Modulation of Reproductive Behaviours by Neuropeptides in *Drosophila*



A thesis submitted for the degree of *Doctor of Philosophy*

by

Sebastian Elliott Birtles

University College

University of Oxford

Hilary 2022

Abstract

All animals must adapt their behaviour according to their environment and intrinsic state. At the level of the nervous system, this demands flexibility in the output of neural circuits. Neuropeptides, the largest class of signalling molecule in the nervous system, are crucial for engendering this flexibility, signalling through G-protein coupled receptors to modulate neuronal and synaptic properties. This study investigated the role of neuropeptides in the modulation of reproductive behaviours in *Drosophila melanogaster*. These behaviours are 'hard-wired' into the nervous system by the action of developmental genes yet display considerable flexibility in response to environmental and physiological factors. Though several neuropeptides are thought to modulate reproductive behaviour, the underlying mechanisms remain poorly understood. Knowledge of where neuropeptides act represents a particularly significant gap in our understanding. Thus, a key objective of this study was to elucidate the distribution of neuropeptide receptors (NPRs). Intersectional genetic approaches were used to target and label cells expressing *NPRs,* revealing broad expression across distinct compartments of the nervous system. Crucially, expression of several *NPRs* was evidenced in the neural circuits governing reproductive behaviour, facilitating direct modulation of these circuits. Analysis of single-cell RNA sequencing data corroborated these findings and provided additional insights, for example: revealing low expression-levels of *NPRs,* expression in transcriptionally distinct cell types, and co-expression in single cells.

Subsequent investigations focused on SIFaR, the receptor for SIFamide: a neuropeptide thought to modulate reproduction, feeding and sleep. Newly generated *SIFaR* mutant alleles produced abnormal reproductive behaviours, supporting a role for SIFaR signalling in their regulation. RNA interference was used to disrupt SIFaR exclusively in the neural circuits underlying these behaviours. This, too, produced robust phenotypes, suggesting that SIFamide modulates reproductive behaviour via direct signalling through these neurons. To study downstream effects of SIFaR signalling, FRET and calcium imaging experiments were designed and piloted, and an optimal paradigm for further investigation was established.

A range of approaches were employed to investigate how SIFamide performs its integrative functions. Immunohistochemical investigation, and analysis of hemibrain connectome data, supported the existence of just 4 SIFamide-producing neurons in the brain, with the latter providing evidence of heterogeneity in their morphologies, uncovering a potential mechanism for the modulation of discrete behaviours by SIFamide. Finally, approaches to assess the function of SIFamide in specific physiological contexts were tested. This included generation of a 'FlpSTOP' allele, facilitating temporally and spatially restricted disruption of *SIFaR*, and the design and validation of experiments to assay reproductive behaviours in the context of starvation.

Collectively, this data elucidates the expression profile of key neuropeptide receptors in *Drosophila* and, by establishing novel tools and approaches, paves the way for future investigation into neuropeptide signalling mechanisms.

Table of Contents

Abstract **2**

Table of Contents **3**

List of Accompanying Materials **9**

Acknowledgements **10**

Author's Declaration **11**

Abbreviations **12**

Chapter 1: Introduction **14**

1.1 Genes, neural circuits, and behaviour 15

1.1.1 The nervous system and behaviour 15

1.1.2 Neural circuits 16

1.1.3 Genes and behaviour 17

1.2 Modulation of neural circuits and behaviour 18

1.2.1 Neuromodulation 18

1.2.2 Neuropeptides 19

1.2.2.1 What are neuropeptides? 19

1.2.2.2 Neuropeptide synthesis & processing 19

1.2.2.3 Neuropeptide release & signalling 20

1.3 Neurogenetics in *Drosophila* 21

1.3.1 Prevalent genetic tools in *Drosophila* 22

1.3.2 The *Drosophila* nervous system 23

1.3.2.1 Neuropeptides in the *Drosophila* nervous system 25

1.3.3 Reproductive behaviour in *Drosophila* 26
1.3.3.1 Male reproductive behaviour 26
1.3.3.2 Female reproductive behaviour 27
1.3.3.3 Genetic and neural correlates of *Drosophila* reproductive behaviours 28
1.3.3.4 Modulation of reproductive behaviour by neuropeptides 29
1.3.3.4.1 Myoinhibitory peptide (MIP) and sex peptide (SP) 30
1.3.3.4.2 SIFamide 31
1.3.3.4.3 Corazonin 32
1.3.3.4.4 Neuropeptide F 33
1.3.3.4.5 Tachykinin 34
1.4 Study objectives 35
Chapter 2: Materials and Methods 36
2.1 *Drosophila* husbandry & behavioural assays 37
2.1.1 *Drosophila* stocks 37
2.1.2 Behavioural assays 37
2.1.3 Automated behaviour tracking 38
2.2 Generation of novel *Drosophila* lines 40
2.2.1 *SPR-GAL4* and *SPR-p65AD* lines 40
2.2.2 *UAS-SIFaR* line 40
2.2.3 $SIFaR^{deletion}$ allele 42
2.3 Western blotting 42
2.3.1 Protein extraction 42
2.3.2 SDS-PAGE and electrophoretic transfer 42
2.3.3 Immunolabelling and detection 43
2.4 RT-qPCR 43

