressed genes (this group includes core clock genes just mentioned) were detected with FDR-adjusted p-values < 0.01, 81 were detected with FDR-adjusted p-values < 0.05, and 99 were detected at an FDR-adjusted significance level < 0.10. Figure 3.5c shows a heatmap of the 47 genes detected at an FDR adjusted p-value < 0.01.

Looking more closely at all genes found to be significantly differentially expressed at an FDR-adjusted p-value < 0.10, it is interesting to see that very few of the genes seen cycling in third antennal segments have been detected to cycle in heads or dissected brains. Figure 3.5d shows a Venn diagram depicting the sets of genes called differentially expressed by the meta-analysis done on the microarray data from fly heads, the RNA-seq analysis done on dissected brains, and the third antennal segment sequencing data presented in this thesis. The lack of overlap of all three datasets is particularly striking. Only 6 genes are found in all three datasets (they happen to be the core molecular clock components), and no more than 16 genes overlap between any two of the lists of cycling genes.

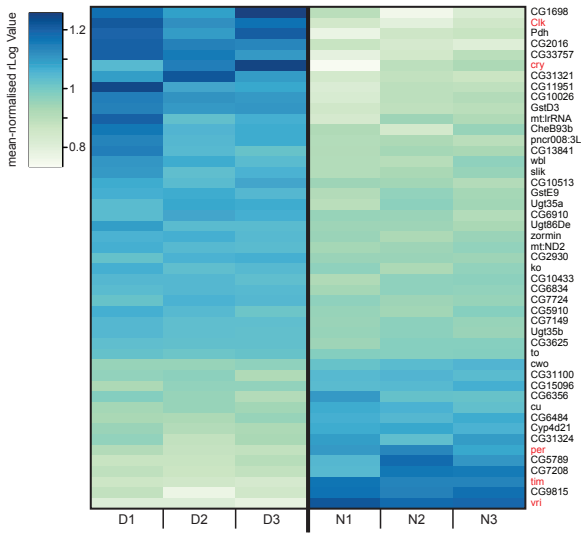
This could be for several reasons. First, different statistical methods were used in each case. The newest methods are proving to be more robust, so this is probably less of a problem. Second, different tissues presumably have different cycling transcripts due to epigenetic factors, and it could be that only the core clock genes cycle in all tissues. In the studies being compared, the preparations are not on unique tissues persay, as the head includes the brain and third antennal segments. Based on the anatomy, there should be more overlap between the third antennal segments and head (as well as the head and brain) than there is between the brain and third antennal segments. This is exactly what is seen in the Venn diagram where only 2 genes (CG17691 and Ugt86De) are shared uniquely between the brain and the third antennal segments, whereas 9 genes are shared between the head and the brain and 10 genes are shared between the head and the third antennal segment.



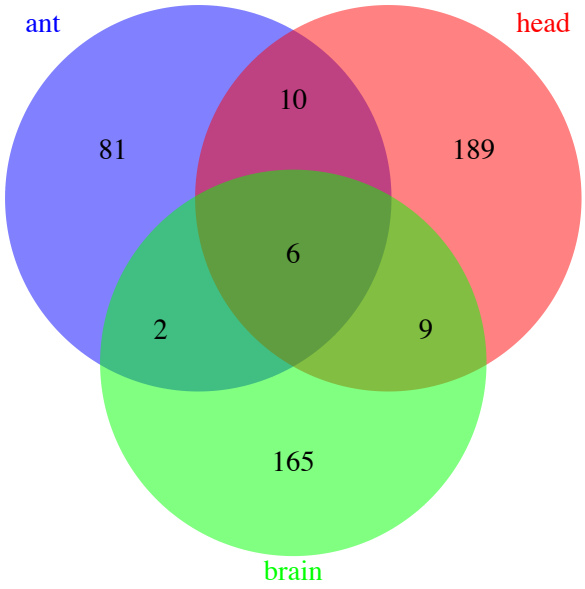
(a) *A typical comparison between normalised expression levels in biological replicates*



(b) *The D and N replicates cluster well in the first 2 principal components*



(c) *Heatmap showing the differentially expressed genes at an FDR < 0.01, core clock genes are in red*



(d) *Venn diagram showing overlap between third antennal segment, head, and brain circadian-regulated genes*

Figure 3.5: *Differential expression analysis using* DESeq2

The obvious next question is what exactly the cycling genes in the third antennal segment are doing if they are different than the cycling genes in other tissues. From physiological studies, DM are known to have a strong circadian cycle of odour detection sensitivity in the third antennal segment as measured by electroantennogram (EAG) responses to odours (Krishnan *et al.*, 1999). In the same study, it was demonstrated that a functional molecular clock in the central brain does not drive the antennal circadian cycle (mutants where the *per* gene is only functional in the central clock cells do not exhibit circadian cycling in odour detection sensitivity), thus suggesting the presence of an autonomous peripheral clock in the third antennal segment.

A brief look at the enriched GO terms in the 99 cycling genes detected at FDR-adjusted p-values < 0.10 can begin to answer the nature of the cycling genes. The biological process terms (Table 3.2) are almost all related to rhythms, homeostasis, and response to environmental chemicals. The molecular function GO terms (Table 3.3) mostly have to do with transmembrane transport and fit nicely with the known cycling of odour detection sensitivity.

GO ID	Term	Fisher's p-value
GO:0009648	photoperiodism	1.3E-05
GO:0043279	response to alkaloid	2.9E-05
GO:0007623	circadian rhythm	3.1E-05
GO:0048511	rhythmic process	3.6E-05
GO:0048148	behavioural response to cocaine	4.3E-05
GO:0042752	regulation of circadian rhythm	4.7E-05
GO:0007622	rhythmic behavior	5.0E-05
GO:0045475	locomotor rhythm	8.7E-05
GO:0008062	eclosion rhythm	1.0E-04
GO:0001659	temperature homeostasis	1.2E-04

Table 3.2: *Top ten most significantly enriched BP GO terms*

GO ID	Term	Fisher's p-value
GO:0015291	secondary active transmembrane transporter activity	3.70E-04
GO:0046983	protein dimerization activity	4.30E-04
GO:0015020	glucuronosyltransferase activity	1.04E-03
GO:0022857	transmembrane transporter activity	1.62E-03
GO:0005215	transporter activity	2.68E-03
GO:0022891	substrate-specific transmembrane transporter activity	3.83E-03
GO:0008509	anion transmembrane transporter activity	4.26E-03
GO:0022804	active transmembrane transporter activity	4.58E-03
GO:0046943	carboxylic acid transmembrane transporter activity	6.97E-03
GO:0005316	phosphate:sodium symporter activity	7.45E-03

Table 3.3: *Top ten most significantly enriched MF GO terms*

Finally, a quick look at the KEGG pathways that are linked to the differentially expressed genes with unadjusted p-values < 0.05 (688 total genes) points to 6 regulated pathways with pathway-wide q-values < 0.360. Unadjusted p-values were used to determine which genes were used as input into this analysis because if many genes in the same pathway have significant unadjusted p-values, it is unlikely to be a coincidence. All pathways reported in Table 3.4 have > 10 regulated genes.

During the day, drug metabolism and the metabolism of xenobiotics are significantly upregulated, as well as glutathione metabolism and starch and sucrose metabolism. At night, the ribosome pathway as well as the oxidative phosphorylation pathway are significantly upregulated. It is hard to speculate exactly what functions all these pathways might play based on sequencing data alone because it could be that the cycling of the resulting proteins is delayed from the cycling of the mRNA transcripts by post-transcriptional or post-translational mechanisms. It does make sense however that cell maintenance pathways like ribosomal construction and oxidative phosphorylation are activated at night whereas metabolic functions dealing with energy and interaction with volatile exotic chemicals are highly active during the day.

Pathway	Up/Down	# of Genes	q-value
Drug metabolism - cytochrome P450	Up	19	9.80E-03
Metabolism of xenobiotics by cytochrome P450	Up	18	9.80E-03
Glutathione metabolism	Up	15	3.20E-02
Starch and sucrose metabolism	Up	10	5.42E-02
Ribosome	Down	23	1.94E-02
Oxidative phosphorylation	Down	12	3.60E-01

Table 3.4: *Regulated pathways with* up *referring genes up-regulated during the day and* down *meaning genes up-regulated at night*

The data presented confirm the ability of the RNA preparation to detect differential expression. Several interesting circadian cycling gene expression modules were detected, some of which could provide clues as to why the third antennal segment is so sensitive to odours in the middle of the night (Krishnan *et al.*, 1999).

3.3 The Kir2.1/Kir2.1-nc System

The Kir2.1/Kir2.1-nc system designed in this thesis was created in the same vein as the dORK/dORK-nc system also available in DM (Nitabach *et al.*, 2002). The difference is that dORK-nc does not localise to the plasma membrane, which is why it cannot conduct current (personal communication with M. Nitabach). For the purposes of this work, it was absolutely vital that the non-conducting channel localise exactly like the conducting channel. Transcriptome measurements can be very sensitive, and differential localisation of overexpressed genes between test cells and control cells could in itself produce gene expression differences which would confound the effects of the silencing. This section will validate a *transcriptomics-friendly* system for conditionally silencing populations of neurons.

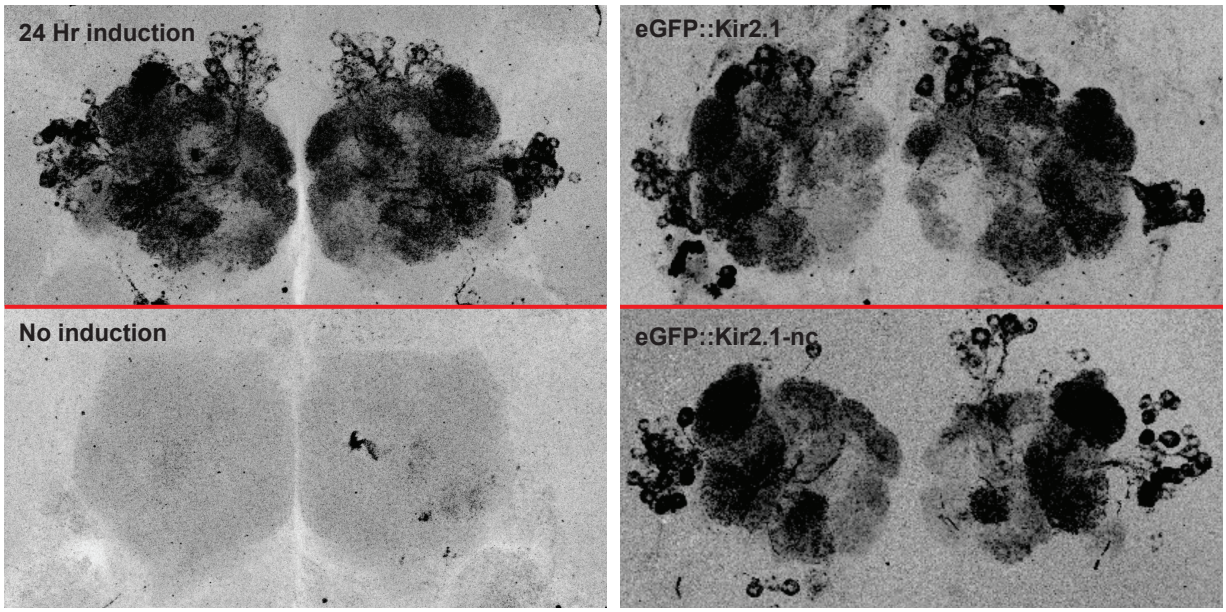
3.3.1 Kir2.1: Expression and Localisation

Kir2.1 (KCNJ2 by gene name) is an inwardly rectifying potassium channel found in humans that is widely used by neuroscientists as a neural silencer because it hyperpolarises neurons and decreases their input resistance, which renders the neuron less excitable (Johns *et al.*, 1999). It was first identified because mutations in the Kir2.1 coding region result in a variety of disorders effecting the QT interval in the cardiac waveform (Priori *et al.*, 2005). Because of its importance in human health, many reports have been published on its function. One study looked at several single amino acid mutations that completely abolish current flow through the channel without affecting its localisation (Haruna *et al.*, 2007). The G146S mutant (referred to in this thesis as Kir2.1-nc) does exactly that, and was cloned for *in vivo* testing.

When flies possessing a UAS-eGFP::Kir2.1 transgene are crossed to the GH146-GAL4 driver line, all of the progeny are dead by early larval stages. However, crossing GH146-GAL4 driver lines to flies harbouring the UAS-eGFP::Kir2.1-nc transgene results in a tube full of happy, healthy progeny. This suggests that the Kir2.1-nc mutation does abolish current flow *in vivo*, in line with what has been shown in cultured cells (Haruna *et al.*, 2007).

Next, the $GAL80^{ts}$ induction system was used in an attempt to turn on expression of the eGFP-Kir2.1/eGFP-Kir2.1-nc transgenes in adult flies, the goal being to rescue the larval lethality and conditionally silence the $GH146^{+}$ cell population in adults. One, then two, and then three doses of tubulin-promoter driven $GAL80^{ts}$ were tested before finally establishing that three doses is enough to rescue the larval lethality of eGFP::Kir2.1 overexpression. Figure 3.6a shows that after 24 hours of induction at 31°C, there is clear expression of eGFP-Kir2.1 that is not present if the flies are left at 21°C. Figure 3.6b

demonstrates that there is no obvious difference in the localisation of the conducting and non-conducting eGFP::Kir2.1 channels. Everything seems to agree with the cell culture report (Haruna *et al.*, 2007), but just because proteins localise does not mean that they function. To be thorough, the system was validated both behaviourally and physiologically.



(a) *GH146-GAL4, 3 x tubp-GAL80^{ts} drives UAS-eGFP::Kir2.1 induction* **(b)** *eGFP::Kir2.1 and eGFP::Kir2.1-nc localise identically in GH146^{+} cells*

Figure 3.6: *Validation of the Kir2.1/Kir2.1-nc induction system*

3.3.2 Behavioural Validation

The most obvious first pass experiment was to validate the Kir2.1/Kir2.1-nc system in the same way that the dORK/dORK-nc system had been validated, in the pigment dispersing factor (*pdf*) circuit. Silencing the *pdf*$^{+}$ neurons has been shown to abolish the circadian locomotory rhythms of the fly (Nitabach *et al.*, 2002). Figure 3.7 shows that adult-onset expression of Kir2.1 using the pdf-GAL4 driver and a single dose of tubulin-promoter driven *GAL80^{ts}* abolishes locomotor rhythms when the flies are left in constant darkness at 31°C. Three-day old female flies that had been raised at 21°C were placed in thin-walled glass tubes and monitored for 5 days at 31°C in DD. Figure 3.7 shows the locomotory behaviour for days 3 and 4. Importantly, the Kir2.1-nc flies remain rhythmic throughout the entire period. It is worth pointing out that the rhythms, especially in the morning, are muted because of the high temperature. Taken together, these data support the functionality of the Kir2.1/Kir2.1-nc system in a tried and trusted

behavioural paradigm.

Next, it was important to assess the impact of adult-onset eGFP::Kir2.1 expression in $GH146^{+}$ cells, the site at which it is expressed for the transcriptomics studies of the third antennal segment. To get an idea of what to expect, GH146-GAL4 flies were first crossed to flies bearing a UAS-$shibire^{ts}$ transgene and naive odour preferences were measured for three odours (methyl salicylate, ethyl butyrate, and geranyl acetate). After being raised at 21°C and spending a total of 35 minutes at 31°C (20 minutes pre-heating and 15 minutes on the airflow in the multiplex), flies expressing $shibire^{ts}$ in the $GH146^{+}$ cells were found to be anosmic for all odours tested (Figure 3.8b).

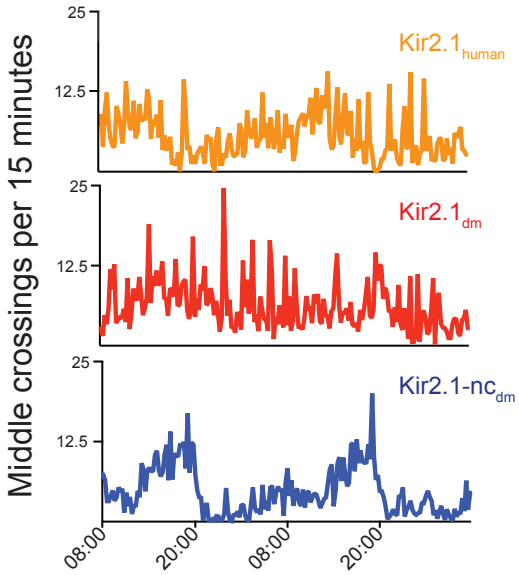
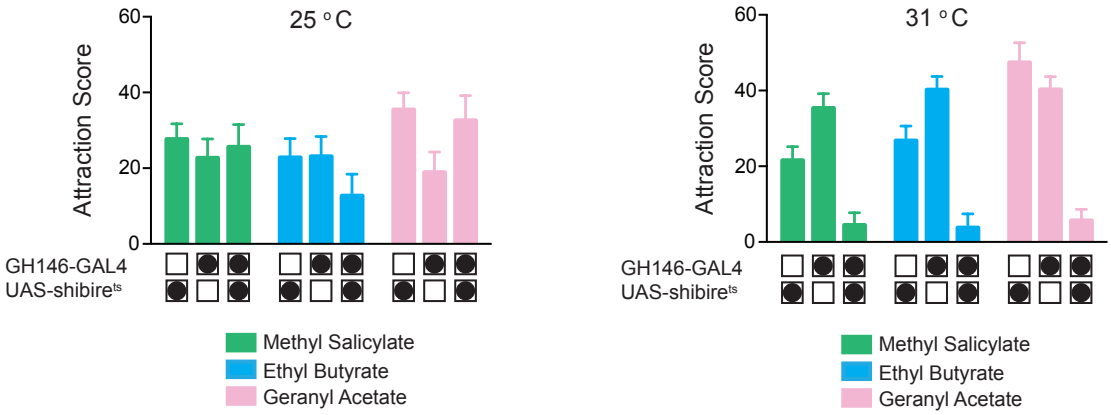


Figure 3.7: *Adult-onset expression of Kir2.1 in the pdf^{+} circuit abolishes circadian locomotory behaviour (n=16 for each)*

Both parental lines crossed to CS demonstrated clear preferences when exposed to the same conditions, which in this case showed them to be attracted to all odours tested. Flies that were switched to 25°C instead of 31°C showed normal odour preferences (Figure 3.8a). This suggests that immediate blockade of the $GH146^{+}$ population of cells is sufficient to seriously hinder the fly's ability to detect odours. It is not quite possible to say that detection of all odours is abolished because only 3 were sampled here, but comprehensive studies of naive olfactory behaviour have reported similar anosmia when blocking $GH146^{+}$ cells with $shibire^{ts}$ (Parnas *et al.*, 2013).

After determining the impact of immediate $GH146^{+}$ blockade, similar experiments were carried out using the Kir2.1/Kir2.1-nc system instead of $shibire^{ts}$. The only difference was that induction of gene expression with $GAL80^{ts}$ required longer times at the elevated temperature. Both GH146-GAL4, 3 x tubp-$GAL80^{ts}$, UAS-eGFP::Kir2.1 and GH146-GAL4, 3 x tubp-$GAL80^{ts}$, UAS-eGFP::Kir2.1-nc flies were kept at 31°C for 12, 24, 48, and 96 hours in food vials and their naive odour preferences were then measured for the same odours as in the $shibire^{ts}$ experiments. All behaving flies were age-matched to be 5 days old, so the induction was simply started earlier for longer inductions.

The results show many complex trends and need to be addressed one odour at a time. Figure 3.9a shows the odour preferences when flies were presented with methyl salicylate (MS) after the different induction periods. MS is detected in the environment only



(a) *Expression of shibire^{ts} in the GH146^{+} cell population at the permissive temperature has no effect on olfactory preferences*

(b) *Expression of shibire^{ts} in the GH146^{+} cell population at the restrictive temperature abolishes olfactory preferences*

Figure 3.8: *The effects of shibire^{ts} in GH146^{+} cells*

by a single, highly specific ORN class (Or10a) and thus is considered a very *narrow* odour (Hallem & Carlson, 2006). This is true when thinking about ORN space, but it is possible that a narrow ORN activation does not result in a narrow PN activation due to the lateral excitation known to occur in the antennal lobe (Olsen *et al.*, 2007). Or10a provides the input for glomerulus DL1 in the antennal lobe, which is ennervated by *GH*146^{+} PNs (Wong *et al.*, 2002). With increasing induction time and thus increasing levels of expression of eGFP::Kir2.1 in the postsynaptic PNs, attraction to MS steadily decreased until 48 hours, when the flies began to show no preference between the odour and mineral oil. The eGFP::Kir2.1-nc flies show consistent attraction across all induction periods, but the magnitude of attraction is dependent on how long they have spent at the elevated temperature. This dependence seems to show increasing attraction up until 24 hours of induction and then steady decline of attraction after that.

Figure 3.9b tells a much different story for a more *broad* odour, ethyl butyrate (EB). EB is detected in the environment by at least 8 different ORN classes (Hallem & Carlson, 2006). Whether this translates into a broader representation at the PN level is not clear because it has not been explicitly measured. What is clear however, is that the same eGFP::Kir2.1 flies who lose the ability to detect MS, seem to maintain a relatively constant attraction to EB at all induction times. Once again, the magnitude of attraction for the eGFP::Kir2.1-nc flies is dependent on how long they have spent at the elevated temperature, with the similar trend of attraction increasing until 24 hours and then decreasing again at 48 and 96 hours of induction.

In the case with geranyl acetate (GA), another *narrow* odour that is specific for a single ORN class (Or82a) (Laissue & Vosshall, 2008), the situation is different yet again. Or82a-expressing ORNs form synapses with PNs postsynaptic to the *GH*146^{+} glomerulus VA6.

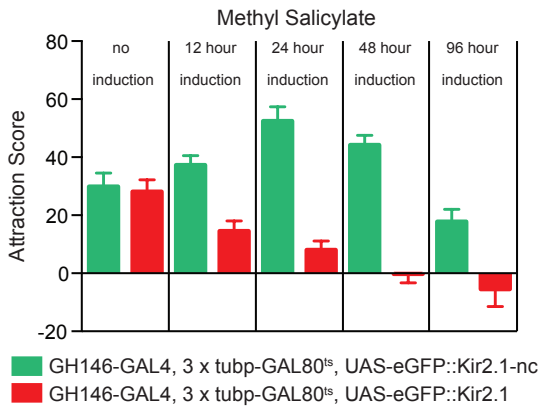
It seems that attraction to GA is very quickly reduced with short induction periods and then attraction returns as time passes even with continued induction. At 96 hours, the flies have almost recovered the full attraction shown by their eGFP::Kir2.1-nc controls. There is again a time dependence of attraction in the eGFP::Kir2.1-nc controls depending on how long the flies had been kept at the elevated temperature.

Taken together, the results are very puzzling. Because GH146-GAL4 does not cover all PNs, and does not drive at a uniform strength in $GH146^{+}$ PNs, it is hard to explain what exactly is happening. An optimistic interpretation would point out that it seems for two of the three odours that after 96 hours of induction there are only very small differences between the eGFP::Kir2.1 and eGFP::Kir2.1-nc fly's attraction scores, which would suggest homeostatic adjustment. However, it is also true that after 96 hours at the elevated temperature the eGFP::Kir2.1-nc flies are less attracted to all odours tested, and it could be the case that the control flies are losing interest because they are stressed from the heat, and that is what brings the preferences closer together. But the fact that the eGFP::Kir2.1 flies are *more* attracted after 96 hours in some cases goes against that logic. In summary, conclusions beyond the fact that electrical silencing of the $GH146^{+}$ cell population has an effect on olfactory preferences that is dependent on the identity of the odour and the induction period would be over-interpreting the information at hand.

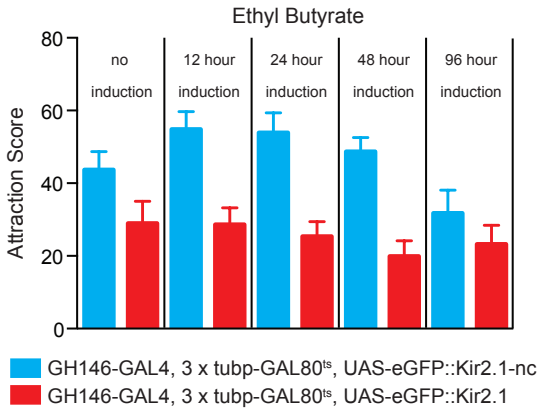
In an attempt to at least clear up some of the questions raised by these experiments, another silencing tool, tetanus toxin light chain (TeTxLc), which cleaves synaptobrevin, was tested in the same way as the eGFP::Kir2.1/eGFP::Kir2.1-nc system. The idea was to see if another neural silencing system also dependent on $GAL80^{ts}$-mediated induction of gene expression would impact behaviour similarly to the Kir2.1/Kir2.1-nc system.

Kir2.1 acts directly on the excitability of the neurons, while TeTxLc only stops their output (Keller *et al.*, 2002). Homeostatic processes at the DM larval NMJ have been shown to depend on the neuron's ability to depolarise, which should not be altered in the case with TeTxLc overexpression (Paradis *et al.*, 2001). If the effects of TeTxLc look more like $shibire^{ts}$, which also blocks release at the output of the neuron, it would suggest that the reason for the residual olfactory preferences for some odours in eGFP::Kir2.1 flies was because of a homeostatic adjustment as a result of altered neuronal excitability.

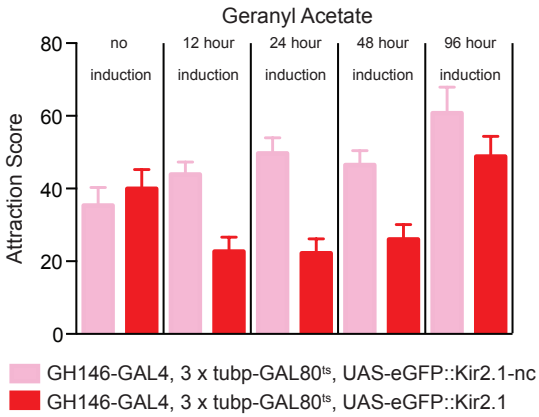
The behavioural data presented in Figure 3.10 for TeTxLc overexpression also tells a complex story. Methyl salicylate attraction is abolished quickly and completely using TeTxLc as a neural silencer (Figure 3.10a). Ethyl butyrate attraction slowly decreases with increasing TeTxLc expression (Figure 3.10b) whereas geranyl acetate attraction is not significantly changed by TeTxLc (Figure 3.10c). The presence of residual olfactory preferences with some odours in this case looks similar to the eGFP::Kir2.1 situation, suggesting that the longer induction periods or the efficacy of silencing are more likely



(a) *Increasing induction time of eGFP::Kir2.1 expression in GH146$^+$ cells decreases attraction to methyl salicylate*



(b) *Increasing induction time of eGFP::Kir2.1 expression in GH146$^+$ cells only slightly decreases attraction to ethyl butyrate*

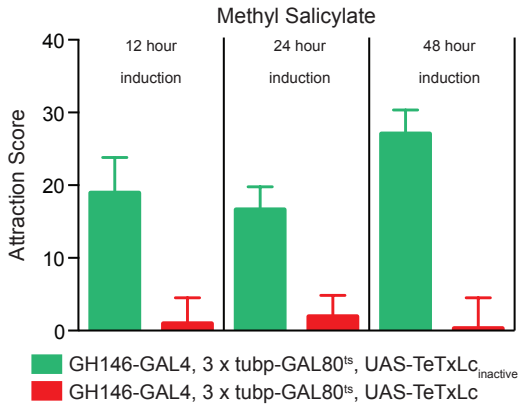


(c) *Increasing induction time of eGFP::Kir2.1 expression in GH146$^+$ cells only slightly decreases attraction to geranyl acetate, and some attraction is recovered after 96 hours of induction*

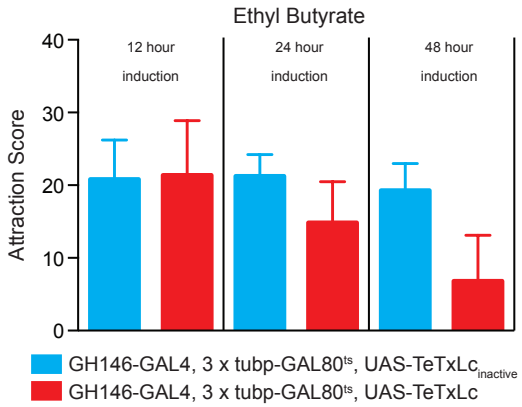
Figure 3.9: *The effects of eGFP::Kir2.1 expression in the GH146$^+$ cell population*

to be the reasons that flies retain some of their olfactory preferences in Kir2.1/TeTxLc-mediated silencing approaches but not with *shibire^{ts}*. It does not rule out however that time for homeostatic adjustment is the main factor in both situations.

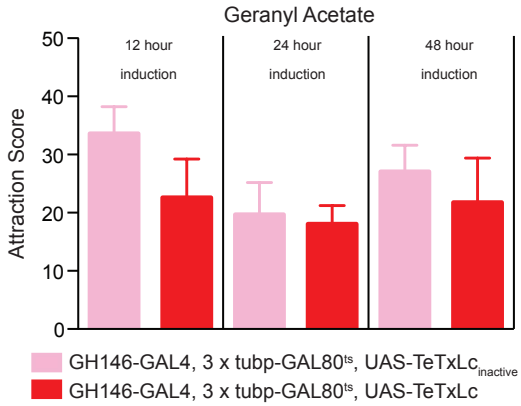
Although the behavioural experiments do raise a lot of questions about exactly what is happening and why the flies lose their ability to detect some odours (but not others) with some silencers (but not all silencers), they do answer the question of whether Kir2.1 has an impact in PNs. The answer is yes and it appears the impact is dependent on the time period in which Kir2.1 expression begins and/or the strength of that expression. It also is clear from this behavioural work that it is the PN expression of GH146 enhancer trap pattern that mediates the change in olfactory choice behaviour and not the glial sheath expression. Glial sheath silencing would not have an odour-specific effect.



(a) *12 hours induction time of TeTxLc expression in $GH146^{+}$ cells is sufficient to abolish attraction to methyl salicylate*



(b) *Increasing induction time of TeTxLc expression in $GH146^{+}$ cells decreases attraction to ethyl butyrate*



(c) *Increasing induction time of TeTxLc expression in $GH146^{+}$ cells does not effect attraction to geranyl acetate*

Figure 3.10: *The effects of TeTxLc expression in the $GH146^{+}$ population of cells*

3.3.3 Electrophysiological Validation

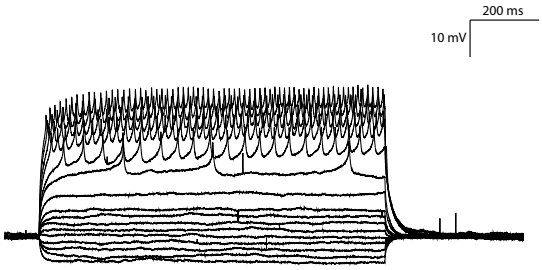
To take a very close look at the effects of adult-onset eGFP::Kir2.1 expression in PNs, intracellular recordings were taken from five day old adult female flies that had been kept at the induction temperature (31°C) for 24 hours. Whole-cell recordings were taken from $GH146^{+}$ eGFP::Kir2.1 and eGFP::Kir2.1-nc expressing PNs while injecting increasing current steps to assay the intrinsic excitability of each neuron. To interpret the results of this experiment, it is important to consider that the $GH146^{+}$ population of PNs encompasses many different classes of PN with different arborisation patterns and thus intrinsically different passive membrane properties (input resistances and membrane time constants). As has been shown, PNs with larger dendritic fields have lower input resistances and faster membrane time constants (Kazama & Wilson, 2008). Because this data includes recordings from randomly sampled $GH146^{+}$ PNs, there is added variability sim-

ply due to the the fact that the PNs are of different classes. It would have been better to use a more specific driver that directed gene expression to a specific PN class. Then, it would have been easier to compare the eGFP::Kir2.1 and eGFP::Kir2.1-nc measurements with less variation caused by broad PN sampling. However, after recording from a few randomly sampled $GH146^{+}$ PNs, it became clear that the neural silencing effects of Kir2.1 are strong enough that such targeted experiments are unnecessary.

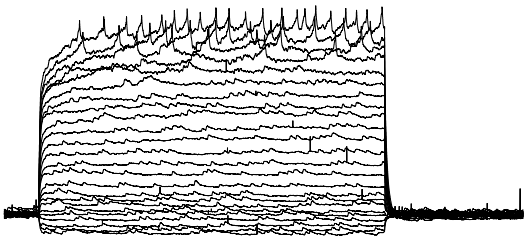
Taking a look at the data, Figure 3.11a shows an example of how a control (eGFP::Kir2.1-nc) PN behaved when injected with increasing current steps. Very quickly, the neuron begins to fire repeatedly with increasing frequency as the amplitude of the current step increases. In the case with an eGFP::Kir2.1 PN, the threshold at which firing begins to occur is greatly increased and the frequency of the evoked spiking after crossing this threshold decreases significantly (Figure 3.11b). As can be seen in Figure 3.11c and Figure 3.11d, the input resistance drops considerably as well as the membrane time constant when expressing eGFP::Kir2.1 versus the non-conducting control. Another important observation is that the firing properties of eGFP::Kir2.1-nc expressing PNs look very similar to PNs expressing only a GFP marker (personal communication with colleagues).

A final and convincing picture of the effects of eGFP::Kir2.1 in PNs can be seen by looking at the smoothed cumulative distribution functions (CDFs) in Figure 3.11e. The green CDF represents the percentage of recorded eGFP::Kir2.1-nc PNs firing action potentials at a minimum of 1 Hz for the corresponding current injection strengths on the x-axis. The red CDF presents the same relationship for the eGFP::Kir2.1 expressing PNs. There is a clear separation between the amount of injected current required to induce action potential firing in the two populations of neurons. The EC_{50} values plotted on the graph specify how much current is required to induce action potentials in half of the population of PNs at a 1 Hz threshold. Figure 3.11f shows that for increasing frequency thresholds (in 5 Hz intervals from 5 Hz to 40 Hz), the EC_{50}'s of both populations of neurons follow a similar linear upward trend, but of course separated by a large offset referred to in the figure as the *Kir effect*.

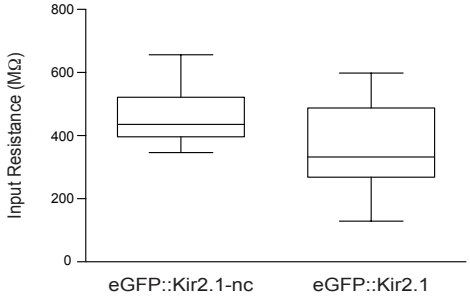
Having evaluated the Kir2.1/Kir2.1-nc system from both a behavioural and electrophysiological standpoint, it is clear that the system does function as designed. Neurons expressing eGFP::Kir2.1 require more input current to evoke a response, but can still fire action potentials if receiving large inputs. Increasing the induction time seems to increase the silencing effect, as seen from the behavioural experiments, and the control flies maintain olfactory function under all conditions. The system is applicable to silence the $GH146^{+}$ cell population and deep sequence the transcriptome of the third antennal segment.



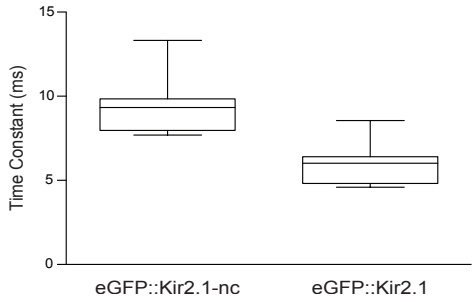
(a) *Current injection evokes firing in an eGFP::Kir2.1-nc expressing PN*



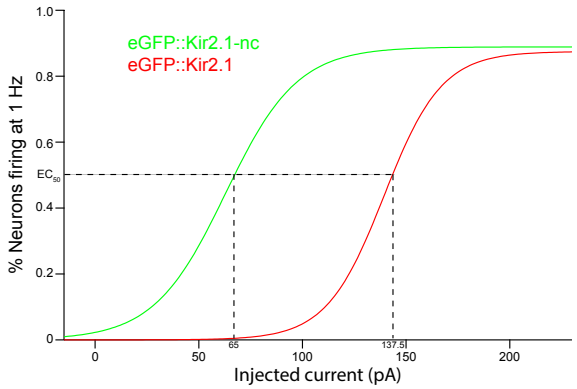
(b) *Current injection struggles to evoke firing in an eGFP::Kir2.1 expressing PN*



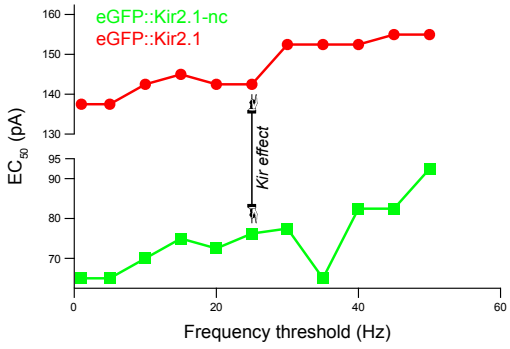
(c) *eGFP::Kir2.1 lowers input resistance, n=8*



(d) *eGFP::Kir2.1 shortens the time constant, n=8*



(e) *Smoothed cumulative distribution functions of eGFP::Kir2.1-nc and eGFP::Kir2.1 PNs for a 1 Hz firing threshold plotted against injected current, n=8*



(f) *EC_{50} plotted against frequency threshold for eGFP::Kir2.1-nc and eGFP::Kir2.1 PNs*

Figure 3.11: *Physiological properties of eGFP::Kir2.1 and eGFP::Kir2.1-nc PNs, n=8*

3.4 Identification of the Homeostatic Machinery

Given that the behavioural effects of eGFP::Kir2.1 expression in $GH146^{+}$ cells can be seen after only 12 hours of induction, in order to maximise the chances of detecting any gene expression regulation, the third antennal segment transcriptome was measured after induction time courses of 12, 24, 48, and 96 hours. To control for age differences, all sequenced flies were age-matched at 5 days old (like in the behaviour). To do that, the induction had to be started at different times in adulthood. Age of expression onset therefore could be the determining factor as opposed to induction strength, and it is not possible from the current data to discriminate between the two.

The results presented in this section will be divided by the two types of sequencing experiments that were carried out to profile the third antennal segment transcriptome. The first section of experiments covers induction times of 12, 24, and 48 hours and these libraries were prepared and sequenced with the exact same methods as the CS libraries analysed in the pilot experiment. In the 96 hour induction experiments, the number of third antennal segments per preparation was reduced from 250 to 50, and the smaller RNA samples were amplified using the SMARTer Ultra Low RNA Kit for Illumina Sequencing discussed in the Methods section. More biological replicates were sequenced in this case because of anticipated higher levels of variability as a result of the amplification.

3.4.1 3′ DGE Results

Flies of the following two genotypes were collected as soon as possible after eclosion and kept at 21°C until it was the right time to switch them to a 31°C induction incubator. The tissue was then processed as described in the Methods section.

w; GH146-GAL4, tubp-$G80^{ts}$/tubp-$G80^{ts}$; UAS-eGFP::Kir2.1/tubp-$G80^{ts}$
w; GH146-GAL4, tubp-$G80^{ts}$/tubp-$G80^{ts}$; UAS-eGFP::Kir2.1-nc/tubp-$G80^{ts}$

The resulting third antennal segment total RNA preparations were then sent to MWG Eurofins and sequenced on an Illumina HiSeq2000 machine using single-end 100 bp reads and the 3′ DGE tag-based technology. The mapping statistics can be found in Table 3.5.

Two notable things are reflected in these mapping statistics. First, the percentages of uniquely mapped reads are much lower than the libraries that had been sequenced in the pilot experiment. Because these preparations were done in exactly the same way, it is hard to explain why the yield of unique reads is so much lower than the pilot study. Because MWG Eurofins insists that everything was done as before, the only plausible explanation is that the larger scale (more libraries) of these experiments meant the preparations had lower RNA yields and that is reflected by a decrease in the number

ID	Total Reads	Mapped Reads	Uniquely Mapped	AV R^2
KIR-48-1	28,004,316	18,815,975(67.19%)	16,487,747(58.88%)	0.9953
KIR-48-2	17,661,910	11,229,908(63.58%)	9,851,297(55.78%)	0.9951
KIR-48-3	27,158,074	15,643,356(57.60%)	13,588,111(50.03%)	0.9957
KIRnc-48-1	40,636,012	26,365,750(64.88%)	23,831,661(58.65%)	0.9770
KIRnc-48-2	26,993,657	13,344,752(49.44%)	11,474,648(42.51%)	**0.9612**
KIRnc-48-3	32,525,963	18,944,019(58.24%)	17,227,196(52.96%)	0.9787
KIR-24-1	31,966,708	17,804,932(55.70%)	15,075,834(47.16%)	0.9911
KIR-24-2	23,942,092	13,998,611(58.47%)	12,528,015(52.33%)	0.9935
KIR-24-3	25,561,056	15,590,606(60.99%)	14,058,188(55.00%)	0.9918
KIRnc-24-1	18,699,305	9,330,034(49.90%)	8,331,334(44.55%)	**0.9882**
KIRnc-24-2	11,172,429	5,655,191(50.62%)	5,049,936(45.20%)	0.9906
KIRnc-24-3	15,315,506	8,767,785(57.25%)	8,020,166(52.37%)	0.9906
KIR-12-1	29,660,291	15,855,218(53.46%)	14,082,433(47.48%)	0.9939
KIR-12-2	28,904,100	15,808,260(54.69%)	14,110,730(48.82%)	0.9943
KIR-12-3	28,938,152	15,549,498(53.73%)	13,323,889(46.04%)	0.9935
KIRnc-12-1	22,023,606	10,963,417(49.78%)	9,701,601(44.05%)	0.9941
KIRnc-12-2	15,850,079	8,864,492(55.93%)	7,922,966(49.99%)	0.9947
KIRnc-12-3	11,834,929	6,237,128(52.70%)	5,546,857(46.87%)	0.9946
AV	24,269,344	13,820,496(56.34%)	12,234,034(49.93%)	0.9897

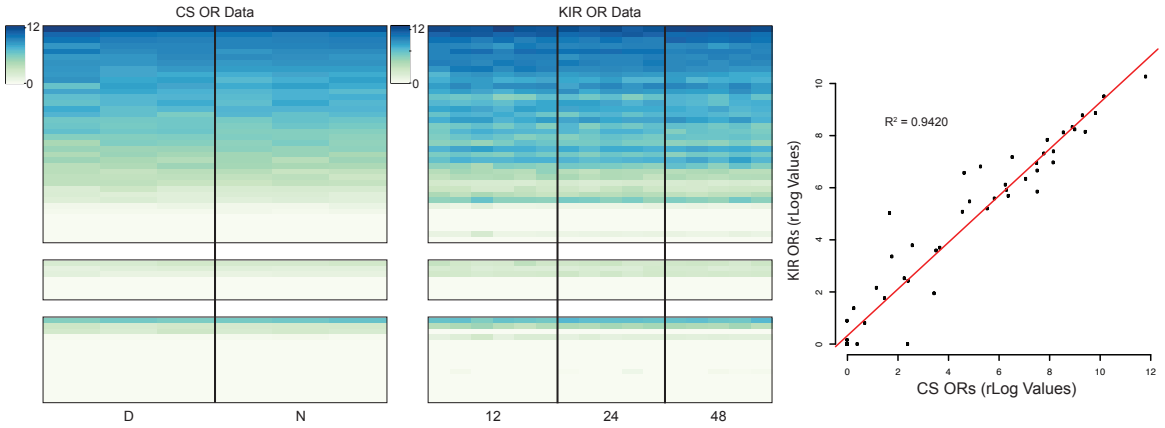
Table 3.5: *Three biological replicates for each of three conditions (12, 24, and 48 hours of induction), the Average R^2 is calculated based on the rLog-transformed values, AV stands for average, numbers in bold identify low-correlation replicates*

of uniquely mapped reads. Second, there are two libraries, KIRnc-48-2 and KIRnc-24-1, that have comparatively low average R^2 values with their biological replicate samples after rLog transformation (<0.99 correlation). These samples were removed from further analysis. All other samples had average correlations above the 0.99 threshold after rLog transformation.

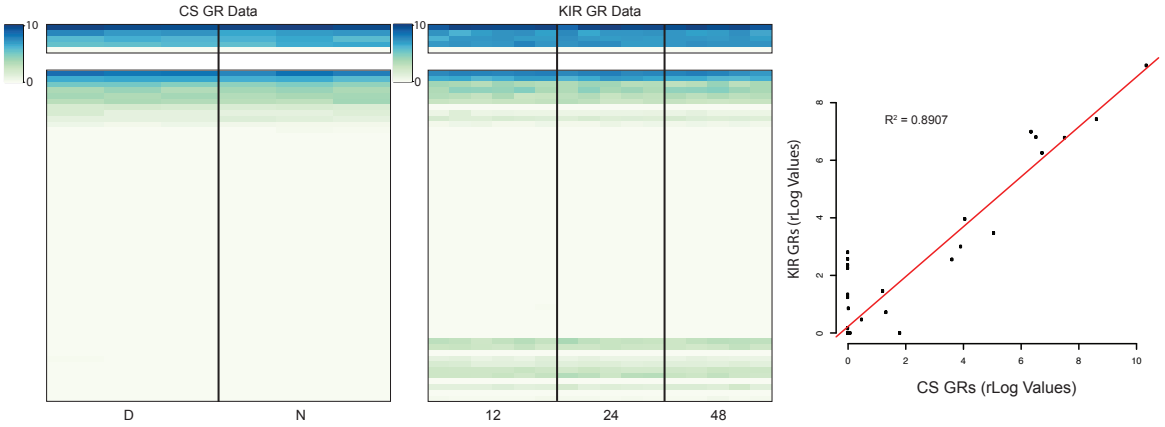
3.4.1.1 Purity Assessment

Having gone to great lengths in the pilot experiment to prove the purity of the preparations, the same litmus tests were used with this *KIR* data as well. Expression values for OR, GR, and IR genes in this experiment can be found in Figure 3.12, but the sort order of the genes has been maintained from the pilot study to show the similarities between the data. Miniaturised versions of Figures 3.1, 3.2, and 3.3 (the expression data from the pilot experiment) are also presented to the left of the *KIR* heatmaps for ease of comparison. Correlations between rLog expression values in this *KIR* data with the pilot experiment *CS* data were very high for every subset of sensory receptor genes.

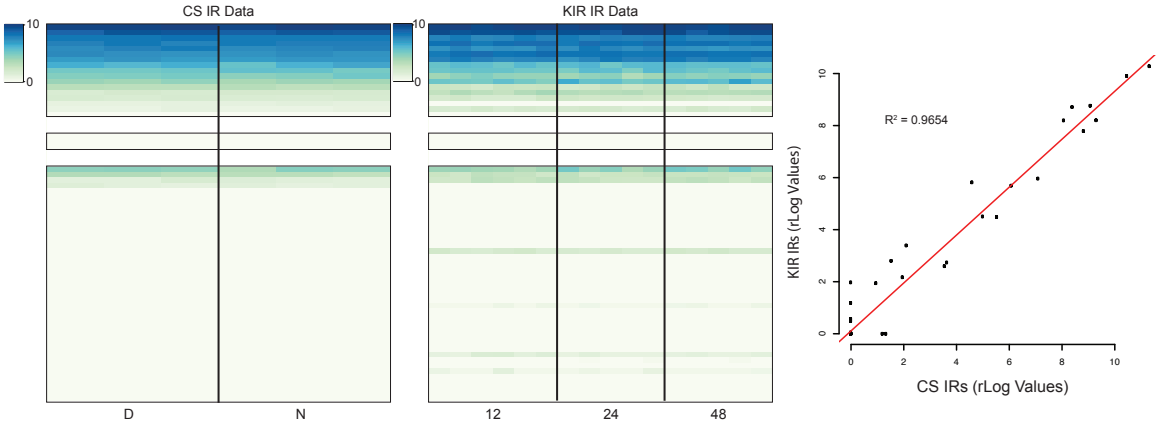
The JO gene set also confirms that this data is comparable to the CS data and is pictured in Figure 3.13.



(a) *Comparison of OR expression in CS sequencing libraries and KIR sequencing libraries*



(b) *Comparison of GR expression in CS sequencing libraries and KIR sequencing libraries*



(c) *Comparison of IR expression in CS sequencing libraries and KIR sequencing libraries*

Figure 3.12: *The sequencing data passes all of the same quality checks that the pilot experiment passed. The results are highly correlated even though different genotypes were being tested under different environmental conditions (heat treatments)*

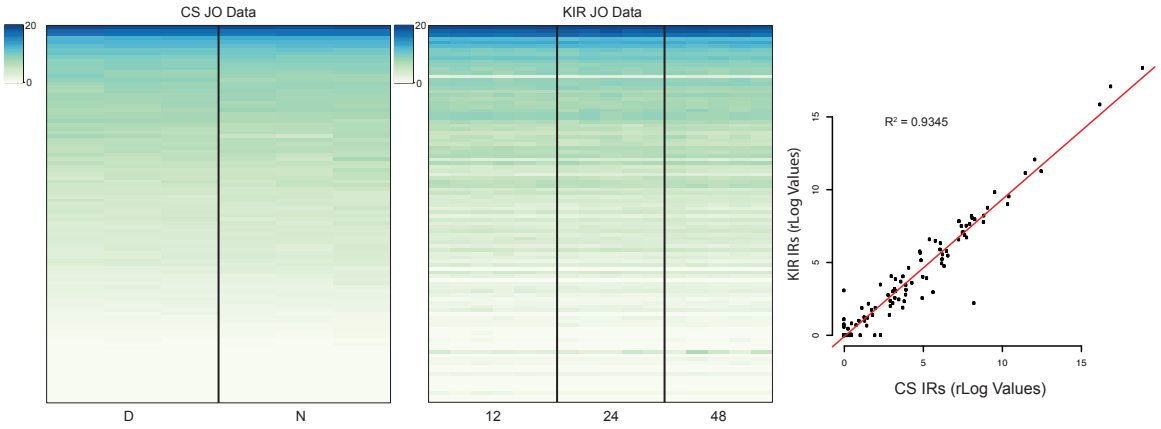


Figure 3.13: *Comparison of JO expression in CS sequencing libraries and KIR sequencing libraries*

Therefore, the quality of the preparations is as good as it was in the the pilot experiment. One interesting observation is that some genes that were not detected before, especially in the GR gene set, have now been detected. This could be due to the different genotypes used in the KIR preparations or due to the heat treatment which might increase expression of certain genes. The fact that they were detected across all libraries and not just a portion of them means it is probably not due to contamination. Further, it is clear that the new GRs and IRs detected are not genes that are known to be found in the labellum or maxillary palps. The data passes the quality checks and is appropriate for differential expression analysis.