2.4.1 RNA extraction & cDNA synthesis 43

2.4.2 qPCR 44

2.5 Immunohistochemistry 45

2.5.1 Tissue dissection and fixation 45

2.5.2 Antibody staining 45

2.6 Fluorescence imaging 46

2.6.1 Confocal microscopy 46

2.6.2 FRET imaging 46

2.6.3 *In-vivo* 2-photon calcium imaging 47

2.6.4 *Ex-vivo* wide-field calcium imaging 48

Chapter 3: Investigating the Distribution of Neuropeptide Receptors 49

3.1 Introduction 50

3.2 Transgenic approaches to map neuropeptide receptor expression 50

3.2.1 Generation of *SPR-GAL4* and *SPR-p65AD* lines by CRISPR-Cas9 51

3.2.2 Generation of *split-GAL4* and *LexA* lines using the Trojan-MiMIC system 52

3.2.3 Applying transgenic tools to map neuropeptide receptor distribution 55

3.2.3.1 *SPR-GAL4*$^{[CRISPR]}$ and *SPR-p65AD*$^{[CRISPR]}$ expression in the CNS 55

3.2.3.2 *SPR-GAL4*$^{[CRISPR]}$ expression in peripheral tissues 57

3.2.3.3 Expression of neuropeptide receptor-MiMIC lines in the nervous system 60

3.3 Mapping neuropeptide receptor expression using single-cell RNA sequencing 63

3.3.1 Quantifying neuropeptide receptor expression in the CNS 64

3.3.2 Expression of neuropeptide receptors across cell types in the CNS 68

3.3.3 Expression of neuropeptide receptors in specific cell types 74

3.3.3.1 *NPR* expression in fast-acting neurotransmitter neuronal subtypes 74

3.3.3.2 *SPR* expression in Glia 74

3.4 Mapping neuropeptide receptor protein distribution 75

3.4.1 Generation of a novel antibody to investigate SIFaR distribution 76

3.4.1.1 SIFaR structure analysis and epitope selection 76

3.4.1.2 Antibody production 77

3.4.1.3 Antibody validation in immunohistochemical assays 78

3.4.1.4 Antibody validation in western blot assays 81

3.5 Summary and discussion 84

Chapter 4: Expression of Neuropeptide Receptors in the Neural Circuits of Reproductive Behaviour 86

4.1 Introduction 87

4.2.1 *SIFaR* 88

4.2.1.1 *SIFaR* expression in *dsx* cells. 88

4.2.1.2 *SIFaR* expression in *fru* cells. 93

4.2.2 *SPR* expression in *dsx* cells 97

4.2.3 *CrzR*, and *NPFR* expression in *dsx* cells 101

4.3 Summary and discussion 104

Chapter 5: Modulation of Reproductive Behaviour by SIFamide 107

5.1 Introduction 108

5.2 Reproductive behaviour in *SIFaR* mutants 108

5.2.1 Generation and characterisation of a *SIFaR* deletion allele 109

5.2.2 *SIFaR-p65AD*$^{[MiMIC]}$ as a *SIFaR* mutant allele 115

5.2.2.1 Reproductive behaviour phenotypes in *SIFaR-p65AD*$^{[MiMIC]}$ homozygous males 115

5.2.2.2 Reproductive behaviour phenotypes in SIFaR-p65AD[MiMIC] homozygous females 120

5.3 Investigating the role of SIFaR signalling in *dsx* neurons..... 121

5.3.1 Phenotypic screen for selection of a suitable RNAi construct..... 122

5.3.2 *SIFaR* knockdown in *dsx* neurons by RNAi affects reproductive behaviours 123

5.4 Generation of a *SIFaR*-FlpSTOP allele 124

5.5 Quantifying SIFaR loss-of-function alleles 128

5.6 Visualising SIFaR signalling in *dsx* circuits 129

5.6.1 Introduction 129

5.6.2 Pilot experiments to measure cAMP modulation via EPAC-FRET 130

5.6.3 Calcium imaging to measure the effect of SIFamide on *SIFaR dsx* neuronal activity 135

5.6.3.1 Pilot *in-vivo* 2-photon calcium imaging experiments..... 135

5.6.3.2 Pilot *ex-vivo* wide-field calcium imaging experiments..... 138

5.7 Summary and discussion 141

Chapter 6: Elucidating SIFamide Signalling Mechanisms and Functions in the Adult Nervous System 143

6.1 Introduction..... 144

6.2 Investigating SIFamide neuron morphology and connectivity. 146

6.2.1 SIFamide expression in the CNS 147

6.2.2 Exploring heterogeneity in SIFamide neuron morphology and connectivity 150

6.3 Investigating the functions of SIFamide signalling in starvation 153

6.3.1 Does SIFamide promote appetitive feeding behaviour in males?..... 154

6.3.2 Does starvation affect courtship behaviour? 157

6.3.3 Follow up experiments to investigate modulation of sexual behaviour by SIFamide during starvation 158

6.4 Summary and discussion 160

Chapter 7: Author's Summary 164

References 172

Accompanying Materials 186

List of Accompanying Materials

Experiments and research conducted alongside this work contributed to the following publication:

Allen, A., Neville, M., Birtles, S., Croset, V., Treiber, C., Waddell, S., Goodwin, S. (2020) 'A single-cell transcriptomic atlas of the adult *Drosophila* ventral nerve cord'. *eLife*, 9, 1–32.

This publication can be found under 'Accompanying Materials' (p.186).

Acknowledgements

To my supervisor, Prof. Stephen Goodwin, thank you for your guidance and encouragement throughout. I always felt that your door was open to me, and that I had your support and confidence. With that knowledge, I had the freedom to express my ideas and pursue my interests. I can't overstate the difference that made to me. It was an honour to carry out my PhD in your lab; I will treasure the experience for life. I extend my deepest gratitude to members of the Goodwin lab: Megan Neville, for your advice and support on practically all aspects of my project, and for your unfaltering patience and kindness; Annika Rings, particularly, for your help with calcium imaging experiments: while the 'destination' may not have been what we hoped, I thoroughly enjoyed the journey; Aaron Allen, for being an oracle on so many subjects, not least during my forays into western blots, qPCR and single-cell transcriptomics. Despite my frequent queries (necessitating the removal of multiple layers of sound proofing on each occasion), not once did you give the impression of not having the time to help. For you, perhaps, this was a curse, but I was tremendously grateful for it; Deniz Erezyilmaz, for your thought-provoking ideas on my project, and for always being so forthcoming with support and advice; and Tetsuya Nojima, for being my inspiration on all things immunohistochemistry related, and for being so generous with your expertise and time.