3.4.1.2 Differential Expression

Sex

An interesting first question that can be addressed from the 12 hour dataset only is what the gene expression differences are in the third antennal segment for males and females. The 24 hour and 48 hour data are not ideal to address this question because all of the 24 hour libraries are from female third antennal segments and all of the 48 hour libraries are from male third antennal segments, thus length of time at the induction temperature will confound true sex differences.

A simple differential expression analysis on the 12 hour data looking only at sex differences with *DESeq2* finds 254 differentially expressed genes at an FDR-adjusted p-value < 0.10. The details of these genes will not be discussed further because it is not crucial for this thesis, but what is important is that some of the genes do provide a good set of positive controls because there are ORs and odorant binding proteins (OBPs) already reported to be differentially expressed between males and females (Zhou *et al.*, 2009).

Table 3.6 shows the 7 differentially expressed ORs/OBPs that were detected at an FDR-adjusted p-value < 0.10. Or67a and Or85a were both found to be upregulated in female third antennal segments, and have thus far not been reported to be sexually dimorphic. A reasonable hypothesis is that females are more sensitive to certain food odours because of the higher nutritional needs associated with producing and laying eggs. The detection of the olfactory machinery known to be sexually dimorphic provides within-run validation of the ability to detect true positives in this data.

Gene	Log2 FC	FDR	Description
Obp19a	-0.893	1.74E-15	pheromonal (Zhou *et al.*, 2009)
Or47b	0.664	1.22E-05	fru^+ ORNS (Dobritsa *et al.*, 2003)
Or85a	-0.797	2.41E-05	food odours
Or22b	-0.807	2.00E-03	sexually dimorphic (Laissue & Vosshall, 2008)
Obp57c	-0.271	4.00E-03	pheromonal (Zhou *et al.*, 2009)
Obp56d	-0.327	1.50E-02	pheromonal (Zhou *et al.*, 2009)
Or67a	-0.378	7.50E-02	food odours

Table 3.6: *Differentially expressed genes between the sexes from the 12 hour induction group, FC = fold change, positive FC means up-regulated in males, FDR = false discovery rate*

12 Hour Induction

For this experiment, the two genotypes of flies were placed at the elevated temperature for 12 hours (2:00 AM until 2:00 PM) starting in the evening of their fourth post-eclosion day. The flies were decapitated between 2:00 PM and 3:00 PM on day 5. Three biological replicates were sequenced for each genotype, one of which was from male flies and the other two replicates were from female flies (as mentioned in the sex section). Sex differences were explicitly controlled for in the regression model entered into *DESeq2* for differential expression analysis. Table 3.7 gives a list of the differentially expressed genes based on the genotype differences alone with an FDR-adjusted p-value < 0.20.

As demonstrated by the GO BP term analysis shown in Table 3.8, many of the differentially expressed genes are involved in cell fate commitment and organ morphogenesis. This is interesting as there is nothing in the literature documenting adult neurogenesis or gliogenesis in the third antennal segment. There are also regulated genes involved in synapse organisation. Because only 29 genes were entered into the enrichment analysis, the p-values are presented as unadjusted p-values given by a Fisher's exact test.

24 Hour Induction

For this experiment, the same two genotypes of flies were placed at the elevated temperature for 24 hours (2:00 PM until 2:00 PM) starting in the afternoon of their fourth post-eclosion day. The flies were decapitated between 2:00 PM and 3:00 PM on day 5. Three biological replicates were sequenced for each genotype, all of which were from

Gene	Log_2 FC	FDR	Description
CG42502	0.529	1.76E-05	no functional information
CG11550	-0.386	2.72E-03	CRAL-TRIO domain
noe	0.319	1.10E-02	no functional information
A2bp1	0.316	1.34E-02	mRNA recognition, transcription factor binding, neuronal development
IM4	-0.328	1.34E-02	defense response, cell morphogenesis
Trc8	-0.549	1.34E-02	zinc ion binding, protein binding, negative regulation of growth
baz	0.302	1.34E-02	protein kinase C binding, multicellular development
dro5	-0.296	1.34E-02	defense response
sxe2	-0.305	1.34E-02	phospholipase activity
GstE3	0.321	3.53E-02	glutathione transferase activity
CG34309	-0.308	4.68E-02	no functional information
sff	0.404	4.70E-02	kinase, ATP binding
CG6490	0.471	5.17E-02	neuronal remodeling, axon pruning
hth	0.225	5.17E-02	transcription factor, organ morphogenesis
CG11873	0.352	6.61E-02	no functional information
pros	0.350	9.32E-02	transcription factor, neuroblast differentiation
CG15201	-0.246	1.07E-01	no functional information
Ank2	0.317	1.15E-01	membrane anchoring, neuron cellular homeostasis
CG17838	0.291	1.15E-01	mRNA binding, RNA recognition
CG6421	-0.326	1.15E-01	lysozyme activity
Hcf	0.522	1.15E-01	transcription factor, chromatin remodeling
Kr-h2	-0.351	1.15E-01	Krupper homolog, transcription factor
Pis	0.267	1.15E-01	phosphatidylinositol synthase
crol	0.251	1.15E-01	metal ion binding, chromatin silencing
sdt	0.315	1.15E-01	guanylate kinase activity, anatomical structure morphogenesis
VhaM8.9	-0.367	1.33E-01	Wnt signaling, endosomal transport
RpL38	-0.260	1.40E-01	structural constituent of ribosome
Acon	-0.341	1.62E-01	aconitase hydratase activity
CG3907	-0.245	1.64E-01	no functional information

Table 3.7: *Differentially expressed genes after 12 hours of induction, FC = fold change, positive FC means upregulated in eGFP::Kir2.1-expressing flies, FDR = false discovery rate adjusted p-value*

GO ID	GO term	All	DE	p-val
GO:0060071	Wnt receptor signalling pathway, planar cell polarity pathway	7	2	2.20E-04
GO:0035567	non-canonical Wnt receptor signalling pathway	8	2	2.90E-04
GO:0016332	establishment or maintenance of polarity of embryonic epithelium	9	2	3.80E-04
GO:0045186	zonula adherens assembly	9	2	3.80E-04
GO:0031345	negative regulation of cell projection organisation	14	2	9.40E-04
GO:0034333	adherens junction assembly	14	2	9.40E-04
GO:0050808	synapse organisation	152	4	1.40E-03
GO:2000027	regulation of organ morphogenesis	70	3	1.51E-03
GO:0090175	regulation of establishment of planar polarity	20	2	1.94E-03
GO:0010628	positive regulation of gene expression	171	4	2.17E-03
GO:0042659	regulation of cell fate specification	23	2	2.57E-03
GO:0009653	anatomical structure morphogenesis	1218	10	2.89E-03
GO:0010453	regulation of cell fate commitment	25	2	3.03E-03
GO:0016055	Wnt receptor signaling pathway	90	3	3.10E-03
GO:0034332	adherens junction organisation	26	2	3.28E-03
GO:0045197	establishment or maintenance of epithelial cell apical/basal polarity	26	2	3.28E-03

Table 3.8: *GO BP terms enriched in the 29 differentially expressed genes detected after 12 hours of induction, p-values are Fisher's exact test derived and have not been adjusted for multiple testing, ALL means the number of times a GO term appears in the gene universe list, DE means the number of times it appears in the differentially expressed gene list*

female flies. As mentioned in the purity section, one of the Kir2.1-nc replicates in this group did not cluster well with the others and was removed from the analysis.

Table 3.9 gives a list of the genes that were detected as differentially expressed after 24 hours of induction at an FDR-adjusted p-value < 0.20. The large fold changes seen here are primarily due to the low expression levels of these genes. Because they are expressed very lowly and have high variability between the two eGFP::Kir2.1-nc biological replicates used in this experiment, the differential expression calls can not be taken too seriously.

Gene	Log_2 FC	FDR	Description
ninaE	-2.640	3.36E-07	rhodopsin-like G protein-coupled receptor
Pdh	-2.630	3.36E-07	photoreceptor dehydrogenase
Yp1	-2.169	1.12E-04	lipase
Ag5r	-1.781	2.70E-02	allergen, CAP domain

Table 3.9: *Differentially expressed genes after 24 hours of induction, FC = fold change, positive FC means upregulated in Kir2.1 flies, FDR = false discovery rate adjusted p-value*

None of the genes detected in the 12 hour induction experiment were detected at the

same level of significance in this 24 hour group. It could be due to a loss of power from only having two Kir2.1-nc replicates, or it could be that the gene expression changes are transient. To check what happened to the genes that came up as significant in the 12 hour dataset, these genes were examined in the 24 hour data to see whether they are detected as regulated in the same direction, but at a lower significance level.

Table 3.10 shows that 13 of the genes called differentially expressed in the 12 hour group have calculated FC values in the same direction. There are 8 genes in which the 24 hour data has a FC in the opposite direction and 8 genes in which the 24 hour data has FCs close to 0. The genes that agree on direction and magnitude of FC (those marked yes in Table 3.10) tend to have lower FDR-adjusted p-values in the 12 hour data, which suggests they are true positives and the 24 hour data lacks the power to detect them.

48 Hour Induction

In this group, the two genotypes of flies were placed at the elevated temperature for 48 hours (2:00 PM until 2:00 PM) starting in the afternoon of their third post-eclosion day. The flies were decapitated between 2:00 PM and 3:00 PM on day 5. Three biological replicates were sequenced for each genotype, all of which were from male flies. One of the Kir2.1-nc replicates in this group did not cluster well with the others and was removed from the analysis.

Table 3.11 gives a list of the genes detected as differentially expressed in this experiment at an FDR-adjusted p-value < 0.20. There is no overlap between this group of differentially expressed genes and the genes from the 12 and 24 hour induction experiments. It could again be due to a loss of power from having only two Kir2.1-nc replicates, but the fact that highly significant genes are detected that were not seen in the other experiments suggests that after 48 hours, new processes have begun to take shape in the third antennal segment.

To illustrate the processes in which these differentially expressed genes function, Table 3.12 shows a GO enrichment analysis for their BP terms. Again, the p-values are not adjusted for multiple testing due to such a small amount of genes being entered into the analysis.

It appears that the prominent regulated processes after 48 hours of induction are either immune-related or involved in the stress response. The two Hsp genes that were detected are known for their transcriptional responses to heat but it is important to point out that both populations of flies were exposed to the same temperature regimen. Therefore, due to the nature of the controls, it is likely that Kir2.1 expression in $GH146^{+}$ cells is the cause of the upregulation and not just a general heat shock response.

Gene	Log_2 FC 24	Log_2 FC 12	Agreement?
CG42502	0.411	0.529	yes
CG11550	-0.267	-0.386	yes
IM4	-0.338	-0.328	yes
Trc8	-0.412	-0.549	yes
baz	0.123	0.302	yes
GstE3	0.168	0.321	yes
CG3907	-0.111	-0.245	yes
CG6490	0.202	0.471	yes
Hcf	0.352	0.522	yes
CG11873	0.254	0.352	yes
pros	0.196	0.350	yes
Ank2	0.237	0.317	yes
CG17838	0.157	0.291	yes
noe	0.028	0.319	maybe
A2bp1	-0.028	0.316	maybe
dro5	0.039	-0.296	maybe
CG34309	-0.088	-0.308	maybe
Kr-h2	-0.019	-0.351	maybe
Pis	0.035	0.267	maybe
crol	-0.080	0.251	maybe
CG15201	0.000	-0.246	maybe
hth	-0.103	0.225	no
CG6421	0.104	-0.326	no
sdt	-0.296	0.315	no
VhaM8.9	0.164	-0.367	no
RpL38	0.229	-0.260	no
Acon	0.129	-0.341	no
sxe2	0.206	-0.305	no
sff	-0.106	0.404	no

Table 3.10: *Comparison between fold changes detected after 12 hours and 24 hours of induction for genes called significantly differentially expressed after 12 hours, FC = fold change,* yes *means that the gene has a non-zero FC in the same direction at 12 and 24 hours,* maybe *means that the gene has a fold change close to 0 after 24 hours which was non-zero after 12 hours, and* no *means that a gene has a FC in the opposite direction after 24 hours of induction*

Gene	Log_2 FC	FDR	Description
CG8620	2.287	1.27E-05	no functional information
CG42502	0.725	3.60E-03	no functional information
Hsp27	1.178	2.73E-02	HSP20-like chaperone
IM3	1.047	3.28E-02	toll signalling pathway, immune response
CR41619	0.958	3.30E-02	no functional information
Fbp2	-1.389	5.65E-02	alcohol dehydrogenase
RpLP2	0.703	5.84E-02	structural constituent of ribosome
Mst57Dc	-1.509	5.91E-02	mating behaviour
S-Lap4	-1.482	5.91E-02	proteolysis
CG5986	1.011	6.44E-02	telomeric stability
CG3224	0.955	7.48E-02	neurogenesis
CR31032	1.438	8.06E-02	no functional information
Trx-2	0.648	8.06E-02	thioredoxin, DNA damage, stress response
Cyp4g1	-1.500	8.23E-02	oxidation-reduction, cytochrome P450
CG7920	-0.673	8.23E-02	lipid metabolism
Hsp67Bc	1.311	8.23E-02	HSP20-like chaperone
Jon65Aiv	-1.461	8.23E-02	proteolysis
mRpS5	0.700	8.23E-02	structural constituent of ribosome
RpS3A	0.561	8.89E-02	structural constituent of ribosome
Tsp39D	0.552	9.69E-02	tetraspanin, cell membrane scaffolding
CG17454	0.574	1.50E-01	RNA binding, tudor domain
CG6770	0.497	1.50E-01	nuclear phosphoprotein, DNA binding
Pkd2	1.376	1.50E-01	cation channel activity
Hsp23	1.013	1.81E-01	HSP20-like chaperone
Obp99d	1.022	1.98E-01	odorant binding, pheromone binding
$snRNA_{U2:34ABa}$	1.306	1.98E-01	mRNA splicing

Table 3.11: *Differentially expressed genes after 48 hours of induction, FC = fold change, positive FC means upregulated in Kir2.1 flies, FDR = false discovery rate adjusted p-value*

GO ID	GO term	All genes	DE genes	p-val
GO:0009408	response to heat	57	3	8.20E-04
GO:0009266	response to temperature stimulus	83	3	2.43E-03
GO:0051704	multi-organism process	305	5	2.69E-03
GO:0050832	defense response to fungus	33	2	5.19E-03
GO:0009620	response to fungus	36	2	6.16E-03
GO:0051707	response to other organism	164	3	1.61E-02
GO:0009607	response to biotic stimulus	166	3	1.66E-02
GO:0006412	translation	321	4	1.95E-02
GO:0006979	response to oxidative stress	66	2	1.98E-02
GO:0006950	response to stress	695	6	2.18E-02
GO:0006952	defense response	191	3	2.40E-02

Table 3.12: *GO BP terms enriched in the 26 differentially expressed genes detected after 48 hours of induction, p-values are from a Fisher's exact test and have not been adjusted for multiple testing*

Summary

For the sake of comparison, all three conditions (12, 24, and 48 hour) were screened for overlapping differentially expressed genes that had an *unadjusted* p-value < 0.10. The total overlap was 5 genes which are displayed in Table 3.13. This analysis was carried out under the assumption that false positives would likely be included at random in each dataset but true positives are more likely to recur at higher levels of unadjusted significance. It is NOT meant to declare these as significantly differentially expressed, but to illustrate that the overlap even at relatively lax significance levels is very small. This is probably due to the fact that different processes become active at different times. It also could mean that the differential expression calling process is error prone due to the underlying distributional assumptions. Independent validation is therefore necessary before declaring differential expression with absolute certainty.

Gene	Log_2 FC 12	Log_2 FC 24	Log_2 FC 48	Description
CG42502	0.529	0.411	0.725	no functional information
CG7920	-0.255	-0.462	-0.673	lipid metabolism
CG11984	-0.213	-0.471	-0.482	potassium channel regulation
Trc8	-0.549	-0.412	-0.390	negative regulation of growth
Tango5	-0.201	-0.449	-0.373	Golgi organisation
Hcf	0.522	0.352	0.384	chromatin remodeling

Table 3.13: *Overlapping differentially expressed genes after 12, 24, and 48 hours of induction with an unadjusted p-value < 0.1, FC = fold change, positive FC means upregulated in Kir2.1 flies*

Taken together, the data suggest that after 12 hours of induction there is a neurogenic/gliogenic/synaptogenic signal measured in the third antennal segment after Kir2.1-based silencing of $GH146^+$ cell population. No significant changes were detected after 24 hours of induction perhaps due to problems with the Kir2.1-nc libraries. After 48 hours, the transcriptomic environment has turned from proliferative to stressed.

The remaining questions that need to be addressed are exactly which cells in the third antennal segment are regulating transcription of these genes (cellular resolution), and which cells are providing the signal to motivate the change (PNs or glia).

It is unlikely that the measured transcriptional regulation is occurring in the $GH146^+$ ensheathing glial cells. Very few of them are located in the third antennal segment, and that means the power to detect regulation in just those cells would be very low. To verify this, all sequencing libraries were screened for reads mapping to the Kir2.1, GAL4, and $GAL80^{ts}$ transgenes. $GAL80^{ts}$ expression should be present in all cells, whereas Kir2.1 and GAL4 should only be expressed in the $GH146^+$ ensheathing glia that enter the third antennal segment (Sen *et al.*, 2005). The raw read counts for each transgene can be found in Table 3.14. As expected, very few reads mapped to GAL4 and Kir2.1,

and significantly more mapped to $GAL80^{ts}$. There is some variability in the read counts mapping to $GAL80^{ts}$ sequence because they have not been normalised to read depth.

library	Kir2.1	GAL4	$GAL80^{ts}$
KIR-12-1	1	4	87
KIR-12-2	0	10	51
KIR-12-3	0	0	81
KIRnc-12-1	3	0	41
KIRnc-12-2	0	0	49
KIRnc-12-3	1	0	18
KIR-24-1	0	0	120
KIR-24-2	2	0	28
KIR-24-3	0	2	30
KIRnc-24-2	6	0	31
KIRnc-24-3	0	0	11
KIR-48-1	0	1	124
KIR-48-2	3	4	54
KIR-48-3	0	0	87
KIRnc-48-1	12	0	52
KIRnc-48-3	0	0	62
AV	1.75	1.31	57.88

Table 3.14: *Raw read counts mapping to transgene sequences after 12, 24, and 48 hours of induction*

3.4.2 RNA-seq Results

The small number of genes detected to be differentially expressed in the 12, 24 and 48 hour induction experiments suggested that either more replicates or deeper sequencing was required. It could be that almost perfectly isogenic, age-matched third antennal segments raised in the same environmental conditions with the only small genetic difference being expressed in a somewhat distant tissue truly do not have many expression differences. It also could be that the changes are subtle, and a three-replicate experimental design with the existing sequencing depths and detection thresholds was not sensitive enough to detect a majority of the changes.

To address these concerns, the 96 hour induction experiments were carried out using high-amplification, high-sensitivity sequencing technology. This technology was not used before simply because it became commercially available only recently. More replicates were sequenced to increase power for detecting smaller fold changes. The purity of the samples was assessed with the same metrics as before. Although the input RNA was generated similarly (only difference was precipitation steps), only 50 third antennal segments

were collected per preparation, and the resulting RNA was then oligo-dT separated and amplified with 14 cycles of PCR.

The total number of sequenced libraries was 20, which were done in 2 batches. The first batch was 12 samples consisting of 6 replicates from each of the 2 genotypes (all male third antennal segments). Two replicates of each genotype in the first batch (KIR-96-2, KIR-96-3, KIRnc-96-2, and KIRnc-96-3) were sequenced to twice the depth of the others to be able to detect very lowly expressed genes more reliably if needed. The second batch (8 samples total) consisted of another 4 samples of each genotype, with 2 samples of each genotype being third antennal segments from females and 2 samples of each genotype being third antennal segments from males. The mapping statistics separated by batch can be found in Table 3.15.

ID	Total Reads	Mapped Reads	Uniquely Mapped	AV R^2
KIR-96-1	31,213,542	25,030,748(80.19%)	14,405,148(46.15%)	**0.9836**
KIR-96-2	55,190,242	43,522,921(78.86%)	26,736,203(48.44%)	0.9937
KIR-96-3	58,556,838	47,205,802(80.62%)	31,655,311(54.06%)	0.9941
KIR-96-4	29,803,652	24,632,464(82.65%)	17,888,586(60.02%)	0.9933
KIR-96-5	28,647,909	23,564,548(82.26%)	17,152,909(59.87%)	0.9937
KIR-96-6	31,030,362	23,420,613(75.48%)	17,208,103(55.46%)	0.9925
KIRnc-96-1	29,748,012	24,036,358(80.80%)	15,070,224(50.66%)	0.9928
KIRnc-96-2	61,127,179	48,487,739(79.32%)	30,257,167(49.50%)	0.9933
KIRnc-96-3	58,864,472	49,116,496(83.44%)	36,050,234(61.24%)	0.9939
KIRnc-96-4	30,582,966	25,278,565(82.66%)	17,865,165(58.42%)	0.9928
KIRnc-96-5	27,223,175	22,593,695(82.99%)	16,178,509(59.43%)	0.9935
KIRnc-96-6	29,951,768	24,706,089(82.49%)	17,520,016(58.49%)	0.9908
KIR-96-7	26,014,249	21,066,849(80.98%)	14,655,033(56.33%)	0.9932
KIR-96-8	30,622,108	24,934,568(81.43%)	17,368,348(56.72%)	0.9928
KIR-96-9	35,563,235	27,845,191(78.30%)	14,971,400(42.10%)	**0.9869**
KIR-96-10	26,388,734	21,106,014(79.98%)	13,725,816(52.01%)	0.9904
KIRnc-96-7	32,514,120	26,933,883(82.84%)	19,342,591(59.49%)	0.9940
KIRnc-96-8	27,902,225	23,365,715(83.74%)	17,388,108(62.32%)	0.9936
KIRnc-96-9	28,058,965	21,639,543(77.12%)	11,934,889(42.54%)	0.9911
KIRnc-96-10	28,737,234	23,114,269(80.43%)	14,720,484(51.22%)	0.9928
AV	33,624,276	27,140,537(80.63%)	18,099,756(54.12%)	0.9921

Table 3.15: *The experimental design for the 96-hour induction experiment, batches are separated in the table, the AV R^2 is calculated based on the rLog-transformed values, AV stands for average, and numbers in bold identify low-correlation replicates which were discarded*

Again, libraries KIR-96-1 and KIR-96-9 had comparatively low average correlation among within-batch biological replicates and were discarded from the downstream analysis. The mapping percentages are slightly misleading. Due to the amplification steps taken in this experiment, there were many primer sequences that needed to be trimmed from the

raw reads. Therefore, the percentages are slightly higher than the table suggests because there were reads thrown out from the raw reads that never entered the mapping process. Even as presented in the table, the mapping quality is in line with published work using the same methods (Hughes *et al.*, 2012)

3.4.2.1 Purity Assessment

The extra amplification made it possible to detect almost all of the *expected* sensory receptor genes (ORs, GRs, and IRs), even the ones that were below the detection thresholds for the techniques used in the pilot experiment. Figure 3.14 shows that all of the expected ORs except for Or23a give a strong signal across all libraries. Another interesting observation is that three ORs annotated to be expressed in the maxillary palps are detected at low levels in all third antennal segment libraries across all preparations (both RNA-seq and 3′ DGE). The other 4 palp ORs are not detected in any of the preparations. It is therefore possible that the ORNs expressing these receptors are in both organs, which has not been reported.