I also thank past members of the Goodwin lab: Matthew Herbert, for his wise words in navigating a PhD, and for always being up for a beer (or four), ostensibly to discuss this serious business; Ammerins de Haan, for her help and practical tips in the wet-lab, without which I would have wasted considerable time; and Carolina Rezával, for getting me from *Drosophila* 'virgin' to competent identifier of the *Drosophila* virgin.

I thank Prof. Scott Waddell, for his engagement with my work at both Transfer and Confirmation of Status, and his constructive feedback on both occasions, and Prof. Manuela Zaccolo, for her useful contributions as Transfer of Status assessor.

I thank my *Viva* examiners, Prof. Clive Wilson and Prof. Christian Wegener, who not only read my thesis but gave it careful consideration and, through their thoughtful questions and comments, encouraged and inspired me to do the same.

I am immeasurably grateful for the love and support of all my family and friends, without which I could not have completed this work and retained my sanity. Particular thanks go to my mum, Jane Parker, for her support on every front, especially during the low points of my write-up. And Antonia Skinner, for her continued encouragement, and her interesting suggestions as to what the subject of my PhD should be.

Finally, I thank the subjects of my experiments, for their collective, and involuntary, sacrifice. In studying their brains, inevitably I pondered over the extent of their consciousness. Reflecting on the nature of some experiments described in this thesis, I hope it is limited…

Author's Declaration

The work presented in this thesis is my own, except where otherwise indicated. This work is being submitted for the degree of *Doctor of Philosophy* at the University of Oxford. None of this work has been submitted previously for any other degree at this University, nor for a qualification at any other institution.

All experiments were performed in The Centre for Neural circuits and Behaviour, University of Oxford, with the following exceptions:

i) Wide-field FRET imaging was carried out in the laboratory of Manuela Zaccolo in the department of Physiology, Anatomy and Genetics, Oxford.

ii) Confocal FRET imaging was carried out in the Micron Bioimaging Facility, Department of Biochemistry, Oxford.

Abbreviations

7-TM – 7-transmembrane

Abg – Abdominal ganglion

Ach – Acetylcholine

AD – Activation domain

AHLS – Adult haemolymph-like saline

AL – Antennal lobe

AZ – Active zone

cAMP – Cyclic adenosine monophosphate

CBF – Cell body fiber

cDNA – Complimentary DNA

CDS – Coding sequence

CFP – Cyan fluorescent protein

CNS – Central nervous system

CrzR – Corazonin receptor

CS – Canton-S

DA – Dopamine

DBD – DNA binding domain

DFSB – Dorsal fan-shaped body

DMSO – Dimethyl sulfoxide

DNA – Deoxyribonucleic acid

DSB – Double-strand break

dSTIM – Drosophila stromal interaction molecule

Dsx/*dsx* – Doublesex/*doublesex*

EPAC – Exchange protein directly activated by cAMP

FAN – Fast-acting neurotransmitter

FRET – Förster resonance energy transfer

FRT – Flippase recognition target

Fru/*fru* – Fruitless/*fruitless*

Fsk - Forskolin

GABA – Gamma aminobutyric acid

GECI – Genetically encoded calcium indicator

GFP – Green fluorescent protein

GOI – Gene of interest

GPCR – G-protein coupled receptor

gRNA – Guide RNA

HDR – Homology-directed repair

HEK – Human embryonic kidney

IBMX – 3-isobutyl-1-methylxanthine

IHC - Immunohistochemistry

KDa - Kilodalton

LDCV – Large dense-core vesicle

LH – Lateral horn

LOF – Loss-of-function

MB – Mushroom body

mGFP – membrane-localising GFP

MiMIC – Minos mediated integration cassette

nGFP – Nucleus-localising GFP

NHEJ – Non-homologous end-joining

NPFR – Neuropeptide F receptor

NPR – Neuropeptide receptor

OA - Octopamine

PCR – Polymerase chain reaction

PMR – Post-mating response

PNS – Peripheral nervous system

qPCR – Quantitative PCR

RNAi – RNA interference

ROI – Region of interest

SA – Splice acceptor

scRNAseq – Single-cell RNA sequencing

SCV – Small clear vesicle

SIFaR – SIFamide receptor

SLP – Superior lateral protocerebrum

SMP – Superior medial protocerebrum

SPR – Sex peptide receptor

TE – Transposable element

TF – Transcription factor

TkR86C – Tachykinin receptor (spliceform 86C)

Tra - Transformer

tSNE - t-distributed stochastic neighbor embedding

UAS – Upstream activation sequence

UTR – Un-translated region

VNC – Ventral nerve cord

WB – Western blot

YFP – Yellow fluorescent protein

Chapter 1: Introduction

1.1 Genes, neural circuits, and behaviour

1.1.1 The nervous system and behaviour

Behaviour can be defined as the internally coordinated responses of whole living organisms to internal and/or external stimuli (Levitis et al., 2009). That behaviour originates from the nervous system was established by the ancient Greeks (Toker et al., 2003), yet only relatively recently have casual links between the nervous system and behaviour started to be established. In the late 18th century, Luciani Galvani's experiments revealed that the nervous system operates using intrinsically generated electricity (Bennett, 1999). Aided by advances in microscopy and histology in the following century, the individual cells that make up the nervous system: neurons, were identified. From this emerged Santiago Ramón y Cajal's neuron doctrine: that neurons function as individual units, each capable of receiving signals at specialised input regions, generating electrical impulses to encode these signals, and propagating these impulses to spatially segregated output sites (Finger, 2005). Cajal's iconic drawings illustrate the distinct functional compartments of the neuron: the cell body, or soma, which contains the nucleus and specialised organelles; the widely arborising input regions - dendrites, the elongated processes along which electrical impulses are propagated - axons, and the axon terminals, where contact is made with downstream neurons. The electrical impulses generated and propagated by neurons arise from the movement of charged ions, principally sodium (Na^+), and potassium (K^+), across the neuronal membrane. Using the giant squid axon, Hodgkin, Huxley and Katz described, with mathematical precision, how discrete nerve impulses, or action potentials, are initiated and propagated by the movement of these ions (Hodgkin & Huxley, 1952). It is in the temporal sequence of action potentials generated by neurons that information is encoded in the nervous system.

Following heated debate for many years over the mechanism of signal transmission at synapses; the gaps between adjacent neurons, it became clear that in most cases signalling between neurons is mediated by the release of chemicals (Valenstein, 2005). The chemicals

involved in fast synaptic transmission - fast acting neurotransmitters (FANs), are packaged into spherical vesicles at the axon terminal. Action potentials arriving at the terminal produce an influx of calcium (Ca^{2+}), via voltage-gated calcium channels, triggering exocytosis: the secretion of the contents of vesicles into the synapse. After diffusing the short (~30nm) distance from pre-synaptic to post-synaptic neuron, FANs bind cognate receptors. Binding of the receptor changes the ionic conductance of the post-synaptic neuron and, thus, the chemical signal is converted back into an electrical one. Some FAN/receptor interactions are excitatory: they depolarise the post synaptic neuron to increase action potential firing probability, while others are inhibitory; hyperpolarising the membrane via anionic currents to decrease firing probability.

1.1.2 Neural circuits

To segregate incoming information and perform discrete behavioural responses, neurons wire up to form larger functional units known as neural circuits. To produce behaviour, a neural circuit must comprise neurons that:

i) detect stimulus information

ii) process the information and coordinate an appropriate response

iii) execute that response

For example, neural circuits underlying motor responses to external stimuli generally comprise: 1st order sensory neurons to detect and transduce stimulus information, afferent projection neurons to relay signals to the central nervous system (CNS) (though sensory neurons can, in some cases, project directly to the CNS), interneurons, through which signals are transmitted within the CNS, and efferent projection neurons, which relay signals back to the peripheral musculature that will execute the response. Thus, behaviours arise from the coordinated activity of populations of neurons, excitatory and inhibitory, interconnected to form functionally

relevant neural circuits. The nervous system must process and respond to a multitude sensory information. As a result, even in the simplest organisms, neural circuit configurations are invariably far more complex than the sensory-inter-motor neuron loop described above. This is illustrated by the nervous system of the worm *Caenorhabditis elegans* (*C. elegans*), for which a complete wiring diagram, or connectome, was established over 30 years ago (White et al, 1986). Though *C. elegans* has only 302 neurons, these neurons form approximately 7000 chemical synapses, and even neurons encoding very simple information form many connections. For example, part of the 'touch-avoidance' response (Chalfie et al., 1985), in which touch of the tail triggers acceleration, requires just two sets of synapses between 3 neurons: the touch-sensing PLM neuron, an interneuron and a motor neuron. Yet, the PLM sensory neuron makes 31 connections with 11 types of neuron (Bargmann & Marder, 2013). The extent of connectivity in neural circuits has some important implications. First, multiple circuit configurations can give rise to a given behaviour and, conversely, a single circuit can perform a range of operations, depending upon which connections are utilised (Bargmann, 2012). Secondly, circuits performing discrete functions are interconnected and capable of influencing each other and, thus, should not be considered as entirely isolated functional units (Getting, 1989).

1.1.3 Genes and behaviour

That genes control an organism's physical characteristics was established by some of the earliest work in genetics. The link between genes and behaviour remained more enigmatic until Seymour Benzer's work in *Drosophila*, which established that the mutation of discrete genes confers specific behaviour consequences (Greenspan, 2008). Benzer also contributed much to our molecular understanding genes; namely, that they are 'no more or less, the DNA sequence that codes for the amino-acid sequence of a protein' (Gayon, 2016). The complement of proteins expressed by cells dictates almost all aspects of their form and function. In the context of neurons, genes encode the signalling factors that guide synapse

formation and, thus, the establishment of neural circuits; the enzymes that synthesise neurotransmitters; the ion channel and receptor proteins that dictate neuronal and synaptic response properties, and so on. Genes, thus, specify behaviour by determining the configuration of neural circuits, the properties of their constituent neurons and synapses and, consequently, the range of outputs they can produce.

1.2 Modulation of neural circuits and behaviour

1.2.1 Neuromodulation

Genes, thus, lay down the blueprint for nervous systems that, even in the simplest organisms, can detect and encode a multitude of information, and produce a rich repertoire of behavioural responses. However, the constantly changing environments animals inhabit demand a level of behavioural flexibility that even the most complex neural circuit configurations cannot achieve (Getting, 1989). Thus, a diverse network of neuromodulators exists ‘superimposed’ upon the wiring diagram of the nervous system (Bargmann & Marder 2013). Neuromodulators take the form of monoamines, such as dopamine and serotonin, neuropeptides, fatty acids, and gases, such as nitric oxide (Marder et al., 2014). The precise actions of these neuromodulators vary, but the basic modus operandi does not: they modify the properties of neurons and synapses to change the way in which they encode information (Nusbaum et al., 2001). In some cases, they may also alter the configuration of a neural circuit, changing the identity of the neurons that participate in encoding specific information (Bargmann, 2012). In this way, neuromodulators engender flexibility in the output of neural circuits, while maintaining their functional integrity (Marder, 2012).