As for the GRs, Figure 3.15 shows that the same receptors that were detected to have strong third antennal segment expression in the 3′ DGE experiments were detected again. Many unstudied GRs were found to be expressed at very low levels which could be a result of the heat treatment. One interesting finding was that Gr32a (in red in the heatmap) is robustly downregulated by Kir2.1 overexpression in the $GH146^{+}$ cell population compared to Kir2.1-nc controls. Gr32a mutant males show increased courtship to males as well as mated females (Miyamoto & Amrein, 2008). Downregulation of this GR could therefore result in the males becoming less sexually selective.

The IR gene set, shown in Figure 3.16, also passes the quality checks. All IRs reported to be expressed in third antennal segments were detected and a few new IRs were also seen but at lower levels. It could be that they are present at very low levels in the tissue under normal conditions and expression is increased after long periods at high temperatures. Higher temperatures do change the vapour pressure of volatile chemicals which could have strong implications for the olfaction machinery as a whole, but this will not be dwelt upon in this thesis.

The final purity check was the JO-enriched gene set. The heatmap from the pilot experiment is included on the left in Figure 3.17 for ease of comparison. Both heatmaps are sorted the same way to highlight their similarities. Again the fact that many of the JO-enriched genes are not detected in the third antennal segment supports the purity of the preparations and confirms that there is minimal second antennal segment contamination.

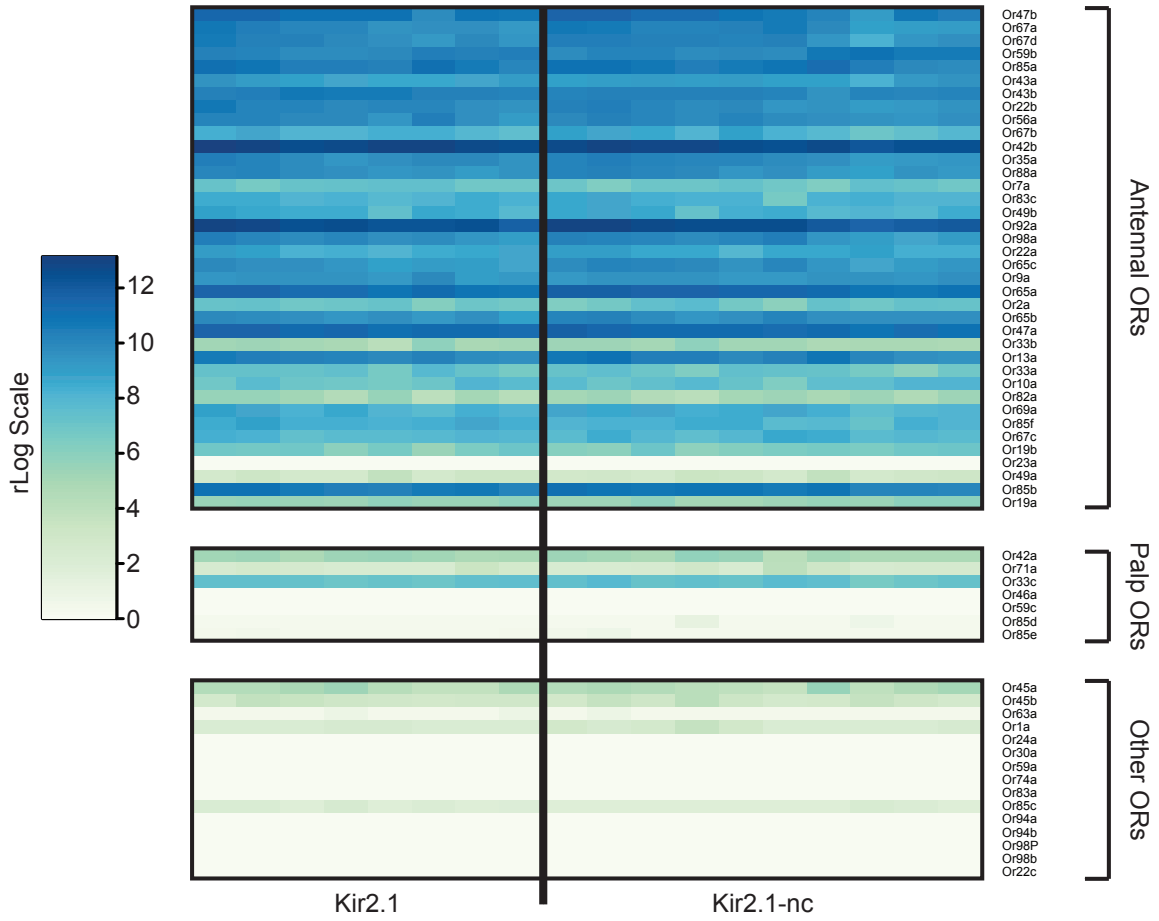


Figure 3.14: *Heatmap of expression rLog values across sequencing libraries for OR genes, annotated at right with expected expression tissue based on literature review. Or83b has not been included in the list because its expression values were so high that it obscured the scale*

As supported by careful analysis of the relevant gene sets, the quality of the RNA preparation and library generation protocols seems up to par with prior experiments.

3.4.2.2 Differential Expression

Due to the high number of biological replicates, it was possible to analyse this data with both parametric and non-parametric differential expression analysis tools. The standard parametric *DESeq2* analysis provided a straightforward way to control for sex and batch, so that analysis will take preference. Taking a highly stringent FDR-adjusted p-value of < 0.05, 32 genes were called differentially expressed and can be found in Table 3.16.

To compare these genes with differentially expressed genes called by another approach, a non-parametric analysis was done with *SAMseq*. *SAMseq* provides a way to process

Gene	Log_2 FC	FDR	Description
PGRP-SD	-1.643	2.70E-11	immune response
Dro	-2.753	1.00E-10	immune response
Hsp26	1.254	1.16E-10	HSP20-like chaperone
pirk	-1.501	1.05E-08	immune response
Obp56a	-1.193	1.48E-06	response to pheromone
CG10332	-2.095	1.92E-06	no functional information
CG42305	1.049	3.85E-06	no functional information
Gr32a	-2.357	7.99E-06	pheromone detection, mating behaviour
Hsp23	1.418	5.93E-05	HSP20-like chaperone
ple	-0.539	1.38E-04	dopamine metabolism
CG42821	-0.958	2.84E-04	no functional information
ST6Gal	-0.681	8.23E-04	sialytransferase activity
IM2	-1.184	1.03E-03	immune response
Hsp67Bc	1.210	1.36E-03	HSP20-like chaperone
CG5346	-0.543	1.89E-03	oxidation/reduction process
Jhe	-1.605	4.35E-03	hormone esterase activity
CG16700	-0.365	5.41E-03	transmembrane amino acid transporter
CG8303	-0.787	7.29E-03	fatty acyl-CoA metabolism
CG8788	0.402	9.99E-03	no functional information
Gmap	-0.468	9.99E-03	Rab GTPase binding, Golgi organisation
ZnT63C	-0.473	1.11E-02	cation efflux
Hnf4	-0.323	1.62E-02	nuclear hormone receptor
CG12290	-0.373	1.90E-02	GPCR
IM1	-0.773	3.07E-02	immune response
CG18302	0.362	3.15E-02	lipid metabolism
CG3301	0.431	3.15E-02	glucose/ribitol metabolism
Mmp1	-0.399	3.15E-02	programmed cell death
AIF	-0.453	3.93E-02	apoptosis inducing factor
CG15890	-0.305	3.93E-02	transmembrane substrate transporter
IM3	-0.756	3.94E-02	immune response
Buffy	0.491	4.64E-02	inhibits programmed cell death
CG34456	0.512	5.00E-02	no functional information

Table 3.16: *Differentially expressed genes after 96 hours of induction, FC = fold change, positive FC means upregulated in Kir2.1 flies, FDR = false discovery rate adjusted p-value*

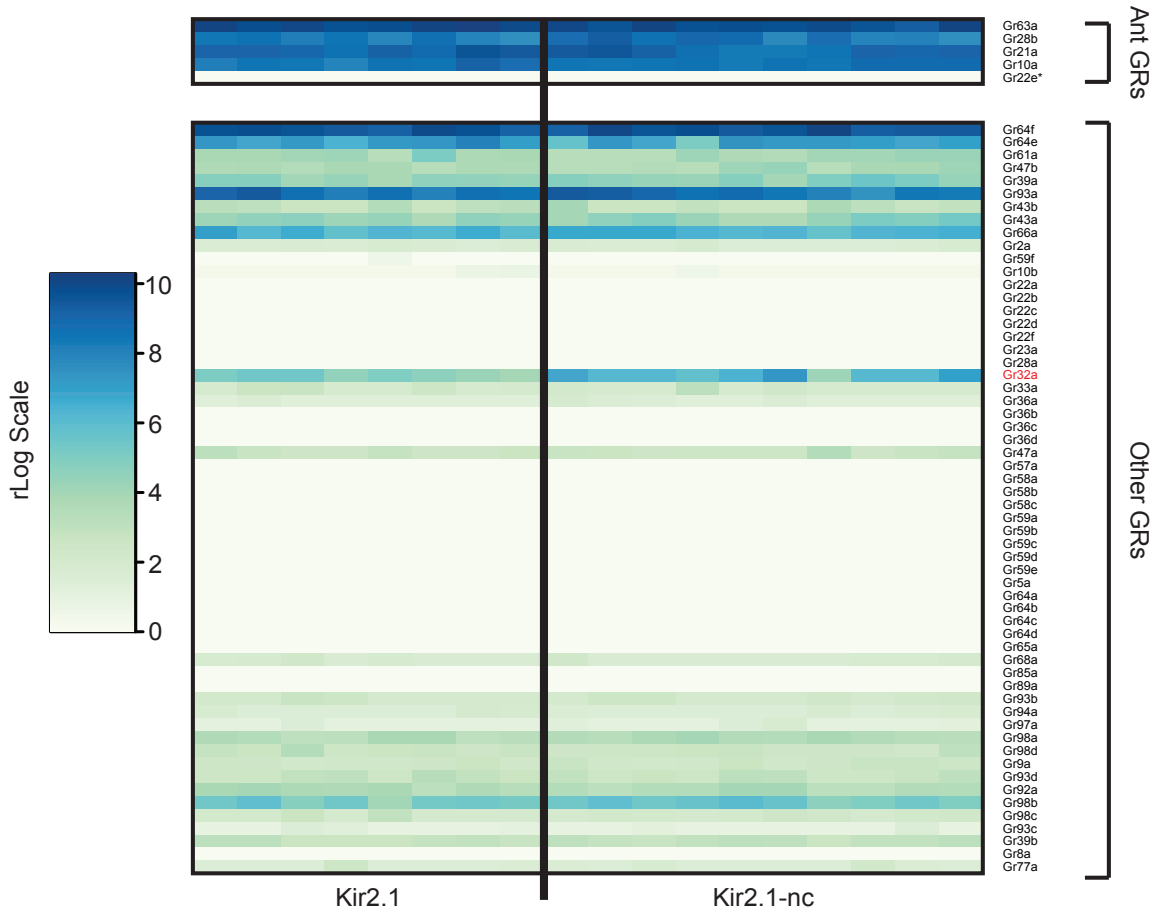


Figure 3.15: *Heatmap of expression rLog Values across sequencing libraries for GR genes, annotated at right with expected expression tissue based on literature review, * - detected with GAL4 driver but not* in situ

different batches, but for the experimental design in this case where batch 2 only has 4 biological replicates, it is not ideal. Also, the sex differences need be processed as a separate batch, which limits the permutation design even more. The result of this is an overly conservative FDR-adjusted p-value.

For the same stringent FDR-adjusted p-value < 0.05 threshold, *SAMseq* finds only 10 differentially expressed genes (all of which were found by *DESeq2*). For an FDR-adjusted p-value < 0.30, *SAMseq* finds 47 differentially expressed genes, 37 of which have FDR-adjusted p-values < 0.20 as reported by *DESeq2*. The two methods do settle on the same genes, and determining the overlap depends heavily on the FDR cutoffs chosen for each method. Normally, non-parametric analyses are better for experiments with many replicates, but in this case, the batch and sex structure of the replicates is better suited for the standard parametric tools.

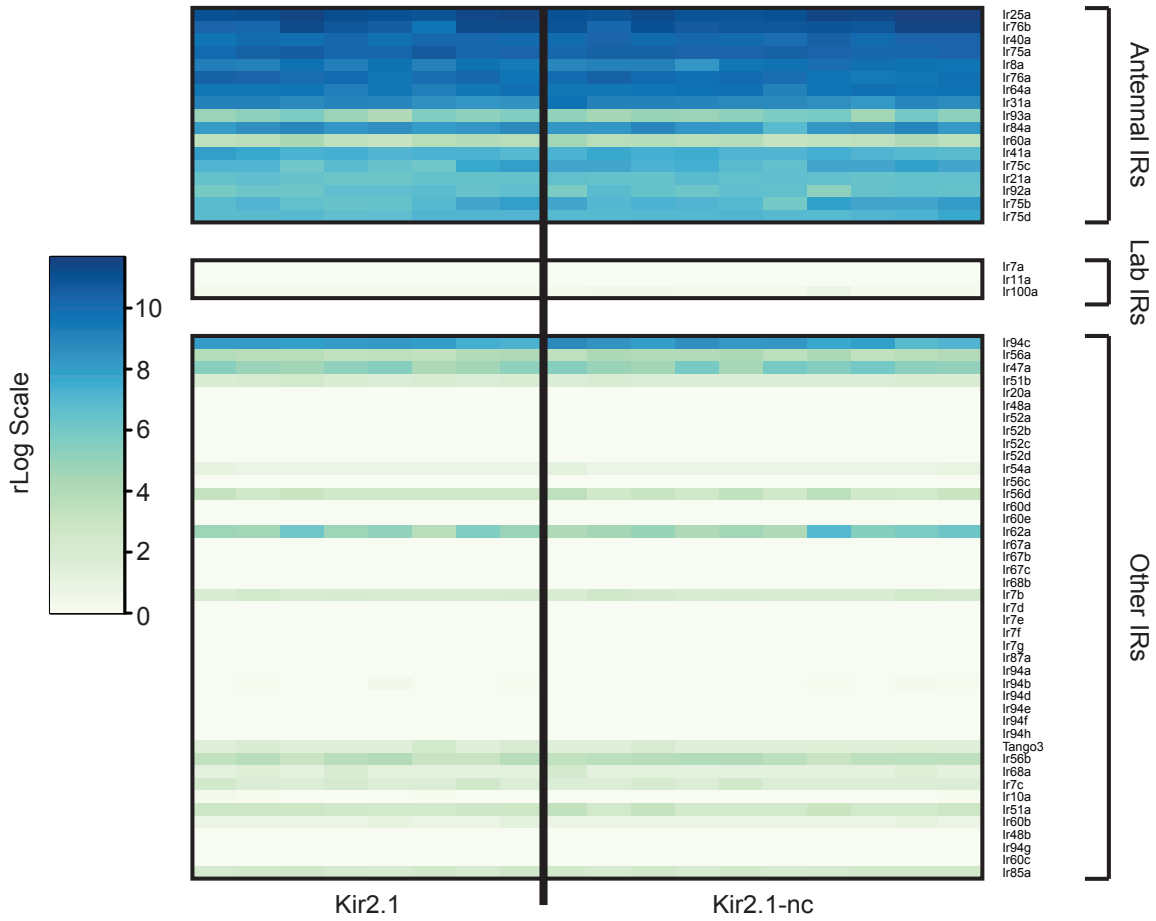


Figure 3.16: *Heatmap of expression rLog Values across sequencing libraries for IR genes, annotated at right with expected expression location based on literature review*

Two important processes are very clear from the differentially expressed genes shown in Table 3.16. All 6 of the significantly differentially expressed immune-related genes are strongly down-regulated (PGRP-SD, Dro, pirk, IM1, IM2, and IM3). At the same time, 3 chaperones known to be involved in the heat-shock response and neuroprotection are up-regulated (Hsp26, Hsp23, and Hsp67Bc). This is in agreement with the 48 hour induction data. A final interesting observation is that 3 genes involved in programmed cell death are differentially expressed. It seems that genes encouraging apoptosis are down-regulated (Mmp1 and AIF), while genes inhibiting apoptotic cell death are up-regulated (Buffy). These gene expression changes taken together imply a neuroregenerative environment where neuronal turnover is being slowed and stress responses are highly active.

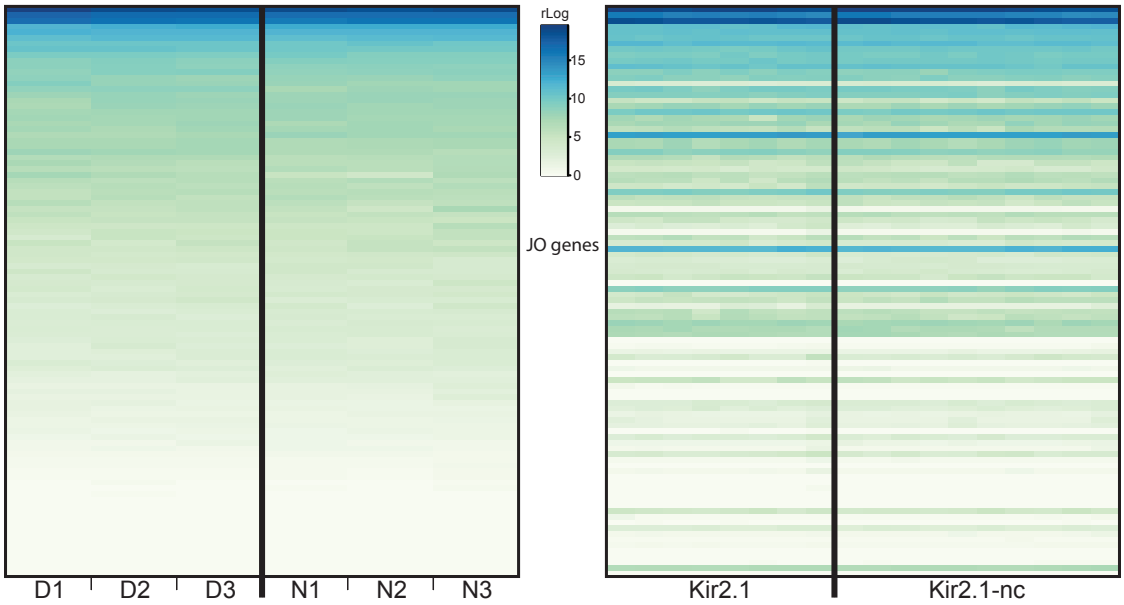


Figure 3.17: *Two heatmaps showing expression of the 100 JO enriched genes sorted by expression value in third antennal segments on the left from the pilot experiment with CS flies and on the right from the RNA-seq experiment with 96 hour-induction of Kir2.1 expression.*

3.4.3 RT-qPCR and Sequencing Platform Validation

It is important to validate the gene expression changes detected in transcriptome-wide screens with independent measurement techniques. Several genes were chosen for RT-qPCR validation based on a differential expression analysis of only batch 1 of the 96-hour induction experiments. The cDNA libraries were generated independently, but from the same starting RNA material and using the same library preparation protocols as the RNA-seq. Some of the genes validated do not appear in Table 3.16, only because when batch 2 is included in the analysis the genes no longer fall in the most significant group. They are still differentially expressed (except for Ipk2 which was not confirmed by RT-qPCR), but are missed due to the highly stringent significance thresholds taken for inclusion in Table 3.16.

Figure 3.18a shows the fold change for the 11 chosen genes based on RT-qPCR with the housekeeping gene CycK. CycK was chosen as a housekeeping gene because in the sequencing data it was expressed at similar levels to the differentially expressed genes, and it was consistently expressed at the same level across all groups. Eight of the genes chosen were upregulated in eGFP::Kir2.1 flies based on the sequencing, and 3 of the genes were downregulated. Figure 3.18b plots the fold changes calculated by RT-qPCR and RNA-seq and shows that there is good agreement between the two independent measures of expression level. A good agreement for several genes not only validates the regulation

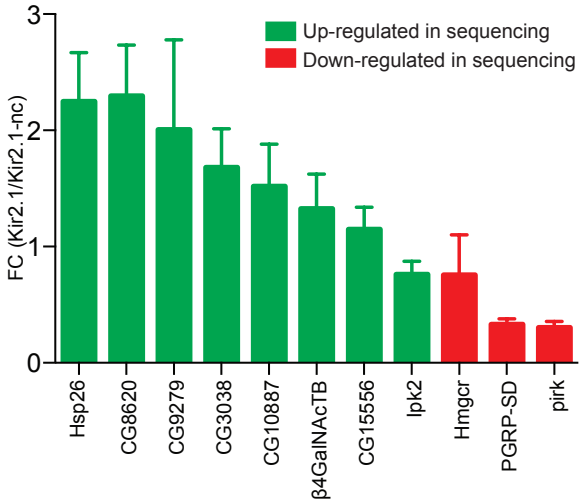
measured for these genes, but also the measurements across the entire dataset.

As a final comparative measure, the average expression level of all expressed genes in all libraries can be plotted to check the agreement between the 3′ DGE measurements (used in 12, 24, and 48-hour induction experiments) and the RNA-seq measurements (96-hour induction only). While the relationship looks appropriately linear, the effect of the extra amplification done in the RNA-seq is apparent in Figure 3.18c. Lowly expressed genes have higher measured expression levels as reported by RNA-seq, however genes with high expression do not follow the same trend. That means lowly expressed genes are amplified more than the highly expressed genes. This is important to keep in mind when assessing the fold changes reported in each case, but there is never a direct comparison for differential expression between libraries generated by 3′ DGE and RNA-seq, so it is not crucial. The general linear trend does confirm that both techniques measure gene expression accurately.

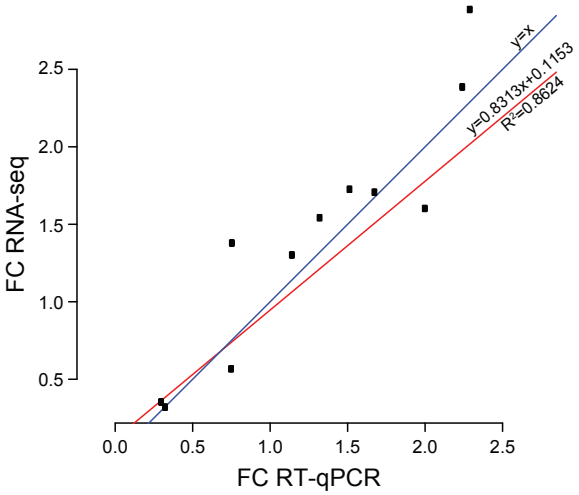
Lastly, the RNA-seq data was also mapped to the sequences of the transgenes and as can be seen in Table 3.17, there are many more reads mapping to the Kir2.1 sequence in this dataset. This means that extra care needs to be taken with these genes because they could be genes expressed in the $GH146^{+}$ glia. Nevertheless, the read counts are still very low in many of the libraries, and several of the differentially expressed genes can also be cross-referenced in the 48 hour induction experiment.

library	Kir2.1	GAL4	$GAL80^{ts}$
KIR-96-2	54	0	2941
KIR-96-3	40	0	3260
KIR-96-4	72	0	1787
KIR-96-5	24	0	1257
KIR-96-6	0	0	1037
KIRnc-96-1	47	0	1037
KIRnc-96-2	1	0	2116
KIRnc-96-3	64	0	2638
KIRnc-96-4	22	1	1412
KIRnc-96-5	17	0	1385
KIRnc-96-6	75	0	1249
KIR-96-7	23	0	1333
KIR-96-8	46	0	1138
KIR-96-10	19	0	1391
KIRnc-96-8	126	4	1502
KIRnc-96-9	121	0	1644
KIRnc-96-9	43	0	1049
KIRnc-96-10	28	0	1355
AV	45.67	0.28	1640.61

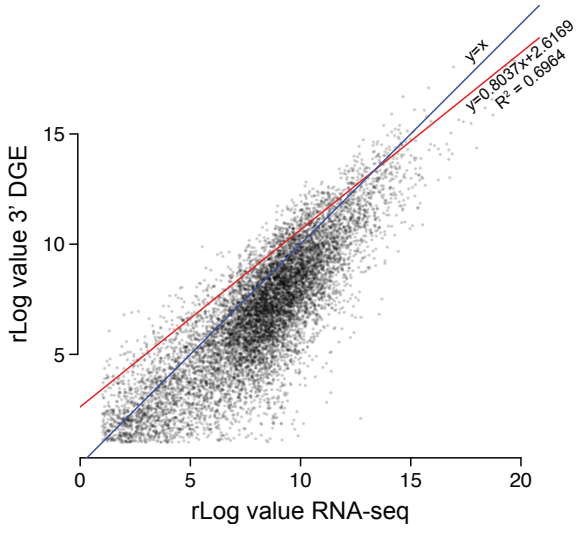
Table 3.17: *Raw read counts mapping to transgene sequences after 96 hours of induction*



(a) *RT-qPCR validation of up and down-regulated genes in the 96-hour induction experiments, CycK is the housekeeping gene*



(b) *Comparison of fold changes calculated by RT-qPCR and RNA-seq, FC = fold change, CycK is the housekeeping gene*



(c) *Comparison of transcriptome-wide average rLog expression values across all expressed genes for 3′ DGE and RNA-seq in third antennal segments*

Chapter 4

Discussion

This work has made significant strides in the design and use of a system to study transsynaptic transcriptional regulation. Understanding how neurons can regulate gene transcription in their presynaptic partners as a result of changes in intrinsic excitability can inform innovative approaches to neural circuit tracing. As mentioned in the Introduction, current tracing methods rely on either the injection of foreign substances or the endogenous overexpression of synapse-crossing proteins, both of which have well-documented shortcomings. The development of novel tracing systems based on the transcriptome analysis carried out in this thesis has been carefully thought out and will be detailed in this section along with a summary of the transcriptome data.