1.2.2 Neuropeptides

1.2.2.1 What are neuropeptides?

Neuropeptides are the largest and most diverse class of signalling molecule in the nervous system, where they function as key modulators of neural circuits and behaviour. The term neuropeptide can be used to describe any peptide signalling molecule that is released by a neuron and which acts on neural substrates (Burbach, 2011). Peptide signalling molecules were, in fact, originally identified outside the nervous system, as hormones: biologically active substances 'carried from the organ where they are produced to the organ which they affect by means of the bloodstream' (Starling, 1904). The same molecules were then found to be present in neurons. Initially, peptide containing neurons in neurosecretory areas, namely the mammalian hypothalamus, were the focus of investigation. This work led to the discovery of numerous hypothalamic peptides, some of which had previously been associated solely with peripheral functions (Schally et al., 1973; Guillemin, 1978). Peptides were then identified in neurons outside these specialised neurosecretory areas in the brain, suggesting their release not only into the circulation, but within the CNS (Zupanc, 1996). It is now established that neuropeptides are distributed throughout the nervous system, functioning as neurotransmitters, co-transmitters and modulators of neuronal activity (Burbach, 2011). The capacity for neuropeptides to perform these diverse roles, in addition to their hormonal functions, makes them ideally suited to the representation of 'global states' that require adaptation of multiple behavioural and physiological functions.

1.2.2.2 Neuropeptide synthesis & processing

Synthesis of a neuropeptide begins with the transcription of a gene encoding a larger precursor, known as a prepropeptide (Nässel & Zandawala, 2019). The prepropeptide includes a 20-25 amino acid 'signal peptide' at the N-terminus, which is required for entry into the endoplasmic reticulum, after which it is cleaved to yield a propeptide. From here, propeptides

enter the regulated secretory pathway and are packaged into large, dense core vesicles (LDCVs). Additional modification of the propeptide sequence then occurs, which influences the properties of the mature neuropeptide. In this way, a range of peptides with distinct biological activity can be generated from a single prepropeptide gene (Burbach, 2011)

As is the case for all proteins, neuropeptide synthesis occurs at the soma. Thus, neuropeptide containing LDCVs are transported along neuronal processes to reach their release sites, where they may be stored for some time (Zupanc, 1996). FANs, which are packaged in small clear vesicles (SCVs), are stored exclusively in the active zone (AZ) of synaptic terminals: a morphologically specialised region that directly opposes the post-synaptic compartment. In contrast, LDCVs in axon terminals are concentrated at a distance from the AZ (Zupanc, 1996). Consistent with the localisation of LDCVs, neuropeptides are thought to be released from sites outside the active zone. This may occur in the vicinity of synapses (peri-synaptic release) but also at a distance from them; from non-synaptic regions of the axon, as well as dendrites (Nässel, 2009).

1.2.2.3 Neuropeptide release & signalling

Since the finding that neuropeptides present at the sympathetic ganglion of bullfrogs can diffuse tens of micrometers to reach their receptors (Jan and Jan, 1982), it has been well established that neuropeptides can act at a distance from their release sites. Thus, neuropeptides released from a particular axon terminal can act on directly opposed synaptic partners, for example to modulate responses to co-released FANs, but also on receptors present on neighbouring dendrites. The latter is sometimes referred to as intrasynaptic spill-over (Nässel, 2009). In addition, there is evidence for the diffusion of neuropeptides over long distances within the nervous system; a process known as volume transmission (Van den Pol, 2012). The capacity of neuropeptides to act at a distance from release sites is further supported by the existence of clear mismatches between the distribution of neuropeptides and their cognate receptors, reported in a range of organisms (Nässel, 2009). The range of action of

neuropeptides is, nonetheless, limited by the presence of peptidases, though various post translational modifications have been shown to render peptides less susceptible to degradation (Isaac et al. 2009), potentially facilitating the phenomenon of volume transmission. Though exocytosis of LDCVs is, like SCVs, is dependent on neuron stimulation (specifically pre-synaptic Ca^{2+} influx), there is evidence that level of stimulation required is greater. Thus, it is generally accepted that there is a higher activation threshold for the release of neuropeptides than there is for FANs (Zupanc, 1996).

With a few exceptions (summarised in: Nässel & Zandalawa, 2019; Jékely et al., 2018) neuropeptides signal through 7 transmembrane G-protein coupled receptors (GPCRs). Binding to the GPCR changes its conformation and initiates signalling through the heterotrimeric G proteins: Gα, Gβ and Gγ. Several types of these G proteins exist, and heterotrimers comprising different combinations activate distinct signalling cascades. For example, when dissociated from the Gβγ dimer, Gα(s) stimulates the activity of adenylyl cyclase, while Gα(i) and Gα(o) inhibit it. Another sub-type of Gα protein; Gα(q/11), targets phospholipase C. Though Gα proteins have, historically, been considered the key mediators of GPCR actions, Gβγ dimer complexes also target a variety of intracellular proteins (Hanlon & Andrew, 2015). The targets of G proteins include enzymes, receptors, ion channels and transcription factors (Hilger et al., 2018). Thus, neuropeptides can modulate neuronal excitability in the short term, for example, via modifications to ion channel proteins, and in the longer term, by modulating the expression of genes that can fundamentally change neuronal characteristics.

1.3 Neurogenetics in *Drosophila*

With a relatively small nervous system which is, nonetheless, capable of producing an array of complex behaviours (locomotion, sleep, aggression, learning and memory and reproduction, to name a few), *Drosophila melanogaster (*hereafter*, Drosophila)* represents an attractive model in which to establish causal links between genes, neural circuits, and behaviour.