4.1 Homeostatic Gene Expression Regulation

4.1.1 Early Proliferation

After 12 hours of adult-onset Kir2.1 expression in the $GH146^{+}$ cell population, there is a strong suggestion of proliferation and differentiation of neural cells in the third antennal segment. Six of the top 20 differentially expressed genes after 12 hours of induction are involved in either growth, neuronal development, or neuronal differentiation.

The most immediate question is whether this is a result of postsynaptic silencing of PNs or silencing of the ensheathing glia of the antennal nerve. It is not certain that Kir2.1 expression in ensheathing glia will have any effect on the physiology of the glial cells. First of all, they do not spike. Also, many reports have found that ensheathing glia have membrane potentials at or around the reversal potential for potassium (Butt & Kalsi, 2006; Bay & Butt, 2012). One of the purported functional roles of ensheathing glia is to potassium buffer the enclosed axons and assure efficient redistribution of the potassium ions (Kofuji *et al.*, 2002). If the main function of the ensheathing glia is to redistribute

potassium and in normal conditions their membrane potential is already at the potassium reversal potential, then small amounts of Kir2.1 overexpression should not disrupt the normal function of these cells. Nevertheless, it will be important to finely pinpoint the source of the neurogenic signal by crossing in a Repo-GAL80 transgene.

Another interesting question is whether the measured changes in gene expression are explained by the birth of new cells, or that existing cells activate neural circuit remodelling modules that are used in development and normally quiescent in stable circuits. The *S/G2/M-Green* transgene available for DM allows endogenous expression of a cell cycle phase-monitoring protein that can identify dividing cells (Nakajima *et al.*, 2011). S/G2/M-Green is a fluorescent-tagged geminin fragment that migrates to the nucleus only during the S, G2, and M phases of the cell cycle. Assuming that dividing cells are normally minimal in the adult third antennal segment (proliferation has not been reported there in adults), it should be possible to cross in an actin promoter-driven S/G2/M-Green transgene (DGRC Kyoto stock # 109678) and look for fluorescent nuclei after adult-onset expression of Kir2.1 in $GH146^{+}$ cells. This can answer the question of whether new neurons are being born or existing neurons are remodelling themselves.

If adult-born ORNs are being created in the third antennal segment as a result of Kir2.1 overexpression in $GH146^{+}$ cells, it would be an important discovery. Several reports have documented proliferating glial cells around the antennal nerve (Kato *et al.*, 2009; von Trotha *et al.*, 2009) in the DM central brain, but these studies found no adult neurogenesis in the central brain. Another report found adult neurogenesis in the DM optic lobe (Fernandez-Hernandez *et al.*, 2013), but this report is controversial because of the use of a perma-twin flippase system instead of the traditional edU and BrdU methods for detecting dividing cells. Also, age-matched controls were not carried out. Therefore, independent validation is needed to confirm this report.

If the third antennal segment is a neurogenic niche, and the birth of new neurons can be stimulated by altering activity in downstream circuit elements (either the ensheathing glia of the antennal nerve or the postsynaptic PNs), this would provide an ideal genetic system to study activity-driven adult neurogenesis. Circuit activity has been shown in mammals to directly regulate the neurogenesis rates and cell survival probabilities in the neurogenic regions of the mammalian olfactory bulb and dentate gyrus (Deng *et al.*, 2010). The genetic tools available in DM would make it easy to identify the type of new ORNs being born, monitor their morphological development and incorporation into the antennal lobe circuit, and to endogenously express actuators and reporters in single adult-born ORNs and observe their impact on downstream circuit elements (iLNs, eLNs and PNs). All of these present interesting and unique experimental approaches to study the birth, integration, and function of adult-born neurons at an unprecedented resolution.

While it is important to carefully scrutinise the highly significant differentially expressed genes and investigate their functional roles, it is also worth considering genes that were expected to be differentially expressed and were not detected, or more precisely that did not meet the stringent false positive thresholds that need to be enforced in transcriptome-wide expression studies. Two interesting sets of genes to search out are genes known to play a role at synapses and genes known to play a role in homeostatic synaptic compensation.

4.1.1.1 Synaptic Genes

It is a bit surprising that only four of the top differentially expressed genes in the 12 hour induction data are related to synaptic functions, and this paucity of synaptic candidates could be for several reasons. It appears that synaptic proteins are not very highly expressed compared to other genes and are regulated within tight ranges, and thus limited statistical power might preclude their detection. If that is indeed the case, and synaptic genes are regulated but not detected, one would expect to find an overabundance of synaptic genes with low p-values that do not reach significance using stringent thresholds and correction for multiple testing.

Looking closely at the data, out of the 152 expressed genes annotated by the GO term for synapse organisation (GO:0050808), 13 of them have unadjusted p-values < 0.05. By chance alone, only $\approx$7 genes would be expected in this group. Looking beyond annotations based on a single GO term and manually curating the data, there are 33 synapse-related genes (Table 4.1) with unadjusted p-values < 0.05 out of $\approx$400 total genes at this significance level. This is a manageable number of candidate synaptic genes to individually test one by one with the understanding that some of them are likely to be false positives. The suspected overabundance of low p-value synaptic genes is thus confirmed in the data.

Another reason that not many synaptic genes were detected could simply be because homeostatic regulation of synaptic genes is not transcriptional. Everything could occur via local translation at the synapse. Local translation at synapses has been widely studied in developmental axon guidance and in plasticity (Yoon *et al.*, 2009), but not in cases examining homeostatic mechanisms.

To speculate, the modules that homeostatically stabilise synapses could be controlled in similar ways as plasticity modules in that there are fast, local changes occurring through local translation and post-translational modification of proteins and at the same time there are slower, longer-lasting transcription-dependent processes being initiated. In support of this hypothesis, it has been reported that presynaptic, but not postsynaptic, homeostatic changes in response to a global reduction in neural activity are transcription-dependent (Han & Stevens, 2009). It has also been shown that transcription is necessary

Gene	Log_2 FC	p-val	Description
baz	0.302	5.40E-05	protein kinase C binding, multicellular development
sff	0.404	3.63E-04	kinase, ATP binding
pros	0.350	9.59E-04	transcription factor, neuroblast differentiation
Ank2	0.317	1.72E-03	membrane anchoring, ion homeostasis
Nsf2	1.072	4.94E-03	synaptic vesicle priming
rab3-GEF	0.969	5.74E-03	GTPase signalling, vesicle priming, synapse assembly
Atpalpha	0.213	7.96E-03	ATPase, Na^+/K^+ pump alpha subunit
CG34139	0.790	1.02E-02	neuroligin-4, cell adhesion, synapse assembly
Khc-73	0.224	1.12E-02	kinesin motor, anterograde transport
sda	-0.220	1.13E-02	peptidase
n-syb	0.179	1.30E-02	exocytosis, neurotransmitter vesicle secretion
Syn1	0.488	1.59E-02	PDZ domain, regulation of synaptic growth
Med	-0.627	1.66E-02	BMP signalling co-factor
nrv1	0.197	1.70E-02	ATPase, Na^+/K^+ pump beta subunit
CkIIbeta	0.187	2.05E-02	casein kinase-2 subunit, neuronal development
Syx6	-0.482	2.26E-02	vesicle release from the Golgi
Dscam	0.500	2.35E-02	cell adhesion, synapse formation
Glut4EF	0.202	2.40E-02	glucose transporter enhancer factor
scrib	-0.203	2.81E-02	PDZ domain, sensory organ development
Rab5	0.153	3.12E-02	GTPase activity, endosomes
hiw	0.364	3.48E-02	ubiquitin ligase activity, synaptic development
Gprk2	0.325	3.48E-02	G protein coupled receptor kinase-2
Rdl	0.172	3.69E-02	$GABA_A$ receptor
CASK	0.189	3.86E-02	membrane-associated guanylate kinase
G-ialpha65A	0.243	3.87E-02	G protein, neuroblast division
para	0.102	3.97E-02	voltage gated Na^+ channel
stmA	-0.265	4.17E-02	synaptic vesicle exocytosis and endocytosis
unc-13-4A	0.368	4.53E-02	vesicle priming
Mmp1	0.213	4.66E-02	peptidase, extracellular matrix organisation
neuroligin	-1.040	4.67E-02	cell adhesion, synapse formation
cpx	0.137	4.81E-02	neurotransmitter secretion, exocytosis
CrebB-17A	0.216	4.86E-02	transcription factor, cyclic AMP response
sgg	0.146	4.86E-02	glycogen synthase kinase, neuronal development

Table 4.1: *Differentially expressed synaptic genes after 12 hours of induction with unadjusted p-values < 0.05*

Gene	Log_2 FC	p-val	Description
rab3-GEF	0.969	5.74E-03	GTPase signalling, vesicle priming, synapse assembly
Khc-73	0.224	1.12E-02	kinesin motor, anterograde transport
CASK	0.189	3.86E-02	membrane-associated guanylate kinase
unc-13-4A	0.368	4.53E-02	vesicle priming
CrebB-17A	0.216	4.86E-02	transcription factor, cyclic AMP response
unc-13	0.245	6.44E-02	vesicle priming

Table 4.2: *Differentially expressed homeostatic genes after 12 hours of induction with unadjusted p-values < 0.075, FC - fold change*

for synaptic un-silencing after relief from long periods of heightened neural activity (Crawford *et al.*, 2012). Thus, it is likely that at least some of these synaptic genes that were found to be slightly below the significance thresholds for differential expression are in fact regulated.

4.1.1.2 Homeostatic Genes

The second set of genes that are important to flesh out of the data have a confirmed role in presynaptic homeostatic processes. All of them are contained within the synaptic genes listed above but they deserve extra attention. None of these genes have been identified at the ORN/PN synapse (most were identified at the NMJ), but it is expected that at least some of the same homeostatic machinery will be used at different places in the nervous system. Additionally, just because a gene has been implicated in homeostatic processes, it does not mean that it will be *transcriptionally* regulated. It could simply be a sensor, or a part of the signalling cascade passing a homeostatic signal to the nucleus. Table 4.2 shows the 6 most significant differentially expressed homeostatic genes.

Unc-13 and unc13-4a are both DM homologs of a mouse active zone protein (Munc-13) believe to play a role in vesicle priming at the active zone cytomatrix (Rosenmund *et al.*, 2002; Augustin *et al.*, 1999). Munc-13 mutants show a reduction in the number of functional synapses as well as vesicle priming deficits (Rosenmund *et al.*, 2002; Augustin *et al.*, 1999). An upregulation in Munc-13 related DM genes might increase the number of release ready vesicles per active zone. CASK is a membrane-associated guanylate kinase (MAGUK) and is a phosphorylation target of CDK5 (Samuels *et al.*, 2007). Phosphorylation by CDK5 mediates a synaptogenic interaction between CASK and liprin-α (Samuels *et al.*, 2007). Rab3 is a small GTPase present on synaptic vesicles. In the GTP-bound state, vesicular release is suppressed (Muller *et al.*, 2011). Local homeostatic activity induces rab3-GAP-mediated GTP-hydrolysis which increases the amount of GDP-bound Rab3, thus increasing vesicular release (Muller *et al.*, 2011). This could in turn emit a signal to the nucleus to increase levels of rab3-GEF, as seen in the data, to reset the

signalling system and stabilise the release properties of the neuron. Rab3 has also been found to play a role in the distribution of proteins in presynaptic compartments, with Rab3 mutants showing fewer active zones with many more T-bars per active zone (Graf *et al.*, 2009). Khc-73 is a kinesin motor family protein, and increased expression of this gene has been associated with an increased number of T-bars at presynaptic release sites (Tsurudome *et al.*, 2010). Finally, CrebB-17A has been shown to promote synaptic growth at the NMJ and also plays a role in long term memory formation in adult flies (Chen & Ganetzky, 2012). It has not been directly implicated in any homeostatic processes but was included here due to its known role in experience-dependent synaptic remodelling.

While the lower significance levels and small fold changes of these genes mean that they need to be independently verified, it brings the discussion to an important point. Neurons seem to be regulating many interacting, interdependent homeostatic processes at once. The transcription/translation of more synaptic release proteins or the alteration of a single homeostatic pathway is unlikely to be responsible in and of itself for homeostatic synaptic release behaviour. For example, changing calcium channel dynamics partially depends on the rearrangement of the vesicle pool for the effects to truly be felt at the synapse. While the loss of an entire homeostatic module might be functionally visible when looking at single gene mutants, the coincident regulation of many pathways in homeostasis-competent animals could dilute the magnitude of expression changes of any single gene.

Further, neurons have been shown to adjust the physical number of release sites, the demographics of synaptic vesicle pools, calcium dynamics, the number of muted synapses, the expression of transporters that clear excess neurotransmitter from the synaptic cleft, and the vesicular filling process all in similar experimental paradigms (Lazarevic *et al.*, 2013). Looking closer at just one of these functional modifications, the physical number of release sites could be regulated by both providing new synaptic machinery or slowing degradation of existing machinery through alteration of the ubiquitin-proteasome system, or both. All of these processes occurring simultaneously means that gene expression changes are likely to be smaller than if there were a single dominant homeostatic pathway. The combined effect of this interdependency and redundancy would be to make the expression level fold changes at the single gene level smaller and therefore more difficult to detect with the current levels of statistical power.

4.1.2 Late Stress

As the Kir2.1 expression induction period is lengthened (to 48 or 96 hours), it seems that the neurogenic/synaptogenic signals after 12 hours of induction disappear and are

replaced by immune system changes and strong increases in the expression of stress genes. This could be because increased metabolic demands on a neuron can be met for a short period of time, but over longer periods of time it takes its toll on the health of the cell.

By far the largest fold changes across all experiments are seen for the four chaperone genes Hsp23, Hsp26, Hsp27, and Hsp67Bc, which are more than doubled. Both the 48 hour induction data and 96 hour induction data agree on the direction and magnitude of the upregulation for these 4 related genes. They are small heat shock proteins known to play a role in neuroprotection and nerve regeneration, and are the DM homologs of the alpha-crystallin complex (Tower, 2011). The fact that expression of these Hsps is known to be induced by heat means that the method used for inducing gene expression could play a role in the regulation of these genes, but Kir2.1 and control flies were both exposed to the same temperature regimen. At the very least it is a synergistic effect of added stress and temperature changes that upregulates their expression. Where they are expressed, what proteins they fold, and whether they all work in concert are all interesting questions that can be answered with the BAC cloning that will be described in the next section.

The differentially expressed immune genes are all strongly downregulated after 96 hours of induction. This effect could reflect a reallocation of cellular energy usage after several days of increased metabolic demands at temperatures that are higher than the normal range for a fly. Again the Kir2.1 is solely responsible for the effect because control flies were exposed to the same temperature, so it is most likely related to accelerated neurodegeneration as a result of the added physiological demands placed by downstream Kir2.1 expression in the circuit.

4.1.2.1 Synaptic and Homeostatic Genes

As for the synaptic genes highlighted in the last section for the 12 hour data, the number of genes found in the 48 and 96 hour datasets at unadjusted p-value thresholds < 0.05 were not more than would be expected by chance, so it is likely that at this stage transcriptional regulation of synaptic genes is minimal. Two homeostatic genes, however, were detected as differentially expressed with unadjusted p-values < 0.05 in the 96 hour induction data and act in the same pathways as some of the genes found in the 12 hour data. CDK5, which as mentioned before phosphorylates CASK facilitating interaction with liprin-α (Samuels *et al.*, 2007), is upregulated by about 20% in Kir2.1 flies after 96 hours of induction, and CCKLR-17D1, which interacts with CrebB-17A, is also upregulated by about 20%. Again these fold changes are very small and need to be validated with RT-qPCR. However, regulation at such low fold changes is not very useful for the next phase

of the project, which is turning homeostatic changes in gene expression into a reporter for synaptic connections.

4.2 Homeostasis-based Circuit Tracing Systems

To achieve the ultimate goal of this work, the expression levels of homeostatically-regulated genes will need to be monitored with an optical reporter. Obviously, larger fold changes of expression make this task easier, so the genes that will be tested first are the ones with the largest changes. The ideal gene would be expressed broadly at low levels and be both strongly and exclusively regulated in presynaptic cells in response to a postsynaptic change in excitability. The Hsp's are the most ideal in the current set of differentially expressed genes.

The first step in this phase of the project will be to make sure that the expression changes are exclusive to the presynaptic cells. The DM olfactory system is an ideal system to address this because it has well-characterised populations of pre- and postsynaptic cells with tools granting independent access to both classes of neurons. These experiments will thus narrow the expression of postsynaptic Kir2.1 into a single glomerulus or a few glomeruli and optically monitor expression of the homeostatically-regulated genes.

NP3062-GAL4 (Figure 4.1) and R81E09-GAL4 are two available driver lines that narrowly express in either one or a few PN classes. NP3062-GAL4 expresses in PNs postsynaptic to 3 AL glomeruli (DM4, DM6, and DL5), and for two of these glomeruli Or-LexA lines have been constructed (Or7a-LexA drives in ORNs presynaptic to DL5 and Or67a drives in ORNs presynaptic to DM6). R81E09-GAL4 drives in PNs postsynaptic to glomerulus VA6, and an Or82a-LexA line has been constructed that drives in the presynaptic ORNs. Using these tools, it will be possible to independently drive effectors postsynaptically, visualise the identity of the presynaptic cells, and monitor their homeostatic behaviour using the genetic systems detailed below.

All of the genetic systems will be based on bacterial artificial chromosome (BAC) recombineering techniques to capture a gene in context with all of its cis-regulatory elements (Stowers, 2011). The reporter will be a destabilised GFP (dGFP) that degrades faster than standard eGFP and thus will prevent aggregation of the protein and potential loss of resolution to detect homeostatically active cells (Lieber *et al.*, 2011).

4.2.1 Genetics

The first generation of reporters will be site-specific recombinations of either QF2 for upregulated genes or QS for downregulated genes into the exact genomic location of

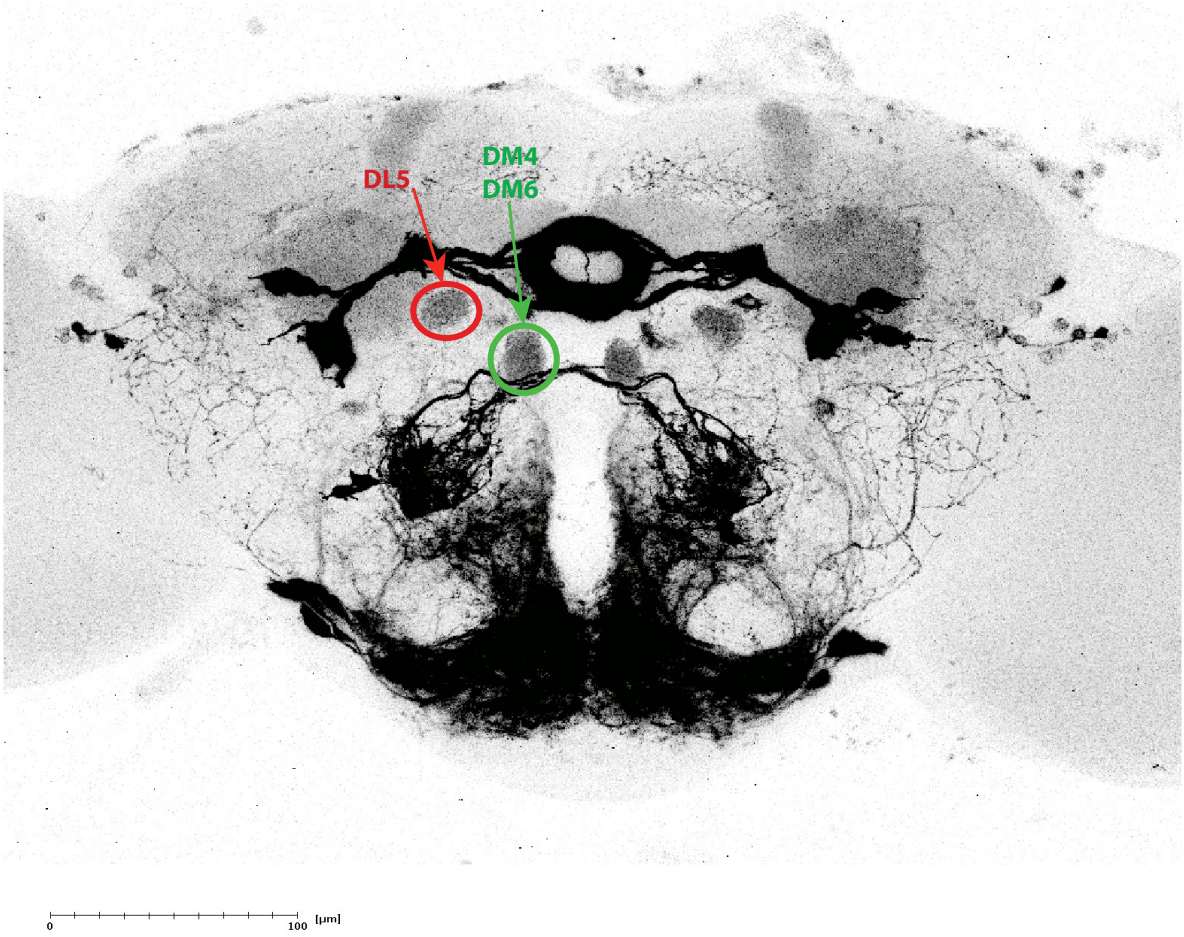


Figure 4.1: *NP3062-GAL4 driven UAS-mCD8::eGFP*

the homeostatic gene within a BAC containing the gene and its cis-regulatory elements (Figure 4.2a). These BACs will then be integrated using the ΦC31 system at the sites in the DM genome receptive for large constructs (Venken *et al.*, 2009).

These preliminary transgenic lines will be used to address several questions. Firstly, it will be important to determine the basal expression levels of the genes of interest and whether they are expressed only in certain areas of the brain. If they are expressed at varying levels in different locations, it will be necessary to normalise reporter signals to regions where homeostatic gene regulation is not occurring, thus looking at relative expression levels instead of absolute.

In the ideal case where reporter expression is ubiquitous and low, it will be important to have ways to adjust the gain of the system in case the magnitude of homeostatic regulation is too subtle to be detected optically (low signal-to-noise ratio). Two potential solutions have been drawn up to increase gain and signal-to-noise ratio (SNR) if this is

true. The *cascade* approach shown in Figure 4.2b relies on the non-linear amplification associated with transcription factor enhancement of expression. In this case, rather than directly driving dGFP expression with QF2, instead QF2 will drive LexA expression which in turn will drive dGFP expression. This is preferable to a QUAS-QF2 positive feedback enhancement because it is not based on an unstable feedback loop, it is simply an extra level of amplification.

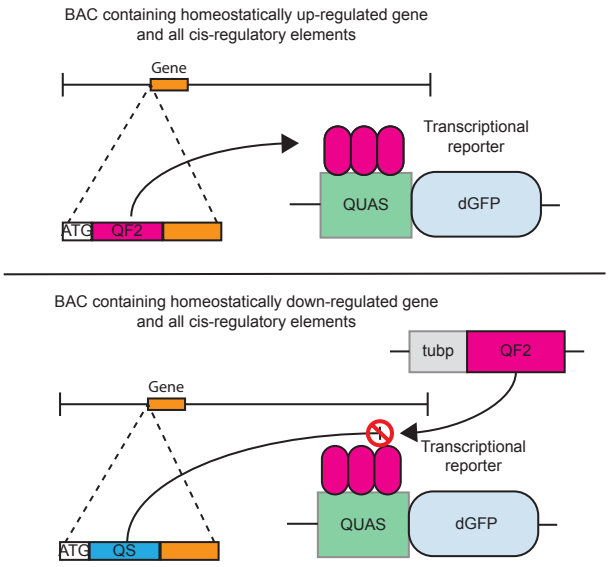
Another approach, shown in Figure 4.2c, employs a genetic *differential amplifier* (Hens, 2014). In this approach, a constitutively active promoter including a LexA repressor binding site (Butala *et al.*, 2009) drives steady QS expression. Upon the induction of the homeostatic gene, bicistronic expression of both QF2 and the repressive LexA (Butala *et al.*, 2009) is activated, and this nonlinearly effects the QF2:QS ratio within the cell and thus the dGFP reporter signal.

While these lines are being constructed, experiments have begun to assess exactly what level of differential expression will be necessary in order to optically detect a difference. Dosage can be varied by crossing in one or two copies of a driver (for example two copies of Or7a-GAL4 driving UAS-dGFP vs. one copy of the driver) to see if it is possible to detect the difference between the two. Obviously, this cannot be done in the same brain so the true testing will begin when the homeostatic gene lines are ready. It will be very important to fine tune basal levels of expression, gain, and SNR.

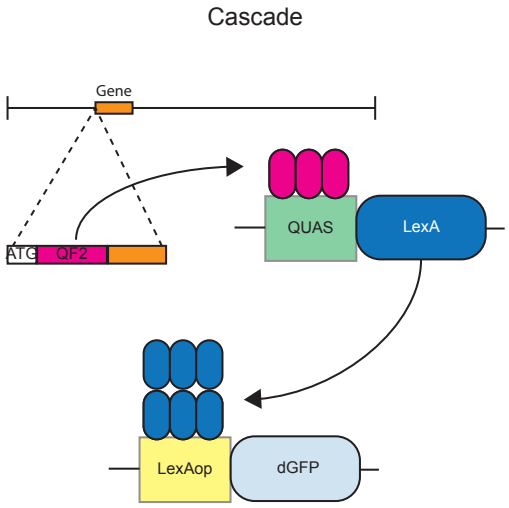
4.2.2 Incorporating Spatial Information - Intersectional Logic

With the ultimate goal being to drive expression of a reporter, sensor, or actuator in neurons that are related by circuit architecture alone, it might be that the best systems will integrate the different circuit tracing technologies and take the form of intersectional approaches based on logical AND gates (Miesenboeck & Kevrekidis, 2005).

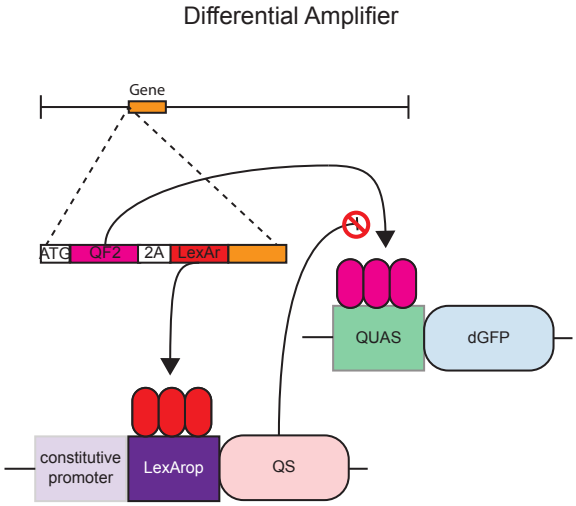
Perhaps this is most easily demonstrated by diagramming a potential genetic system. In the fly, the WGA transgene mentioned in the Introduction travels bidirectionally across synapses, but has not had much use since its discovery (Tabuchi *et al.*, 2000). This is probably because when allowed to travel unchecked, WGA winds up labelling too broadly to be useful. It has been shown that WGA fused to other proteins does not hinder its ability to cross synapses, and in fact WGA carries its fused partner with it (Huh *et al.*, 2010). Therefore, it is possible that the activation domain of the GAL4 transgene (GAL4ad), as used in the split-GAL4 system (Luan *et al.*, 2006), could be fused to the WGA transgene via a tobacco etch virus protease cleavage site (TEVpcs) between the two peptides similar to the way it was used in the TANGO assay (Inagaki *et al.*, 2012).



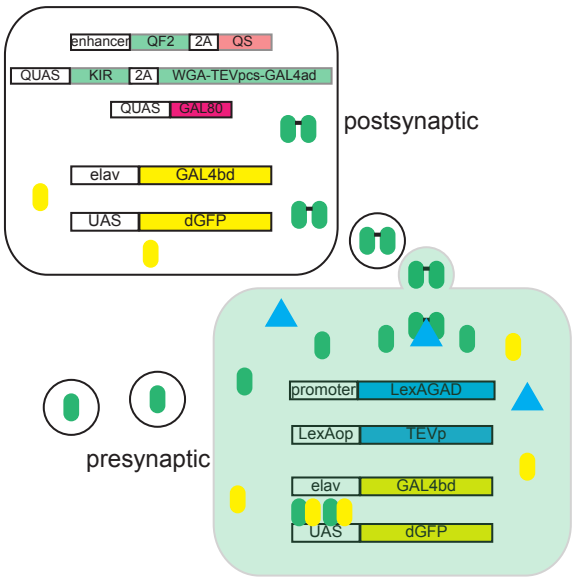
(a) *Scheme for reporting homeostatic gene expression, the QF2 or QS coding sequence recombined into a BAC containing the gene and all of its cis-regulatory elements*



(b) *In case there is need of extra amplification, a cascade approach can be taken that uses the LexA/LexAop system in series with the QF2/QUAS system*



(c) *Another amplification approach using a constitutively active repressible promoter to drive QS and bicistronic expression of QF2 and a repressive LexA*



(d) *An inducible tracing system integrating spatial and homeostatic signals using a WGA-TEVpcs-GAL4ad fusion protein, an elav-driven GAL4bd, and a homeostatic gene-driven TEV protease*

At the same time, homeostatic-based or activity-based gene expression regulation (for example one of the genes found in this thesis) could be used to drive expression of the TEV protease (TEVp), which would separate the GAL4ad from the WGA-TEVpcs-GAL4ad fusion protein (Huh *et al.*, 2010). Importantly, expression of the protease would have to be blocked in the cells expressing the WGA-TEVpcs-GAL4ad fusion protein. It might also be necessary to express the TEVp as a fusion with an endosomal protein (or maybe just with a membrane tag) to bring it in close proximity with the protease cleavage site. Continued transport of cleaved WGA would then not matter because it will have been separated from the GAL4ad after travelling across a single synapse.

This kind of system would likely increase the amount of separated WGA and GAL4ad in cells that are both spatially and homeostatically linked. The expression of WGA-TEVpcs-GAL4ad can be driven by the same enhancer that drives the changes in excitability or activity patterns of a single neuron in the circuit.

The final piece of the system would be pan-neuronal expression of the GAL4 binding domain (GAL4bd), and this could be made inducible by using the Gene Switch versions of GAL4 to give the experimenter the ability to boot the system (Osterwalder *et al.*, 2001). All of the genetic tools mentioned here exist in DM and an overview of the system can be viewed in Figure 4.2d.

4.3 Summary of Future Work

Many important issues have been highlighted in this discussion that need to be summarised. First, now that several candidate homeostatically-regulated genes have been identified, the most crucial questions that need to be answered are the precise cellular source of the homeostatic signal and the identity of the cells in which the regulation is occurring. A GH146-GAL4, Repo-GAL80 driver is under construction to increase the specificity of the postsynaptic perturbation and determine the relative contributions of the postsynaptic neurons and the glia ensheathing the antennal nerve. Or-LexA lines (Or7a, Or67a, and Or82a) have already been generated granting access to specific ORN classes for which there are available GAL4 drivers to silence their cognate PNs exclusively. These tools should allow fine mapping of the transcriptional responses and validation that the responses are indeed synaptic-partner specific.

At the same time, new methods to histochemically separate RNA from specific cells in a heterogeneous sample have become available and could be tested to improve the cellular resolution of the transcriptome measurements (Henry *et al.*, 2012; Thomas *et al.*, 2012). Experiments are also underway to determine whether new neurons are being born in the adult third antennal segment or if existing neurons are activating remodelling modules

that have been characterised only in developing circuits. These experiments depend on an actin promoter-driven S/G2/M-Green transgene that is currently being examined on its ability to detect dividing cells.

Two different assays for measuring glomerulus-specific homeostatic behaviour functionally are also being developed, the first being the ability to count presynaptic sites that are exclusively between a single class of ORN and its cognate PNs, and the second being glomerulus-specific imaging of synaptic release and calcium influx. Both of these assays can be performed when expressing either Kir2.1 or Kir2.1-nc postsynaptically.

Finally, the first generation of BAC cloning is underway to create QF2 lines for Hsp26, rab3-GEF, Nsf2, CG8620, and CG9279. These are the genes with the largest fold changes seen in the data and are thus the starting point for the next phase of the project.

4.4 Conclusion

This work has developed a genetic system and tissue preparation that is able to detect differentially expressed genes in presynaptic circuit elements after adult-onset silencing of the downstream postsynaptic circuit components. No other report has published a transcriptome-wide account of such expression changes in adult flies. While the project has come a long way, there is still a long way to go in taking the homeostatically-regulated genes, validating them, and turning them into circuit tracing tools. Along the way, novel circadian-regulated genes were detected in the third antennal segment, the full set of sensory receptors expressed in the third antennal segment was laid out, a few novel sexually dimorphic OR genes were detected, and many candidate homeostatic gene expression changes were observed. This work raises interesting questions and also overcame many technical difficulties to create an ideal system for transcriptomics-based studies and adult-onset neuronal silencing.

Bibliography

Adams, Mark D., & Celniker, Susan E. et al. 2000. The Genome Sequence of Drosophila melanogaster. *Science*, **287**(5461), 2185–2195.

Alexa, Adrian, RahnenfÃijhrer, JÃűrg, & Lengauer, Thomas. 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, **22**(13), 1600–1607.

Anders, Simon, & Huber, Wolfgang. 2010. Differential expression analysis for sequence count data. *Genome Biology*, **11**(10), R106.

Andrews, S. 2012. FastQC A Quality Control tool for High Throughput Sequence Data. *http://www.bioinformatics.babraham.ac.uk/projects/fastqc/*.

Arenkiel, Benjamin R., & Ehlers, Michael D. 2009. Molecular genetics and imaging technologies for circuit-based neuroanatomy. *Nature*, **461**(7266), 900–907.

Asmann, Yan, Klee, Eric, Thompson, E Aubrey, Perez, Edith, Middha, Sumit, Oberg, Ann, Therneau, Terry, Smith, David, Poland, Gregory, Wieben, Eric, & Kocher, Jean-Pierre. 2009. 3' tag digital gene expression profiling of human brain and universal reference RNA using Illumina Genome Analyzer. *BMC Genomics*, **10**(1), 531.

Augustin, Iris, Rosenmund, Christian, Sudhof, Thomas C., & Brose, Nils. 1999. Munc13-1 is essential for fusion competence of glutamatergic synaptic vesicles. *Nature*, **400**(6743), 457–461.

Baines, Richard A., Uhler, Jay P., Thompson, Annemarie, Sweeney, Sean T., & Bate, Michael. 2001. Altered Electrical Properties in DrosophilaNeurons Developing without Synaptic Transmission. *The Journal of Neuroscience*, **21**(5), 1523–1531.

Bay, Virginia, & Butt, Arthur M. 2012. Relationship between glial potassium regulation and axon excitability: A role for glial Kir4.1 channels. *Glia*, **60**(4), 651–660.

Beier, Kevin T., Borghuis, Bart G., El-Danaf, Rana N., Huberman, Andrew D., Demb, Jonathan B., & Cepko, Constance L. 2013. Transsynaptic Tracing with Vesicular

Stomatitis Virus Reveals Novel Retinal Circuitry. *The Journal of Neuroscience*, **33**(1), 35–51.

Bellen, Hugo J., Tong, Chao, & Tsuda, Hiroshi. 2010. 100 years of Drosophila research and its impact on vertebrate neuroscience: a history lesson for the future. *Nat Rev Neurosci*, **11**(7), 514–522.

Benton, Richard, Vannice, Kirsten S., Gomez-Diaz, Carolina, & Vosshall, Leslie B. 2009. Variant Ionotropic Glutamate Receptors as Chemosensory Receptors in Drosophila. *Cell*, **136**(1), 149 – 162.

Beshel, Jennifer, & Zhong, Yi. 2013. Graded Encoding of Food Odor Value in the Drosophila Brain. *The Journal of Neuroscience*, **33**(40), 15693–15704.

Bier, E, Vaessin, H, Shepherd, S, Lee, K, McCall, K, Barbel, S, Ackerman, L, Carretto, R, Uemura, T, & Grell, E. 1989. Searching for pattern and mutation in the Drosophila genome with a P-lacZ vector. *Genes & Development*, **3**(9), 1273–1287.

Bischof, Johannes, Maeda, Robert K., Hediger, Monika, Karch, FranÃğois, & Basler, Konrad. 2007. An optimized transgenesis system for Drosophila using germ-line-specific ÏĘC31 integrases. *Proceedings of the National Academy of Sciences*, **104**(9), 3312–3317.

Bliss, T. V. P., & Collingridge, G. L. 1993. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature*, **361**(6407), 31–39.

Bliss, T. V. P., & Lomo, T. 1973. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *The Journal of Physiology*, **232**(2), 331–356.

Bolger, Tony. 2011. Trimmomatic: A flexible read trimming tool for Illumina NGS data. *The Journal of Physiology*, June.

Brand, A.H., & Perrimon, N. 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*, **118**(2), 401–415.

Briggman, Kevin L, & Denk, Winfried. 2006. Towards neural circuit reconstruction with volume electron microscopy techniques. *Current Opinion in Neurobiology*, **16**(5), 562 – 570. <ce:title>Neuronal and glial cell biology / New technologies</ce:title>.

Burrone, Juan, O'Byrne, Michael, & Murthy, Venkatesh N. 2002. Multiple forms of synaptic plasticity triggered by selective suppression of activity in individual neurons. *Nature*, **420**(6914), 414–418.

Busto, Germain U., Cervantes-Sandoval, Isaac, & Davis, Ronald L. 2010. Olfactory Learning in Drosophila. *Physiology*, **25**(6), 338–346.

Butala, M., Å¡gur Bertok, D., & Busby, S.J.W. 2009. The bacterial LexA transcriptional repressor. *Cellular and Molecular Life Sciences*, **66**(1), 82–93.

Butt, Arthur M., & Kalsi, Amanpreet. 2006. Inwardly rectifying potassium channels (Kir) in central nervous system glia: a special role for Kir4.1 in glial functions. *Journal of Cellular and Molecular Medicine*, **10**(1), 33–44.

Cajal, Santiago Ramon Y. 1906. The structure and connexions of neurons. *In: Nobel Lecture*.

Callaway, Edward M. 2008. Transneuronal circuit tracing with neurotropic viruses. *Current Opinion in Neurobiology*, **18**(6), 617 – 623.

Carlsson, Mikael A., Diesner, Max, Schachtner, Joachim, & NÃd'ssel, Dick R. 2010. Multiple neuropeptides in the Drosophila antennal lobe suggest complex modulatory circuits. *The Journal of Comparative Neurology*, **518**(16), 3359–3380.

Carpenter, Frederic W. 1905. The Reactions of the Pomace Fly (Drosophila ampelophila Loew) to Light, Gravity, and Mechanical Stimulation. *The American Naturalist*, **39**(459), pp. 157–171.

Castellucci, Vincent, Pinsker, Harold, Kupfermann, Irving, & Kandel, Eric R. 1970. Neuronal Mechanisms of Habituation and Dishabituation of the Gill-Withdrawal Reflex in Aplysia. *Science*, **167**(3926), 1745–1748.

Catalan, Ana, Hutter, Stephan, & Parsch, John. 2012. Population and sex differences in Drosophila melanogaster brain gene expression. *BMC Genomics*, **13**(1), 654.

Ceriani, M. Fernanda, Hogenesch, John B., Yanovsky, Marcelo, Panda, Satchidananda, Straume, Martin, & Kay, Steve A. 2002. Genome-Wide Expression Analysis in DrosophilaReveals Genes Controlling Circadian Behavior. *The Journal of Neuroscience*, **22**(21), 9305–9319.

Chen, Xu, & Ganetzky, Barry. 2012. A neuropeptide signaling pathway regulates synaptic growth in Drosophila. *The Journal of Cell Biology*, **196**(4), 529–543.

Chou, Ya-Hui, Spletter, Maria L, Yaksi, Emre, Leong, Jonathan C S, Wilson, Rachel I, & Luo, Liqun. 2010. Diversity and wiring variability of olfactory local interneurons in the Drosophila antennal lobe. *Nat Neurosci*, **13**(4), 439–449.

Cirelli, Chaira, & Bushey, Daniel. 2008. Sleep and Wakefulness inDrosophila melanogaster. *Annals of the New York Academy of Sciences*, **1129**(1), 323–329.

Claridge-Chang, Adam, Wijnen, Herman, Naef, Felix, Boothroyd, Catharine, Rajewsky, Nikolaus, & Young, Michael W. 2001. Circadian Regulation of Gene Expression Systems in the Drosophila Head. *Neuron*, **32**(4), 657 – 671.

Claridge-Chang, Adam, Roorda, Robert D., Vrontou, Eleftheria, Sjulson, Lucas, Li, Haiyan, Hirsh, Jay, & MiesenbÃűck, Gero. 2009. Writing Memories with Light-Addressable Reinforcement Circuitry. *Cell*, **139**(2), 405 – 415.

Clark, Tyson A., Sugnet, Charles W., & Ares, Manuel. 2002. Genomewide Analysis of mRNA Processing in Yeast Using Splicing-Specific Microarrays. *Science*, **296**(5569), 907–910.

Clyne, Peter J., Warr, Coral G., & Carlson, John R. 2000. Candidate Taste Receptors in Drosophila. *Science*, **287**(5459), 1830–1834.

Consortium, The ENCODE Project. 2004. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science*, **306**(5696), 636–640.

Cooper, Steven J. 2008. From Claude Bernard to Walter Cannon. Emergence of the concept of homeostasis. *Appetite*, **51**(3), 419 – 427.

Couto, Africa, Alenius, Mattias, & Dickson, Barry J. 2005. Molecular, Anatomical, and Functional Organization of the Drosophila Olfactory System. *Current Biology*, **15**(17), 1535 – 1547.

Cowan, W.M., Gottlieb, D.I., Hendrickson, Anita E., Price, J.L., & Woolsey, T.A. 1972. The autoradiographic demonstration of axonal connections in the central nervous system. *Brain Research*, **37**(1), 21 – 51.

Cowan, W.M., Südhof, T.C., Stevens, C.F., & Institute, Howard Hughes Medical. 2001. *Synapses*. Synapses. Johns Hopkins University Press.

Crawford, Devon C., Jiang, Xiaoping, Taylor, Amanda, Moulder, Krista L., & Mennerick, Steven. 2012. Differential Requirement for Protein Synthesis in Presynaptic Unmuting and Muting in Hippocampal Glutamate Terminals. *PLoS ONE*, **7**(12), e51930.

Crick, Francis. 1995. *Astonishing Hypothesis: The Scientific Search for the Soul*. A Touchstone Book. Scribner.

Croset, Vincent, Rytz, Raphael, Cummins, Scott F., Budd, Aidan, Brawand, David, Kaessmann, Henrik, Gibson, Toby J., & Benton, Richard. 2010. Ancient Protostome Origin of Chemosensory Ionotropic Glutamate Receptors and the Evolution of Insect Taste and Olfaction. *PLoS Genet*, **6**(8), e1001064.

Dacks, Andrew M., Green, David S., Root, Cory M., Nighorn, Alan J., & Wang, Jing W. 2009. Serotonin Modulates Olfactory Processing in the Antennal Lobe of Drosophila. *Journal of Neurogenetics*, **23**(4), 366–377. PMID: 19863268.

Damasio, A. 2008. *Descartes' Error: Emotion, Reason and the Human Brain*. Random House.

DasGupta, Shamik, & Waddell, Scott. 2008. Learned Odor Discrimination in Drosophila without Combinatorial Odor Maps in the Antennal Lobe. *Current Biology*, **18**(21), 1668 – 1674.

Davis, Graeme W. 2006. HOMEOSTATIC CONTROL OF NEURAL ACTIVITY: From Phenomenology to Molecular Design. *Annual Review of Neuroscience*, **29**(1), 307–323. PMID: 16776588.

De Gois, StÃľphanie, SchÃdfer, Martin K.-H., Defamie, Norah, Chen, Chu, Ricci, Anthony, Weihe, Eberhard, Varoqui, HÃlÃÍne, & Erickson, Jeffrey D. 2005. Homeostatic Scaling of Vesicular Glutamate and GABA Transporter Expression in Rat Neocortical Circuits. *The Journal of Neuroscience*, **25**(31), 7121–7133.

Deng, Wei, Aimone, James B., & Gage, Fred H. 2010. New neurons and new memories: how does adult hippocampal neurogenesis affect learning and memory? *Nat Rev Neurosci*, **11**(5), 339–350.

Denk, Winfried, & Horstmann, Heinz. 2004. Serial Block-Face Scanning Electron Microscopy to Reconstruct Three-Dimensional Tissue Nanostructure. *PLoS Biol*, **2**(11), e329.

der Weel, M.J. Dolleman-Van, Wouterlood, F.G., & Witter, M.P. 1994. Multiple anterograde tracing, combining Phaseolus vulgaris leucoagglutinin with rhodamine- and biotin-conjugated dextran amine. *Journal of Neuroscience Methods*, **51**(1), 9 – 21.

Diegelmann, Soeren, Bate, Michael, & Landgraf, Matthais. 2008. Gateway cloning vectors for the LexA-based binary expression system in Drosophila. *fly*, **2**(1933-6934), 236–239.

Dobritsa, Anna A., van der Goes van Naters, Wynand, Warr, Coral G., Steinbrecht, R.Alexander, & Carlson, John R. 2003. Integrating the Molecular and Cellular Basis of Odor Coding in the Drosophila Antenna. *Neuron*, **37**(5), 827 – 841.

Donlea, Jeffrey M., Thimgan, Matthew S., Suzuki, Yasuko, Gottschalk, Laura, & Shaw, Paul J. 2011. Inducing Sleep by Remote Control Facilitates Memory Consolidation in Drosophila. *Science*, **332**(6037), 1571–1576.

Dubnau, Josh, Chiang, Ann-Shyn, Grady, Lori, Barditch, Jody, Gossweiler, Scott, McNeil, John, Smith, Patrick, Buldoc, Francois, Scott, Rod, Certa, Uli, Broger, Clemens, & Tully, Tim. 2003. The staufen/pumilio Pathway Is Involved in Drosophila Long-Term Memory. *Current Biology*, **13**(4), 286 – 296.

Dunipace, Leslie, Meister, Stephan, McNealy, Corum, & Amrein, Hubert. 2001. Spatially restricted expression of candidate taste receptors in the Drosophila gustatory system. *Current Biology*, **11**(11), 822 – 835.

Ekstrand, Mats I., Enquist, L.W., & Pomeranz, Lisa E. 2008. The alpha-herpesviruses: molecular pathfinders in nervous system circuits. *Trends in Molecular Medicine*, **14**(3), 134 – 140.

Emery, Patrick, So, W.Venus, Kaneko, Maki, Hall, Jeffrey C, & Rosbash, Michael. 1998. CRY, a Drosophila Clock and Light-Regulated Cryptochrome, Is a Major Contributor to Circadian Rhythm Resetting and Photosensitivity. *Cell*, **95**(5), 669 – 679.

Fernandez-Hernandez, Ismael, Rhiner, Christa, & Moreno, Eduardo. 2013. Adult Neurogenesis in Drosophila. *Cell Reports*, **3**(6), 1857 – 1865.

Fields, B.N., Knipe, D.M., & Howley, P.M. 2007. *Fields' Virology*. Fields' Virology, no. v. 1. Wolters Kluwer Health/Lippincott Williams & Wilkins.

Finger, S. 2005. *Minds Behind the Brain: A History of the Pioneers and Their Discoveries*. Oxford University Press, USA.

Frank, C. Andrew, Kennedy, Matthew J., Goold, Carleton P., Marek, Kurt W., & Davis, Graeme W. 2006. Mechanisms Underlying the Rapid Induction and Sustained Expression of Synaptic Homeostasis. *Neuron*, **52**(4), 663 – 677.

Frank, C. Andrew, Pielage, Jan, & Davis, Graeme W. 2009. A Presynaptic Homeostatic Signaling System Composed of the Eph Receptor, Ephexin, Cdc42, and CaV2.1 Calcium Channels. *Neuron*, **61**(4), 556 – 569.

Furshpan, E. J., & Potter, D. D. 1957. Mechanism of Nerve-Impulse Transmission at a Crayfish Synapse. *Nature*, **180**(4581), 342–343.

Ganguly-Fitzgerald, Indrani, Donlea, Jeff, & Shaw, Paul J. 2006. Waking Experience Affects Sleep Need in Drosophila. *Science*, **313**(5794), 1775–1781.

Glanzman, David L. 2010. Common Mechanisms of Synaptic Plasticity in Vertebrates and Invertebrates. *Current Biology*, **20**(1), R31 – R36.

Golgi, Camillo. 1906. The neuron doctrine - theory and facts. *In: Nobel Lecture*.

Goold, Carleton P., & Davis, Graeme W. 2007. The {BMP} Ligand Gbb Gates the Expression of Synaptic Homeostasis Independent of Synaptic Growth Control. *Neuron*, **56**(1), 109 – 123.