Moreover, 100 years of genetic studies in *Drosophila* have resulted in a huge array of tools that permit the manipulation of single genes and neurons, in order that their relative contributions to behaviour can be established (Venken et al., 2011a)

1.3.1 Prevalent genetic tools in *Drosophila*

The sequencing of the entire genome in *Drosophila* (Adams et al., 2000) paved the way for technologies that permit mutation of specific genes (mutagenesis), to assay their function, and the introduction of transgenes into the *Drosophila* genome (transgenesis), to label and manipulate cells expressing a gene of interest.

Several widely used mutagenic strategies utilise Transposable Elements (TE's): naturally occurring elements that mobilise and insert at different regions in the genome in eukaryotes (Venken & Bellen, 2007). The 3 most popular in *Drosophila* are the P element, *Piggyback*, and *Minos* (Ryder and Russell, 2003). Insertion of these TE's within genes can disrupt their expression, thus TE insertions are mutagenic. Strategies have allowed controlled transposition of TE's be induced throughout genome, yielding a huge collection of *Drosophila* lines with TE's present in most coding genes. TE insertions do vary in terms of their degree of mutagenicity, however, with some inducing only partial gene loss-of-function (LOF) (Venken & Bellen, 2014). To ensure complete LOF, genetic deletions can be generated via Flp- mediated recombination between TE's flanking the target gene, that have been engineered to contain Flp recognition sites (FRTs) (Thibault et al., 2004; Golic and Golic, 1996; Ryder et al., 2004) (This method will be described in more detail in section 5.2.1). TE's have also been engineered to contain integration sites, facilitating their replacement with elements that confer greater mutagenicity. For example, an engineered version of the *Minos* TE: the Minos-Mediated Integration Cassette (MiMIC), incorporates inverted *attP* sites allowing recombinase-mediated cassette exchange (RMCE). Fisher et al. targeted MiMIC insertions with an engineered FlpSTOP construct, which contains several elements that disrupt transcription and translation, generating strong LOF alleles for several genes (Venken et al., 2011; Fisher et al., 2017). More recently established

gene editing methods are also used frequently in *Drosophila* to induce mutations. The most common is CRISPR-Cas9, which will be is described in more detail in section 3.3.2.

Another prevalent approach in *Drosophila* involves the insertion of exogenous transgenes into the genome, facilitating labelling and manipulation of cells that express a gene of interest (GOI). Restriction of transgene expression to specific cells (i.e., those expressing a GOI) is facilitated by the availability of binary expression systems. The most popular of these is GAL4/UAS. GAL4 is a transcriptional activator, identified in yeast, that binds to an upstream activation sequence (UAS) to drive the expression of genes downstream of the UAS (Fischer et al., 1988; Brand and Perrimon, 1993). Expressed ectopically, GAL4 does not drive the expression of any endogenous genes in *Drosophila*. If an additional transgene engineered downstream of a UAS sequence is present, however, expression of that transgene is activated by GAL4. The expression of GAL4 itself is directed by the *cis*-regulatory elements surrounding it. Coupling of GAL4 expression to an endogenous GOI can, thus, be achieved by incorporating *cis*-regulatory elements of the GOI into the inserted GAL4 construct. Alternatively, various methods permit the insertion of a GAL4 encoding construct directly into the GOI (these methods, and their practical implications, will be discussed in more detail in section 3.2). Split-binary systems can be used to further refine the expression of transgenes. In the case of split-GAL4 (Luan et al., 2006), GAL4 is split into two domains: an activation domain (AD) and DNA binding domain (DBD), each of which is expressed under the control of *cis*-regulatory elements of two different GOIs. The AD and DBD can heterodimerise via leucine zippers, thus functional GAL4 is reconstituted only in cells where the expression of the GOIs overlaps. In this way, the existence (or absence) of GOI co-expression can be identified, and the effects of transgenes can be targeted to more restricted neuronal populations.

1.3.2 The *Drosophila* nervous system

The central nervous system *Drosophila* (CNS) comprises the brain and ventral nerve cord (VNC). The brain contains an estimated ~100,000 neurons, about a million times fewer than

the human brain (Venken et al., 2011a). The diversity of neuronal and glial sub-types is, however, largely shared, with the *Drosophila* brain containing an array of sensory neurons, interneurons, motor neurons and neurosecretory cells (Croset et al., 2017). In addition, communication between neurons in the *Drosophila* nervous system relies upon many of the same molecules as in vertebrates. Fast synaptic transmission is mediated by the FANs Acetyl Choline (Ach), Glutamate and Gamma-aminobutyric acid (GABA), and biogenic amines, such as Dopamine (DA) and Seratonin (5-HT), and many neuropeptides are also shared, which perform analogous neuromodulatory functions (Venken et al., 2011a).

In the *Drosophila* brain, all neuron somata are located in a rind layer on the brain surface. From the soma, cell body fibres (CBFs) project inwards, branching into axons and dendrites in the inner region known as the neuropil, where the synapses are located (Strausfeld & Meinertzhagen, 1998) (Scheffer et al., 2020). The neuropil itself is organised into a variety of distinct anatomical compartments, or neuromeres, each associated with different functions.

Ascending and descending neuronal processes enter and exit the brain via the cervical connective, which connects the brain to the VNC- located in the thorax. The organisation of the VNC into a rind and neuropil layers is the same as in the brain. The VNC is sometimes compared to the vertebrate spinal cord, for it is where efferent motor neurons controlling most organs and body muscles exit the CNS, and afferent sensory neurons from the same tissues enter it, via a series of nerve trunks. The VNC is organised accordingly, into several discrete neuromeres along the dorso-ventral axis, each associated with innervation of different body regions (Court et al., 2020). The peripheral nervous system (PNS) in *Drosophila* refers to the afferent and efferent neurons innervating the musculature, organs, and other peripheral tissues, as well as the various sensory organs with which 1st order sensory neurons are associated.

1.3.2.1 Neuropeptides in the *Drosophila* nervous system

In excess of 40 genes encoding peptides and their cognate receptors are present in the *Drosophila* genome (Nässel & Zandalawa, 2019). Of these, a significant number appear to be expressed by neurons. Indeed, extracts of *Drosophila* brain were found to contain 42 neuropeptides encoded by 18 different genes (Yew et al., 2009). More recently, scRNAseq approaches have confirmed the expression of dozens of neuropeptides in the nervous system (Croset et al., 2017; Allen et al., 2019). Peptides are found in most types of neuron in *Drosophila* (Nässel & Winther, 2010). Many, however, express in a relatively small number of specialised 'neurosecretory' cells, sometimes referred to as peptidergic neurons. These neurons are specially adapted for the storage and release of neuropeptides. Their transcriptional profile includes a variety of genes implicated in the synthesis and processing of peptide precursors (Allen et al., 2019). Morphologically, they typically have enlarged somata and extensive axonal arborisations, allowing the storage of large quantities of peptide (which scRNAseq data suggests are expressed at very high levels (Allen et al.,2019)) and their release throughout the CNS. Regions of the brain, such as the pars intercerebralis (PI) - the *Drosophila* equivalent of the mammalian hypothalamus, contain a particularly high density of peptidergic neuron somata. Secretory neurons may also innervate the corpora cardiaca or corpora allata, facilitating secretion into the haemolymph (Nässel & Winther, 2010).

Peptides released from neurosecretory cells can be considered to be extrinsic to the circuits that they modulate. The presence of neuropeptides co-expressed with FANs within local circuit elements, where they function as co-transmitters, is more established in vertebrates than in *Drosophila* (Nässel, 2019). scRNAseq and other expression data has, however, identified a number of neuropeptides that are expressed in larger numbers of neurons in the *Drosophila* nervous system, co-expressed with FANs (Croset et al., 2017; Allen et al., 2019). The most notable example of this is short neuropeptide F (sNPF), which is found in many thousands of neurons, predominantly in the antennal lobe (AL) and mushroom body (MB) that also release Ach (Carlsson et al., 2010; Knapek et al., 2013).

1.3.3 Reproductive behaviour in *Drosophila*

The set of behaviours surrounding reproduction are among the most intensely studied in *Drosophila.* In part, this is because they are among the most overt and easily quantifiable. As such, the consequences of genetic and/or neuronal manipulations are often more tangible (Pavlou & Goodwin, 2013; Villella & Hall, 2008). In addition, the clear sexual dimorphism that exists in reproductive behaviour must relate to structural and functional differences at the level of the nervous system. Comparing neuroanatomical features of the male and female nervous system can, therefore, provide more general insights into how nervous system structure relates to function.

1.3.3.1 Male reproductive behaviour

Reproductive behaviour in *Drosophila* males centers upon male courtship: a complex sequence of behaviours that have evolved to facilitate copulation and successful sperm transfer. To initiate courtship, the male is guided by visual, olfactory, and auditory cues, which alert him to the presence of a female (Cook, 1979). The male initiates courtship by orientating towards the female and actively pursuing her. He extends his foreleg and 'taps' the female, allowing him to sample non-volatile pheromones present on her abdomen. (Manning, 1959; Venard et al., 1989; Greenspan and Ferveur, 2000; Amrein and Thorne, 2005; Thorne et al., 2005). Next, the male produces the most overt aspect of courtship. He extends one wing and vibrates it, generating a courtship song. The song incorporates two characteristic elements: *pulse song*; repetitive sound pulses, and *sine song*; a lower frequency (~60Hz) extended sound, with the two modes serving different functions. For example, sine song enhances the receptivity of the female, while the interval between pulses (ipi) (~34ms in *Drosophila)* is crucial for species recognition (Bennet-Clark and Ewing, 1969; von Schilcher, 1976; Ritchie et al., 1999). The complexity of the song structure thus provides the male with an opportunity to demonstrate his fitness, and, for the female, a tangible measure of this fitness. If the female is

receptive to courtship, the male guides himself into position by grasping the female with his forelegs and curling his abdomen to achieve intromission. Engagement with the female is maintained until sperm and seminal fluids have been transferred, after which the male detaches himself (Villella & Hall, 2008).

1.3.3.2 Female reproductive behaviour

During courtship, the behaviour of the female is far less conspicuous and, therefore, more difficult to quantify. For this reason, it has been the subject of significantly fewer studies (Ferveur, 2010). However, it is the female that ultimately decides whether copulation will occur. This decision she bases both on her assessment of the male's suitability; conveyed via a range of sensory stimuli, most notably courtship song (Dickson, 2008), and her own reproductive status at that point in time. Initially, when courted by a male, females display almost continuous movement, demanding active pursuit by the male, and may also produce active rejection behaviours, including kicking and ovipositor (Connolly and Cook, 1973; Ejima et al., 2001; Spieth & Ringo, 1983). If the female decides to sanction mating, she will cease rejection behaviours, slow movement, and open her vaginal plate to facilitate an attempt at copulation (Spieth, 1974; Spieth and Ringo, 1983). Thus, while the male is ostensibly the active party during courtship, the burden of decision-making falls to the female, and the execution of that decision is very much an active process.