Goold, Carleton P., & Nicoll, Roger A. 2010. Single-Cell Optogenetic Excitation Drives Homeostatic Synaptic Depression. *Neuron*, **68**(3), 512 – 528.

Graf, Ethan R., Daniels, Richard W., Burgess, Robert W., Schwarz, Thomas L., & DiAntonio, Aaron. 2009. Rab3 Dynamically Controls Protein Composition at Active Zones. *Neuron*, **64**(5), 663 – 677.

Graveley, Brenton R., Brooks, Angela N., Carlson, Joseph W., Duff, Michael O., Landolin, Jane M., Yang, Li, Artieri, Carlo G., van Baren, Marijke J., Boley, Nathan, Booth, Benjamin W., Brown, James B., Cherbas, Lucy, Davis, Carrie A., Dobin, Alex, Li, Renhua, Lin, Wei, Malone, John H., Mattiuzzo, Nicolas R., Miller, David, Sturgill, David, Tuch, Brian B., Zaleski, Chris, Zhang, Dayu, Blanchette, Marco, Dudoit, Sandrine, Eads, Brian, Green, Richard E., Hammonds, Ann, Jiang, Lichun, Kapranov, Phil, Langton, Laura, Perrimon, Norbert, Sandler, Jeremy E., Wan, Kenneth H., Willingham, Aarron, Zhang, Yu, Zou, Yi, Andrews, Justen, Bickel, Peter J., Brenner, Steven E., Brent, Michael R., Cherbas, Peter, Gingeras, Thomas R., Hoskins, Roger A., Kaufman, Thomas C., Oliver, Brian, & Celniker, Susan E. 2011. The developmental transcriptome of Drosophila melanogaster. *Nature*, **471**(7339), 473–479.

Greenspan, Ralph J., & Ferveur, Jean-Francois. 2000. COURTSHIP IN DROSOPHILA. *Annual Review of Genetics*, **34**(1), 205–232. PMID: 11092827.

Gross, Charles G. 1987. *Encyclopedia of Neuroscience*. BirkhŁuser.

Hallem, Elissa A., & Carlson, John R. 2006. Coding of Odors by a Receptor Repertoire. *Cell*, **125**(1), 143 – 160.

Han, Edward B., & Stevens, Charles F. 2009. Development regulates a switch between post- and presynaptic strengthening in response to activity deprivation. *Proceedings of the National Academy of Sciences*.

Hao, Linhui, Sakurai, Akira, Watanabe, Tokiko, Sorensen, Ericka, Nidom, Chairul A., Newton, Michael A., Ahlquist, Paul, & Kawaoka, Yoshihiro. 2008. Drosophila RNAi screen identifies host genes important for influenza virus replication. *Nature*, **454**(7206), 890–893.

Harbers, M., & Kahl, G. 2012. *Tag-based Next Generation Sequencing*. Molecular Plant Biology. Wiley.

Harbison, Susan, Chang, Sherman, Kamdar, Kim, & Mackay, Trudy. 2005. Quantitative genomics of starvation stress resistance in Drosophila. *Genome Biology*, **6**(4), R36.

Haruna, Yoshisumi, Kobori, Atsushi, Makiyama, Takeru, Yoshida, Hidetada, Akao, Masaharu, Doi, Takahiro, Tsuji, Keiko, Ono, Seiko, Nishio, Yukiko, Shimizu, Wataru, Inoue, Takehiko, Murakami, Tomoaki, Tsuboi, Naoya, Yamanouchi, Hideo, Ushinohama, Hiroya, Nakamura, Yoshihide, Yoshinaga, Masao, Horigome, Hitoshi, Aizawa, Yoshifusa, Kita, Toru, & Horie, Minoru. 2007. Genotype-phenotype correlations of KCNJ2 mutations in Japanese patients with Andersen-Tawil syndrome. *Human Mutation*, **28**(2), 208–208.

Haucke, Volker, Neher, Erwin, & Sigrist, Stephan J. 2011. Protein scaffolds in the coupling of synaptic exocytosis and endocytosis. *Nat Rev Neurosci*, **12**(3), 127–138.

Hayworth, KJ, Kasthuri, N, Schalek, R, & Lichtman, JW. 2006. Automating the Collection of Ultrathin Serial Sections for Large Volume TEM Reconstructions. *Microscopy and Microanalysis*, **12**(8), 86–87.

Helfrich-Forster, Charlotte, Tauber, Marcus, Park, Jae H., Muller, Max Muhlig-Versen, Max, Schneuwly, Stephan, & Hofbauer, Alois. 2000. Ectopic Expression of the Neuropeptide Pigment-Dispersing Factor Alters Behavioral Rhythms in Drosophila melanogaster. *The Journal of Neuroscience*, **20**(9), 3339–3353.

Helmstaedter, Moritz, Briggman, Kevin L., Turaga, Srinivas C., Jain, Viren, Seung, H. Sebastian, & Denk, Winfried. 2013. Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature*, **500**(7461), 168–174.

Henry, Gilbert L., Davis, Fred P., Picard, Serge, & Eddy, Sean R. 2012. Cell type specific genomics of Drosophila neurons. *Nucleic Acids Research*, **40**(19), 9691–9704.

Hens, Korneel. 2014. System for the non-linear amplification of gene expression differences for optical reporters. *Peronal Communication*, February.

Hodgkin, A. L., & Huxley, A. F. 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, **117**(4), 500–544.

Hong, Weizhe, Mosca, Timothy J., & Luo, Liqun. 2012. Teneurins instruct synaptic partner matching in an olfactory map. *Nature*, **484**(7393), 201–207.

Hughes, Michael E., Grant, Gregory R., Paquin, Christina, Qian, Jack, & Nitabach, Michael N. 2012. Deep sequencing the circadian and diurnal transcriptome of Drosophila brain. *Genome Research*, **22**(7), 1266–1281.

Huh, Y, Oh, MS, Leblanc, P, & Kim, KS. 2010. Gene transfer in the nervous system and implications for transsynaptic neuronal tracing. *Expert Opin Biol Ther*, **10**(5), 763–772–.

Inagaki, Hidehiko K., de Leon, Shlomo Ben-Tabou, Wong, Allan M., Jagadish, Smitha, Ishimoto, Hiroshi, Barnea, Gilad, Kitamoto, Toshihiro, Axel, Richard, & Anderson, David J. 2012. Visualizing Neuromodulation In Vivo: TANGO-Mapping of Dopamine Signaling Reveals Appetite Control of Sugar Sensing. *Cell*, **148**(3), 583 – 595.

Jefferis, Gregory S. X. E., Marin, Elizabeth C., Stocker, Reinhard F., & Luo, Liqun. 2001. Target neuron prespecification in the olfactory map of Drosophila. *Nature*, **414**(6860), 204–208.

Jenett, Arnim, Rubin, Gerald M., Ngo, Teri-T.B., Shepherd, David, Murphy, Christine, Dionne, Heather, Pfeiffer, Barret D., Cavallaro, Amanda, Hall, Donald, Jeter, Jennifer, Iyer, Nirmala, Fetter, Dona, Hausenfluck, Joanna H., Peng, Hanchuan, Trautman, Eric T., Svirskas, Robert R., Myers, Eugene W., Iwinski, Zbigniew R., Aso, Yoshinori, DePasquale, Gina M., Enos, Adrianne, Hulamm, Phuson, Lam, Shing Chun Benny, Li, Hsing-Hsi, Laverty, Todd R., Long, Fuhui, Qu, Lei, Murphy, Sean D., Rokicki, Konrad, Safford, Todd, Shaw, Kshiti, Simpson, Julie H., Sowell, Allison, Tae, Susana, Yu, Yang, & Zugates, Christopher T. 2012. A GAL4-Driver Line Resource for Drosophila Neurobiology. *Cell Reports*, **2**(4), 991 – 1001.

Johns, David C., Marx, Ruth, Mains, Richard E., OâĂŹRourke, Brian, & MarbÃąn, Eduardo. 1999. Inducible Genetic Suppression of Neuronal Excitability. *The Journal of Neuroscience*, **19**(5), 1691–1697.

Kato, Kentaro, Awasaki, Takeshi, & Ito, Kei. 2009. Neuronal programmed cell death induces glial cell division in the adult Drosophila brain. *Development*, **136**(1), 51–59.

Katz, L. C., Burkhalter, A., & Dreyer, W. J. 1984. Fluorescent latex microspheres as a retrograde neuronal marker for in vivo and in vitro studies of visual cortex. *Nature*, **310**(5977), 498–500.

Katzel, Dennis, Zemelman, Boris V, Buetfering, Christina, Wolfel, Markus, & Miesenbock, Gero. 2011. The columnar and laminar organization of inhibitory connections to neocortical excitatory cells. *Nat Neurosci*, **14**(1), 100–107.

Kazama, Hokto, & Wilson, Rachel I. 2008. Homeostatic Matching and Nonlinear Amplification at Identified Central Synapses. *Neuron*, **58**(3), 401 – 413.

Kazama, Hokto, & Wilson, Rachel I. 2009. Origins of correlated activity in an olfactory circuit. *Nat Neurosci*, **12**(9), 1136–1144.

Keegan, Kevin P, Pradhan, Suraj, Wang, Ji-Ping, & Allada, Ravi. 2007. Meta-Analysis of <italic>Drosophila</italic> Circadian Microarray Studies Identifies a Novel Set of Rhythmically Expressed Genes. *PLoS Comput Biol*, **3**(11), e208.

Keene, Alex C., & Waddell, Scott. 2007. Drosophila olfactory memory: single genes to complex neural circuits. *Nat Rev Neurosci*, **8**(5), 341–354.

Keller, Andreas, Sweeney, Sean T., Zars, Troy, O'Kane, Cahir J., & Heisenberg, Martin. 2002. Targeted expression of tetanus neurotoxin interferes with behavioral responses to sensory input in Drosophila. *Journal of Neurobiology*, **50**(3), 221–233.

Kim, Hyung-Goo, Kishikawa, Shotaro, Higgins, Anne W., Seong, Ihn-Sik, Donovan, Diana J., Shen, Yiping, Lally, Eric, Weiss, Lauren A., Najm, Juliane, Kutsche, Kerstin, Descartes, Maria, Holt, Lynn, Braddock, Stephen, Troxell, Robin, Kaplan, Lee, Volkmar, Fred, Klin, Ami, Tsatsanis, Katherine, Harris, David J., Noens, Ilse, Pauls, David L., Daly, Mark J., MacDonald, Marcy E., Morton, Cynthia C., Quade, Bradley J., & Gusella, James F. 2008. Disruption of Neurexin 1 Associated with Autism Spectrum Disorder. *The American Journal of Human Genetics*, **82**(1), 199–207.

Kitamoto, Toshihiro. 2001. Conditional modification of behavior in Drosophila by targeted expression of a temperature-sensitive shibire allele in defined neurons. *Journal of Neurobiology*, **47**(2), 81–92.

Kleinfeld, David, Bharioke, Arjun, Blinder, Pablo, Bock, Davi D., Briggman, Kevin L., Chklovskii, Dmitri B., Denk, Winfried, Helmstaedter, Moritz, Kaufhold, John P., Lee, Wei-Chung Allen, Meyer, Hanno S., Micheva, Kristina D., Oberlaender, Marcel, Prohaska, Steffen, Reid, R. Clay, Smith, Stephen J., Takemura, Shinya, Tsai, Philbert S., & Sakmann, Bert. 2011. Large-Scale Automated Histology in the Pursuit of Connectomes. *The Journal of Neuroscience*, **31**(45), 16125–16138.

Knaden, Markus, Strutz, Antonia, Ahsan, Jawaid, Sachse, Silke, & Hansson, Bill S. 2012. Spatial Representation of Odorant Valence in an Insect Brain. *Cell Reports*, **1**(4), 392 – 399.

Knight, David, Xie, Wei, & Boulianne, GabrielleL. 2011. Neurexins and Neuroligins: Recent Insights from Invertebrates. *Molecular Neurobiology*, **44**(3), 426–440.

Kobbert, C., Apps, R., Bechmann, I., Lanciego, J.L., Mey, J., & Thanos, S. 2000. Current concepts in neuroanatomical tracing. *Progress in Neurobiology*, **62**(4), 327 – 351.

Kofuji, Paulo, Biedermann, Bernd, Siddharthan, Venkatraman, Raap, Maik, Iandiev, Ian, Milenkovic, Ivan, Thomzig, Achim, Veh, Rodiger W., Bringmann, Andreas, & Reichenbach, Andreas. 2002. Kir potassium channel subunit expression in retinal glial cells: Implications for spatial potassium buffering. *Glia*, **39**(3), 292–303.

Kohler, R.E. 1994. *Lords of the Fly: Drosophila Genetics and the Experimental Life*. History, philosophy, and social studies of science : Biology. University of Chicago Press.

Konopka, RJ, & Benzer, S. 1971. Clock mutants of Drosophila melanogaster. *Proc Natl Acad Sci U S A*, **68**(9), 2112–2116–.

Krashes, Michael J., DasGupta, Shamik, Vreede, Andrew, White, Benjamin, Armstrong, J. Douglas, & Waddell, Scott. 2009. A Neural Circuit Mechanism Integrating Motivational State with Memory Expression in Drosophila. *Cell*, **139**(2), 416 – 427.

Kremer, Malte C., Christiansen, Frauke, Leiss, Florian, Paehler, Moritz, Knapek, Stephan, Andlauer, Till F.M., FÃűrstner, Friedrich, Kloppenburg, Peter, Sigrist, Stephan J., & Tavosanis, Gaia. 2010. Structural Long-Term Changes at Mushroom Body Input Synapses. *Current Biology*, **20**(21), 1938 – 1944.

Krishnan, Balaji, Dryer, Stuart E., & Hardin, Paul E. 1999. Circadian rhythms in olfactory responses of Drosophila melanogaster. *Nature*, **400**(6742), 375–378.

Kristensson, Krister, & Olsson, Yngve. 1971. Uptake and retrograde axonal transport of peroxidase in hypoglossal neurones. *Acta Neuropathologica*, **19**(1), 1–9.

Kurtovic, Amina, Widmer, Alexandre, & Dickson, Barry J. 2007. A single class of olfactory neurons mediates behavioural responses to a Drosophila sex pheromone. *Nature*, **446**(7135), 542–546.

Kuypers, H.G.J.M., Bentivoglio, M., Catsman-Berrevoets, C.E., & Bharos, A.T. 1980. Double retrograde neuronal labeling through divergent axon collaterals, using two fluorescent tracers with the same excitation wavelength which label different features of the cell. *Experimental Brain Research*, **40**(4), 383–392.

Kwon, Jae Young, Dahanukar, Anupama, Weiss, Linnea A., & Carlson, John R. 2011. Molecular and Cellular Organization of the Taste System in the Drosophila Larva. *The Journal of Neuroscience*, **31**(43), 15300–15309.

Lafon, Monique. 2005. Rabies virus receptors. *Journal of NeuroVirology*, **11**(1), 82–87.

Lai, Sen-Lin, & Lee, Tzumin. 2006. Genetic mosaic with dual binary transcriptional systems in Drosophila. *Nat Neurosci*, **9**(5), 703–709.

Lai, Sen-Lin, Awasaki, Takeshi, Ito, Kei, & Lee, Tzumin. 2008. Clonal analysis of Drosophila antennal lobe neurons: diverse neuronal architectures in the lateral neuroblast lineage. *Development*, **135**(17), 2883–2893.

Laissue, PhilippeP., & Vosshall, LeslieB. 2008. The Olfactory Sensory Map in Drosophila. *Pages 102–114 of:* Technau, GerhardM. (ed), *Brain Development in Drosophila melanogaster*. Advances in Experimental Medicine and Biology, vol. 628. Springer New York.

Lanciego, JosÃĺ L., & Wouterlood, Floris G. 2011. A half century of experimental neuroanatomical tracing. *Journal of Chemical Neuroanatomy*, **42**(3), 157 – 183.

Langmead, Ben, & Salzberg, Steven L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Meth*, **9**(4), 357–359.

Larsson, Mattias C., Domingos, Ana I., Jones, Walton D., Chiappe, M.Eugenia, Amrein, Hubert, & Vosshall, Leslie B. 2004. Or83b Encodes a Broadly Expressed Odorant Receptor Essential for Drosophila Olfaction. *Neuron*, **43**(5), 703 – 714.

LaVail, J.H., Tauscher, A.N., Sucher, A., Harrabi, O., & Brandimarti, R. 2007. Viral regulation of the long distance axonal transport of herpes simplex virus nucleocapsid. *Neuroscience*, **146**(3), 974 – 985.

Lazarevic, Vesna, Pothula, Santosh, Andres-Alonso, Maria, & Fejtova, Anna. 2013. Molecular mechanisms driving homeostatic plasticity of neurotransmitter release. *Frontiers in Cellular Neuroscience*, **7**.

Lee, Tzumin, & Luo, Liqun. 1999. Mosaic Analysis with a Repressible Cell Marker for Studies of Gene Function in Neuronal Morphogenesis. *Neuron*, **22**(3), 451 – 461.

Li, Jun, & Tibshirani, Robert. 2013. Finding consistent patterns: A nonparametric approach for identifying differential expression in RNA-Seq data. *Statistical Methods in Medical Research*, **22**(5), 519–536.

Li, Jun, Witten, Daniela M., Johnstone, Iain M., & Tibshirani, Robert. 2012. Normalization, testing, and false discovery rate estimation for RNA-sequencing data. *Biostatistics*, **13**(3), 523–538.

Lichtman, Jeff W., Livet, Jean, & Sanes, Joshua R. 2008. A technicolour approach to the connectome. *Nat Rev Neurosci*, **9**(6), 417–422.

Lieber, Toby, Kidd, Simon, & Struhl, Gary. 2011. DSL-Notch Signaling in the Drosophila Brain in Response to Olfactory Stimulation. *Neuron*, **69**(3), 468 – 481.

Lin, Yiing, Han, Mei, Shimada, Brian, Wang, Lin, Gibler, Therese M., Amarakone, Aloka, Awad, Tarif A., Stormo, Gary D., Van Gelder, Russell N., & Taghert, Paul H. 2002. Influence of the period-dependent circadian clock on diurnal, circadian, and aperiodic gene expression in Drosophila melanogaster. *Proceedings of the National Academy of Sciences*, **99**(14), 9562–9567.

Ling, Daijun, & Salvaterra, Paul M. 2011. Robust RT-qPCR Data Normalization: Validation and Selection of Internal Reference Genes during Post-Experimental Data Analysis. *PLoS ONE*, **6**(3), e17762.

Llinas, Rodolfo R. 2003. The contribution of Santiago Ramon y Cajal to functional neuroscience. *Nat Rev Neurosci*, **4**(1), 77–80.

Lo, Liching, & Anderson, David J. 2011. A Cre-Dependent, Anterograde Transsynaptic Viral Tracer for Mapping Output Pathways of Genetically Marked Neurons. *Neuron*, **72**(6), 938 – 950.

Lopez-Munoz, Francisco, Boya, Jesus, & Alamo, Cecilio. 2006. Neuron theory, the cornerstone of neuroscience, on the centenary of the Nobel Prize award to Santiago RamÃşn y Cajal. *Brain Research Bulletin*, **70**(4âĂŞ6), 391–405.

Love, Michael I, Huber, Wolfgang, & Anders, Simon. 2014. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *bioRxiv*.

Luan, Haojiang, Peabody, Nathan C., Vinson, Charles R., & White, Benjamin H. 2006. Refined Spatial Manipulation of Neuronal Function by Combinatorial Restriction of Transgene Expression. *Neuron*, **52**(3), 425 – 436.

Lunter, Gerton, & Goodson, Martin. 2011. Stampy: A statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Research*, **21**(6), 936–939.

Luo, Liqun, Callaway, Edward M., & Svoboda, Karel. 2008. Genetic Dissection of Neural Circuits. *Neuron*, **57**(5), 634 – 660.

Luo, Weijun, Friedman, Michael, Shedden, Kerby, Hankenson, Kurt, & Woolf, Peter. 2009. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics*, **10**(1), 161.

Luscher, C., Nicoll, R. A., Malenka, R. C., & Muller, D. 2000. Synaptic plasticity and dynamic modulation of the postsynaptic membrane. *Nat Neurosci*, **3**(6), 545–550.

Malmgren, Leslie, Olsson, Yngve, Olsson, Tomas, & Kristensson, Krister. 1978. Uptake and retrograde axonal transport of various exogenous macromolecules in normal and crushed hypoglossal nerves. *Brain Research*, **153**(3), 477 – 493.

Manseau, Lynn, Baradaran, Ali, Brower, Danny, Budhu, Anuradha, Elefant, Felice, Phan, Huy, Philp, Alastair Valentine, Yang, Mingyao, Glover, David, Kaiser, Kim, Palter, Karen, & Selleck, Scott. 1997. GAL4 enhancer traps expressed in the embryo, larval brain, imaginal discs, and ovary of drosophila. *Developmental Dynamics*, **209**(3), 310–322.

Marder, Eve, & Goaillard, Jean-Marc. 2006. Variability, compensation and homeostasis in neuron and network function. *Nat Rev Neurosci*, **7**(7), 563–574.

Marie, Bruno, Pym, Edward, Bergquist, Sharon, & Davis, Graeme W. 2010. Synaptic Homeostasis Is Consolidated by the Cell Fate Gene gooseberry, a Drosophila pax3/7 Homolog. *The Journal of Neuroscience*, **30**(24), 8071–8082.

Maskos, Uwe, Kissa, Karima, St. Cloment, CÃľcile, & BrÃżlet, Philippe. 2002. Retrograde trans-synaptic transfer of green fluorescent protein allows the genetic mapping of neuronal circuits in transgenic mice. *Proceedings of the National Academy of Sciences*, **99**(15), 10120–10125.

Mathew, Dennis, Martelli, Carlotta, Kelley-Swift, Elizabeth, Brusalis, Christopher, Gershow, Marc, Samuel, Aravinthan D. T., Emonet, Thierry, & Carlson, John R. 2013. Functional diversity among sensory receptors in a Drosophila olfactory circuit. *Proceedings of the National Academy of Sciences*, **110**(23), E2134–E2143.

McDonald, Michael J., & Rosbash, Michael. 2001. Microarray Analysis and Organization of Circadian Gene Expression in Drosophila. *Cell*, **107**(5), 567 – 578.

McGuire, Sean E., Mao, Zhengmei, & Davis, Ronald L. 2004. Spatiotemporal Gene Expression Targeting with the TARGET and Gene-Switch Systems in Drosophila. *Sci. STKE*, **2004**(220), pl6.

Menegazzi, Pamela, Yoshii, Taishi, & Helfrich-FÃűrster, Charlotte. 2012. Laboratory versus Nature: The Two Sides of the Drosophila Circadian Clock. *Journal of Biological Rhythms*, **27**(6), 433–442.

Miesenboeck, Gero, & Kevrekidis, Ioannis G. 2005. OPTICAL IMAGING AND CONTROL OF GENETICALLY DESIGNATED NEURONS IN FUNCTIONING CIRCUITS. *Annual Review of Neuroscience*, **28**(1), 533–563. PMID: 16022604.

Miyamoto, Tetsuya, & Amrein, Hubert. 2008. Suppression of male courtship by a Drosophila pheromone receptor. *Nat Neurosci*, **11**(8), 874–876.

Miyamoto, Tetsuya, Chen, Yan, Slone, Jesse, & Amrein, Hubert. 2013. Identification of a <italic>Drosophila</italic> Glucose Receptor Using Ca²⁺ Imaging of Single Chemosensory Neurons. *PLoS ONE*, **8**(2), e56304.

Mosca, Timothy. 2013. The organization of olfactory synapses and their regulation by the cell surface receptors Teneurin-a and Dalek. *In: Brain, Behaviour, and Evolution*.

Motojiro, Yoshihara, & Ito, Kei. 2000. Improved Gal4 screening kit for large-scale generation of enhancer-trap strains. *Dros. Inf. Serv.*, **83**, 199–202.

Muller, H. J. 1927. ARTIFICIAL TRANSMUTATION OF THE GENE. *Science*, **66**(1699), 84–87.

Muller, Martin, Pym, Edward C.G., Tong, Amy, & Davis, Graeme W. 2011. Rab3-GAP Controls the Progression of Synaptic Homeostasis at a Late Stage of Vesicle Release. *Neuron*, **69**(4), 749 – 762.

Muller, Martin, Liu, Karen Suk Yin, Sigrist, Stephan J., & Davis, Graeme W. 2012. RIM Controls Homeostatic Plasticity through Modulation of the Readily-Releasable Vesicle Pool. *The Journal of Neuroscience*, **32**(47), 16574–16585.

Nakajima, Yu-ichiro, Kuranaga, Erina, Sugimura, Kaoru, Miyawaki, Atsushi, & Miura, Masayuki. 2011. Nonautonomous Apoptosis Is Triggered by Local Cell Cycle Progression during Epithelial Replacement in Drosophila. *Molecular and Cellular Biology*, **31**(12), 2499–2512.

Ni, Lina, Bronk, Peter, Chang, Elaine C., Lowell, April M., Flam, Juliette O., Panzano, Vincent C., Theobald, Douglas L., Griffith, Leslie C., & Garrity, Paul A. 2013. A gustatory receptor paralogue controls rapid warmth avoidance in Drosophila. *Nature*, **500**(7464), 580–584.

Nitabach, Michael N, Blau, Justin, & Holmes, Todd C. 2002. Electrical Silencing of Drosophila Pacemaker Neurons Stops the Free-Running Circadian Clock. *Cell*, **109**(4), 485 – 495.

Novikova, L, Novikov, L, & Kellerth, J.-O. 1997. Persistent neuronal labeling by retrograde fluorescent tracers: a comparison between Fast Blue, Fluoro-Gold and various dextran conjugates. *Journal of Neuroscience Methods*, **74**(1), 9 – 15.

Olsen, Shawn R., & Wilson, Rachel I. 2008. Lateral presynaptic inhibition mediates gain control in an olfactory circuit. *Nature*, **452**(7190), 956–960.

Olsen, Shawn R., Bhandawat, Vikas, & Wilson, Rachel I. 2007. Excitatory Interactions between Olfactory Processing Channels in the Drosophila Antennal Lobe. *Neuron*, **54**(1), 89 – 103.

Osterwalder, Thomas, Yoon, Kenneth S., White, Benjamin H., & Keshishian, Haig. 2001. A conditional tissue-specific transgene expression system using inducible GAL4. *Proceedings of the National Academy of Sciences*, **98**(22), 12596–12601.

Owald, David, Khorramshahi, Omid, Gupta, Varun K, Banovic, Daniel, Depner, Harald, Fouquet, Wernher, Wichmann, Carolin, Mertel, Sara, Eimer, Stefan, Reynolds, Eric, Holt, Matthew, Aberle, Hermann, & Sigrist, Stephan J. 2012. Cooperation of Syd-1

with Neurexin synchronizes pre- with postsynaptic assembly. *Nat Neurosci*, **15**(9), 1219–1226.

Paradis, Suzanne, Sweeney, Sean T, & Davis, Graeme W. 2001. Homeostatic Control of Presynaptic Release Is Triggered by Postsynaptic Membrane Depolarization. *Neuron*, **30**(3), 737 – 749.

Parnas, Moshe, Lin, Andrew C., Huetteroth, Wolf, & MiesenbÃűck, Gero. 2013. Odor Discrimination in Drosophila: From Neural Population Codes to Behavior. *Neuron*, **79**(5), 932 – 944.

Perisse, Emmanuel, Yin, Yan, Lin, Andrew C., Lin, Suewei, Huetteroth, Wolf, & Waddell, Scott. 2013. Different Kenyon Cell Populations Drive Learned Approach and Avoidance in Drosophila. *Neuron*, **79**(5), 945 – 956.

Perrat, Paola N., DasGupta, Shamik, Wang, Jie, Theurkauf, William, Weng, Zhiping, Rosbash, Michael, & Waddell, Scott. 2013. Transposition-Driven Genomic Heterogeneity in the Drosophila Brain. *Science*, **340**(6128), 91–95.

Petersen, Sophie A, Fetter, Richard D, Noordermeer, Jasprina N, Goodman, Corey S, & DiAntonio, Aaron. 1997. Genetic Analysis of Glutamate Receptors in Drosophila Reveals a Retrograde Signal Regulating Presynaptic Transmitter Release. *Neuron*, **19**(6), 1237 – 1248.

Pfeiffenberger, Cory, Lear, Bridget C., Keegan, Kevin P., & Allada, Ravi. 2010. Locomotor Activity Level Monitoring Using the Drosophila Activity Monitoring (DAM) System. *Cold Spring Harbor Protocols*, **2010**(11), pdb.prot5518.

Pfeiffer, Barret D., Ngo, Teri-T B., Hibbard, Karen L., Murphy, Christine, Jenett, Arnim, Truman, James W., & Rubin, Gerald M. 2010. Refinement of Tools for Targeted Gene Expression in Drosophila. *Genetics*, **186**(2), 735–755.

Pilgram, Gonneke S. K., Potikanond, Saranyapin, van der Plas, Mariska C., Fradkin, Lee G., & Noordermeer, Jasprina N. 2011. The RhoGAP crossveinless-c Interacts with Dystrophin and Is Required for Synaptic Homeostasis at the Drosophila Neuromuscular Junction. *The Journal of Neuroscience*, **31**(2), 492–500.

Ping, Yong, & Tsunoda, Susan. 2012. Inactivity-induced increase in nAChRs upregulates Shal K+ channels to stabilize synaptic potentials. *Nat Neurosci*, **15**(1), 90–97.

Potter, Christopher J, & Luo, Liqun. 2011. Using the Q system in Drosophila melanogaster. *Nat. Protocols*, **6**(8), 1105–1120.

Power, Maxwell E. 1943. The brain of Drosophila melanogaster. *Journal of Morphology*, **72**(3), 517–559.

Price, Jeffrey L, Blau, Justin, Rothenfluh, Adrian, Abodeely, Marla, Kloss, Brian, & Young, Michael W. 1998. double-time Is a Novel Drosophila Clock Gene that Regulates {PERIOD} Protein Accumulation. *Cell*, **94**(1), 83 – 95.

Priori, Silvia G., Pandit, Sandeep V., Rivolta, Ilaria, Berenfeld, Omer, Ronchetti, Elena, Dhamoon, Amit, Napolitano, Carlo, Anumonwo, Justus, di Barletta, Marina Raffaele, Gudapakkam, Smitha, Bosi, Giuliano, Stramba-Badiale, Marco, & Jalife, JosÃĺ. 2005. A Novel Form of Short QT Syndrome (SQT3) Is Caused by a Mutation in the KCNJ2 Gene. *Circulation Research*, **96**(7), 800–807.

Quiroga, R.Q., & Panzeri, S. 2013. *Principles of Neural Coding*. Taylor & Francis.

Ramskold, Daniel, Luo, Shujun, Wang, Yu-Chieh, Li, Robin, Deng, Qiaolin, Faridani, Omid R, Daniels, Gregory A, Khrebtukova, Irina, Loring, Jeanne F, Laurent, Louise C, Schroth, Gary P, & Sandberg, Rickard. 2012. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotech*, **30**(8), 777–782.

Raviola, Elio, & Mazzarello, Paolo. 2011. The diffuse nervous network of Camillo Golgi: Facts and fiction. *Brain Research Reviews*, **66**(1âĂŞ2), 75–82.

Reiner, Anton, Veenman, C.Leonardus, Medina, Loreta, Jiao, Yun, Mar, Nobel Del, & Honig, Marcia G. 2000. Pathway tracing using biotinylated dextran amines. *Journal of Neuroscience Methods*, **103**(1), 23 – 37.

Rinaman, L, Card, JP, & Enquist, LW. 1993. Spatiotemporal responses of astrocytes, ramified microglia, and brain macrophages to central neuronal infection with pseudorabies virus. *The Journal of Neuroscience*, **13**(2), 685–702.

Robinson, Mark D., & Smyth, Gordon K. 2007. Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics*, **23**(21), 2881–2887.

Robinson, Mark D., & Smyth, Gordon K. 2008. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics*, **9**(2), 321–332.

Robinson, Mark D., McCarthy, Davis J., & Smyth, Gordon K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**(1), 139–140.

Rocca, Julius. 1998. Galen and Greek Neuroscience (Notes Towards a Preliminary Survey). *Early Science and Medicine*, **3**(3), 216–240.

Rosenmund, Christian, Sigler, Albrecht, Augustin, Iris, Reim, Kerstin, Brose, Nils, & Rhee, Jeong-Seop. 2002. Differential Control of Vesicle Priming and Short-Term Plasticity by Munc13 Isoforms. *Neuron*, **33**(3), 411 – 424.

Roux, Sylvie, Colasante, Cesare, Cloment, CÃľcile Saint, Barbier, Julien, Curie, Thomas, Girard, Emmanuelle, MolgÃş, Jordi, & BrÃżlet, Philippe. 2005. Internalization of a GFP-tetanus toxin C-terminal fragment fusion protein at mature mouse neuromuscular junctions. *Molecular and Cellular Neuroscience*, **30**(1), 79 – 89.

Royer, Sebastien, Coulson, Rosalind L., & Klein, Marc. 2000. Switching Off and On of Synaptic Sites atAplysia Sensorimotor Synapses. *The Journal of Neuroscience*, **20**(2), 626–638.

Rubin, GM, & Spradling, AC. 1982. Genetic transformation of Drosophila with transposable element vectors. *Science*, **218**(4570), 348–353.

Sachse, Silke, Rueckert, Erroll, Keller, Andreas, Okada, Ryuichi, Tanaka, Nobuaki K., Ito, Kei, & Vosshall, Leslie B. 2007. Activity-Dependent Plasticity in an Olfactory Circuit. *Neuron*, **56**(5), 838 – 850.

Samuels, Benjamin Adam, Hsueh, Yi-Ping, Shu, Tianzhi, Liang, Haoya, Tseng, Huang-Chun, Hong, Chen-Jei, Su, Susan C., Volker, Janet, Neve, Rachael L., Yue, David T., & Tsai, Li-Huei. 2007. Cdk5 Promotes Synaptogenesis by Regulating the Subcellular Distribution of the {MAGUK} Family Member {CASK}. *Neuron*, **56**(5), 823 – 837.

Sandrock, Alfred W., Dryer, Stuart E., Rosen, Kenneth M., Gozani, Shai N., Kramer, Rainer, Theill, Lars E., & Fischbach, Gerald D. 1997. Maintenance of Acetylcholine Receptor Number by Neuregulins at the Neuromuscular Junction in Vivo. *Science*, **276**(5312), 599–603.

Sasagawa, Yohei, Nikaido, Itoshi, Hayashi, Tetsutaro, Danno, Hiroki, Uno, Kenichiro, Imai, Takeshi, & Ueda, Hiroki. 2013. Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome Biology*, **14**(4), R31.

Schiavo, Giampietro G., Benfenati, Fabio, Poulain, Bernard, Rossetto, Ornella, de Laureto, Patrizia Polverino, DasGupta, Bibhuti R., & Montecucco, Cesare. 1992. Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic cleavage of synaptobrevin. *Nature*, **359**(6398), 832–835.

Schulz, David J, Goaillard, Jean-Marc, & Marder, Eve. 2006. Variable channel expression in identified single and electrically coupled neurons in different animals. *Nat Neurosci*, **9**(3), 356–362.

Scott, Kristin, Jr., Roscoe Brady, Cravchik, Anibal, Morozov, Pavel, Rzhetsky, Andrey, Zuker, Charles, & Axel, Richard. 2001. A Chemosensory Gene Family Encoding Candidate Gustatory and Olfactory Receptors in Drosophila. *Cell*, **104**(5), 661 – 673.

Sen, Anindya, Shetty, Chetak, Jhaveri, Dhanisha, & Rodrigues, Veronica. 2005. Distinct types of glial cells populate the Drosophila antenna. *BMC Developmental Biology*, **5**(1), 25.

Senthilan, Pingkalai R., Piepenbrock, David, Ovezmyradov, Guvanch, Nadrowski, BjÃűrn, Bechstedt, Susanne, Pauls, Stephanie, Winkler, Margret, MÃűbius, Wiebke, Howard, Jonathon, & GÃűpfert, Martin C. 2012. Drosophila Auditory Organ Genes and Genetic Hearing Defects. *Cell*, **150**(5), 1042 – 1054.

Shang, Yuhua, Claridge-Chang, Adam, Sjulson, Lucas, Pypaert, Marc, & MiesenbÃűck, Gero. 2007. Excitatory Local Circuits and Their Implications for Olfactory Processing in the Fly Antennal Lobe. *Cell*, **128**(3), 601 – 612.

Sheng, Morgan. 2001. Molecular organization of the postsynaptic specialization. *Proceedings of the National Academy of Sciences*, **98**(13), 7058–7061.

Sherrington, C.S. 1897. *The central nervous system. In A Textbook of Physiology.* London: Macmillian.

Stocker, R.F., Lienhard, M.C., Borst, A., & Fischbach, K.F. 1990. Neuronal architecture of the antennal lobe in Drosophila melanogaster. *Cell and Tissue Research*, **262**(1), 9–34.

Stowers, R. Steven. 2011. An efficient method for recombineering GAL4 and QF drivers. *fly*, **5**(1933-6934), 371–378.

Su, Chih-Ying, Menuz, Karen, Reisert, Johannes, & Carlson, John R. 2012. Non-synaptic inhibition between grouped neurons in an olfactory circuit. *Nature*, **492**(7427), 66–71.

Suh, Greg S. B., Wong, Allan M., Hergarden, Anne C., Wang, Jing W., Simon, Anne F., Benzer, Seymour, Axel, Richard, & Anderson, David J. 2004. A single population of olfactory sensory neurons mediates an innate avoidance behaviour in Drosophila. *Nature*, **431**(7010), 854–859.

Tabuchi, Katsuhiko, Sawamoto, Kazunobu, Suzuki, Emiko, Ozaki, Koichi, Sone, Masaki, Hama, Chihiro, Tanifuji-Morimoto, Takako, Yuasa, Yoshihiro, Yoshihara, Yoshihiro, Nose, Akinao, & Okano, Hideyuki. 2000. GAL4/UAS-WGA system as a powerful tool for tracing Drosophila transsynaptic neural pathways. *Journal of Neuroscience Research*, **59**(1), 94–99.

Takemura, Shin-ya, Bharioke, Arjun, Lu, Zhiyuan, Nern, Aljoscha, Vitaladevuni, Shiv, Rivlin, Patricia K., Katz, William T., Olbris, Donald J., Plaza, Stephen M., Winston, Philip, Zhao, Ting, Horne, Jane Anne, Fetter, Richard D., Takemura, Satoko, Blazek, Katerina, Chang, Lei-Ann, Ogundeyi, Omotara, Saunders, Mathew A., Shapiro, Victor, Sigmund, Christopher, Rubin, Gerald M., Scheffer, Louis K., Meinertzhagen, Ian A., & Chklovskii, Dmitri B. 2013. A visual motion detection circuit suggested by Drosophila connectomics. *Nature*, **500**(7461), 175–181.

Tea, Joy S., Chihara, Takahiro, & Luo, Liqun. 2010. Histone Deacetylase Rpd3 Regulates Olfactory Projection Neuron Dendrite Targeting via the Transcription Factor Prospero. *The Journal of Neuroscience*, **30**(29), 9939–9946.

Thomas, Amanda, Lee, Pei-Jung, Dalton, Justin E., Nomie, Krystle J., Stoica, Loredana, Costa-Mattioli, Mauro, Chang, Peter, Nuzhdin, Sergey, Arbeitman, Michelle N., & Dierick, Herman A. 2012. A Versatile Method for Cell-Specific Profiling of Translated mRNAs in <italic>Drosophila</italic>. *PLoS ONE*, **7**(7), e40276.

Tomishima, Mark J., & Enquist, Lynn W. 2002. In Vivo Egress of an Alphaherpesvirus from Axons. *Journal of Virology*, **76**(16), 8310–8317.

Tower, John. 2011. Heat shock proteins and Drosophila aging. *Experimental Gerontology*, **46**(5), 355 – 362. Aging studies in Drosophila melanogaster.

Trapnell, Cole, Roberts, Adam, Goff, Loyal, Pertea, Geo, Kim, Daehwan, Kelley, David R, Pimentel, Harold, Salzberg, Steven L, Rinn, John L, & Pachter, Lior. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protocols*, **7**(3), 562–578.

Trapnell, Cole, Hendrickson, David G, Sauvageau, Martin, Goff, Loyal, Rinn, John L, & Pachter, Lior. 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotech*, **31**(1), 46–53.

Trojanowski, John Q., Gonatas, Jacqueline O., & Gonatas, Nicholas K. 1982. Horseradish peroxidase (HRP) conjugates of cholera toxin and lectins are more sensitive retrogradely transported markers than free {HRP}. *Brain Research*, **231**(1), 33 – 50.

Tsurudome, Kazuya, Tsang, Karen, Liao, Edward H., Ball, Robin, Penney, Jay, Yang, Jr-Shiuan, Elazzouzi, Fatima, He, Tao, Chishti, Athar, Lnenicka, Greg, Lai, Eric C., & Haghighi, A. Pejmun. 2010. The Drosophila miR-310 Cluster Negatively Regulates Synaptic Strength at the Neuromuscular Junction. *Neuron*, **68**(5), 879 – 893.

Turrigiano, Gina G., & Nelson, Sacha B. 2004. Homeostatic plasticity in the developing nervous system. *Nat Rev Neurosci*, **5**(2), 97–107.

Turrigiano, Gina G., Leslie, Kenneth R., Desai, Niraj S., Rutherford, Lana C., & Nelson, Sacha B. 1998. Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature*, **391**(6670), 892–896.

Ueda, Hiroki R., Matsumoto, Akira, Kawamura, Miho, Iino, Masamitsu, Tanimura, Teiichi, & Hashimoto, Seiichi. 2002. Genome-wide Transcriptional Orchestration of Circadian Rhythms inDrosophila ,. *Journal of Biological Chemistry*, **277**(16), 14048–14052.

Ugolini, Gabriella. 2011. Chapter 10 - Rabies Virus as a Transneuronal Tracer of Neuronal Connections. *Pages 165 – 202 of:* Jackson, Alan C. (ed), *Research Advances in Rabies*. Advances in Virus Research, vol. 79. Academic Press.

Vandesompele, Jo, De Preter, Katleen, Pattyn, Filip, Poppe, Bruce, Van Roy, Nadine, De Paepe, Anne, & Speleman, Frank. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, **3**(7), research0034.1–research0034.11.

Velculescu, Victor E., Zhang, Lin, Vogelstein, Bert, & Kinzler, Kenneth W. 1995. Serial Analysis of Gene Expression. *Science*, **270**(5235), 484–487.

Venken, Koen J. T., He, Yuchun, Hoskins, Roger A., & Bellen, Hugo J. 2006. P[acman]: A BAC Transgenic Platform for Targeted Insertion of Large DNA Fragments in D. melanogaster. *Science*, **314**(5806), 1747–1751.

Venken, Koen J T, Carlson, Joseph W, Schulze, Karen L, Pan, Hongling, He, Yuchun, Spokony, Rebecca, Wan, Kenneth H, Koriabine, Maxim, de Jong, Pieter J, White, Kevin P, Bellen, Hugo J, & Hoskins, Roger A. 2009. Versatile P[acman] BAC libraries for transgenesis studies in Drosophila melanogaster. *Nat Meth*, **6**(6), 431–434.

von Trotha, Jakob, Egger, Boris, & Brand, Andrea. 2009. Cell proliferation in the Drosophila adult brain revealed by clonal analysis and bromodeoxyuridine labelling. *Neural Development*, **4**(1), 9.

Vosshall, Leslie B. 2000. Olfaction in Drosophila. *Current Opinion in Neurobiology*, **10**(4), 498 – 503.

Vosshall, Leslie B., & Stocker, Reinhard F. 2007. Molecular Architecture of Smell and Taste in Drosophila. *Annual Review of Neuroscience*, **30**(1), 505–533. PMID: 17506643.

Vosshall, Leslie B, Amrein, Hubert, Morozov, Pavel S, Rzhetsky, Andrey, & Axel, Richard. 1999. A Spatial Map of Olfactory Receptor Expression in the Drosophila Antenna. *Cell*, **96**(5), 725 – 736.

Wang, Zhong, Gerstein, Mark, & Snyder, Michael. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*, **10**(1), 57–63.

Weiner, J. 1999. *Time, Love, Memory: A Great Biologist and His Quest for the Origins of Behavior*. Vintage Series. Vintage Books.

Weiss, Linnea A., Dahanukar, Anupama, Kwon, Jae Young, Banerjee, Diya, & Carlson, John R. 2011. The Molecular and Cellular Basis of Bitter Taste in Drosophila. *Neuron*, **69**(2), 258 – 272.

Weyhersmuller, Annika, Hallermann, Stefan, Wagner, Nicole, & Eilers, Jens. 2011. Rapid Active Zone Remodeling during Synaptic Plasticity. *The Journal of Neuroscience*, **31**(16), 6041–6052.

White, J. G., Southgate, E., Thomson, J. N., & Brenner, S. 1986. The Structure of the Nervous System of the Nematode Caenorhabditis elegans. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, **314**(1165), 1–340.

Wickersham, Ian R., Lyon, David C., Barnard, Richard J.O., Mori, Takuma, Finke, Stefan, Conzelmann, Karl-Klaus, Young, John A.T., & Callaway, Edward M. 2007. Monosynaptic Restriction of Transsynaptic Tracing from Single, Genetically Targeted Neurons. *Neuron*, **53**(5), 639 – 647.

Wiesel, T.N., Hubel, D.H., & Lam, D.M.K. 1974. Autoradiographic demonstration of ocular-dominance columns in the monkey striate cortex by means of transneuronal transport. *Brain Research*, **79**(2), 273 – 279.

Wijnen, Herman, Naef, Felix, Boothroyd, Catharine, Claridge-Chang, Adam, & Young, Michael W. 2006. Control of Daily Transcript Oscillations in <italic>Drosophila</italic> by Light and the Circadian Clock. *PLoS Genet*, **2**(3), e39.

Wilson, Rachel I. 2013. Early Olfactory Processing in Drosophila: Mechanisms and Principles. *Annual Review of Neuroscience*, **36**(1), 217–241. PMID: 23841839.

Wilson, Rachel I., & Laurent, Gilles. 2005. Role of GABAergic Inhibition in Shaping Odor-Evoked Spatiotemporal Patterns in the Drosophila Antennal Lobe. *The Journal of Neuroscience*, **25**(40), 9069–9079.

Wilson, Rachel I., Turner, Glenn C., & Laurent, Gilles. 2004. Transformation of Olfactory Representations in the Drosophila Antennal Lobe. *Science*, **303**(5656), 366–370.

Wong, Allan M, Wang, Jing W, & Axel, Richard. 2002. Spatial Representation of the Glomerular Map in the Drosophila Protocerebrum. *Cell*, **109**(2), 229 – 241.

Xie, Lulu, Kang, Hongyi, Xu, Qiwu, Chen, Michael J., Liao, Yonghong, Thiyagarajan, Meenakshisundaram, O'Donnell, John, Christensen, Daniel J., Nicholson, Charles, Iliff, Jeffrey J., Takano, Takahiro, Deane, Rashid, & Nedergaard, Maiken. 2013. Sleep Drives Metabolite Clearance from the Adult Brain. *Science*, **342**(6156), 373–377.

Xu, Wei, & Sudhof, Thomas C. 2013. A Neural Circuit for Memory Specificity and Generalization. *Science*, **339**(6125), 1290–1295.

Yaksi, Emre, & Wilson, Rachel I. 2010. Electrical Coupling between Olfactory Glomeruli. *Neuron*, **67**(6), 1034 – 1047.

Yoon, Byung C., Zivraj, Krishna H., & Holt, Christine E. 2009. Local Translation and mRNA Trafficking in Axon Pathfinding. *Pages 108–138 of:* Koenig, Edward (ed), *Cell Biology of the Axon*. Results and Problems in Cell Differentiation, vol. 48. Springer Berlin Heidelberg.

Younger, Meg A., Muller, Martin, Tong, Amy, Pym, Edward C., & Davis, Graeme W. 2013. A Presynaptic {ENaC} Channel Drives Homeostatic Plasticity. *Neuron*, **79**(6), 1183 – 1196.

Zampieri, Niccolo, Jessell, Thomas M., & Murray, Andrew J. 2014. Mapping Sensory Circuits by Anterograde Transsynaptic Transfer of Recombinant Rabies Virus. *Neuron*, **81**(4), 766 – 778.

Zhang, Xiujun, Hao, Lili, Meng, Lijun, Liu, Meiling, Zhao, Lina, Hu, Fen, Ding, Cunbao, Wang, Yang, He, Baoling, Pan, Yuxin, Fang, Wei, Chen, Jing, Hu, Songnian, & Jia, Mengchun. 2013. Digital Gene Expression Tag Profiling Analysis of the Gene Expression Patterns Regulating the Early Stage of Mouse Spermatogenesis. *PLoS ONE*, **8**(3), e58680.

Zhao, CongJian, Dreosti, Elena, & Lagnado, Leon. 2011. Homeostatic Synaptic Plasticity through Changes in Presynaptic Calcium Influx. *The Journal of Neuroscience*, **31**(20), 7492–7496.

Zhou, Shanshan, Stone, Eric A., Mackay, Trudy F. C., & Anholt, Robert R. H. 2009. Plasticity of the Chemoreceptor Repertoire in <italic>Drosophila melanogaster</italic>. *PLoS Genet*, **5**(10), e1000681.

Zhou, Shanshan, Campbell, Terry G., Stone, Eric A., Mackay, Trudy F. C., & Anholt, Robert R. H. 2012. Phenotypic Plasticity of the Drosophila Transcriptome. *PLoS Genet*, **8**(3), e1002593.