Unlike the male, the female's contribution to successful reproduction does not end at copulation. After mating, the female becomes unreceptive to subsequent male courtship and dramatically increases her rate of ovulation and oviposition. Profound changes to her feeding habits, locomotor behaviour and sleep patterns also occur, so as to maximise the number and quality of her offspring (Carvalho et al., 2006; Kubli, 2003; Kubli, 2008).

1.3.3.3 Genetic and neural correlates of *Drosophila* reproductive behaviours

In *Drosophila,* sex-specific differentiation of the neural circuits governing reproductive behaviours is regulated by two transcription factor (TF) proteins: Fruitless (Fru) and Doublesex (Dsx). (Goodwin & Hobert, 2021). In the nervous system, these TFs regulate gene expression in specific neurons to guide their formation into sex-specific and sexually dimorphic neural circuits that engender reproductive behaviours. The *fru* and *dsx* genes are, in fact, terminal genes of the sex determination hierarchy. Early events in sex determination result in the expression of Transformer (Tra) protein in females, and an absence of this protein in males. The presence of Tra results in splicing of *dsx* to give rise to Dsx^F*:* the female-specific isoform of the protein. Female-specific splicing of *fru* introduces a stop codon into the mRNA, meaning that it is not translated. Consequently, females lack functional Fru protein. Absence of Tra in males results in 'default' splicing of dsx and fru to give rise to male-specific dsx^M and Fru^M proteins. Thus, in males, sexual differentiation of specific cells and tissues is orchestrated by Dsx^M and Fru^M in tandem, while in females, DsxF performs this function alone (Pavlou & Goodwin, 2013).

Initially, the importance of *fru* and *dsx* in reproductive behaviour was identified through behavioural analysis of males and females harbouring mutations in these genes (Goodwin & Hobert, 2021). How and why these genes are important was subsequently revealed by examining their patterns of expression in the nervous system (Siwicki & Kravtiz, 2009). The expression of Dsx is highly restricted in in the CNS of both sexes. In males, *dsx* is expressed in ~ 150 neurons in each brain hemisphere, forming 10 distinct anatomical clusters. In the female brain, 30-40 neurons in each hemisphere express *dsx*, grouped into 7-8 discrete anatomical clusters (Lee et al. 2002; Nojima et al. 2021; Rideout et al. 2007, 2010). *dsx* neurons have also been identified in the VNC of both sexes. In both cases, most of these neurons are located in the Abg. There are, however, male-specific *dsx* neurons in more anterior segments of the VNC. Close analysis of sex differences in *dsx* expression throughout the CNS has revealed that all *dsx* clusters are sexually dimorphic in terms of cell numbers

and/or connectivity (Nojima et al., 2021). A restricted neuronal population in the PNS also expresses *dsx*: a population of sensory neurons in the foreleg in both sexes (~72 in males and ~58 in females) (Rideout et al., 2010), and a small population of sensory neurons in the female reproductive tract (Rezával et al.,2012).

A larger, but overlapping, neuronal population expresses Fru^M protein in males: about 1700 neurons in the adult CNS, constituting ~20 anatomical clusters located across discrete functional compartments. Fru^M is also expressed in a variety of sensory neurons throughout the male PNS, including the legs, antennae, proboscis, wing joints and genitalia (Manoli et al., 2005).

The expression of Fru and Dsx in sensory neurons, interneurons and motor neurons in the adult nervous system suggests these neurons form a circuit capable of receiving and processing information relevant to reproductive behaviours and executing sex-specific behavioural responses (Pavlou & Goodwin, 2013; Goodwin & Hobert, 2021). Consistent with this, manipulating the activity of *dsx*- and *fru*- expressing neurons has been shown to elicit profound effects on reproductive behaviours in both sexes (Auer & Benton 2016). Thus, while *dsx* and *fru* neurons undoubtedly do not function in isolation to produce reproductive behaviour, they appear to constitute the core neural circuitry underlying these behaviours.

1.3.3.4 Modulation of reproductive behaviour by neuropeptides

The advantage of genetic determination, or ‘hardwiring’, of reproductive behaviours is clear: both males and females are innately capable of producing them. Like all behaviours, however, it is essential that these behaviours can be adapted according to environment and physiological context. In both sexes, mating itself alters the performance of reproductive behaviours: females become less sexually receptive after mating, prioritising behaviours that will maximise the production of viable progeny, while males start to suppress courtship behaviour after repeated copulation (Zhang et al., 2016). Reproductive behaviours are also modulated by learning and social environment, and by the need to perform other vital functions,

such as feeding and sleep (Ellendersen & Von Phillipsborn, 2017). The mechanisms by which these contexts are represented in the nervous system and translated into behavioural adaptation, remain only partly understood. There is, however, growing evidence that neuropeptides form an important part of this process. Considered below are some of the neuropeptides that have been implicated in the modulation of sexual (and, in some cases, specifically reproductive) behaviours.

1.3.3.4.1 Myoinhibitory peptide (MIP) and sex peptide (SP)

MIP and SP are distinct peptides with differing action in the CNS, but which share a common receptor: the sex-peptide receptor (SPR). Strictly speaking, SP is not a neuropeptide: it is produced by the male accessory gland (Chen et al., 1988). However, SP signalling through SPR underlies what is the perhaps the most studied example of reproductive behaviour modulation in *Drosophila:* the female post-mating response (PMR). After mating, females undergo a dramatic change in their behaviour, becoming unreceptive to male courtship and increasing their rate of egg laying (Kubli,2003). Oher, more subtle, changes in behaviour have also been described more recently (Isaac et al., 2010; Hussain et al., 2016). These post-mating effects are triggered by SP, present in the male seminal fluid, acting on SPR expressed in sensory neurons in the female reproductive tract to modulate their input to central circuits (Kubli, 2003; Liu & Kubli, 2003; Chapman et al., 2003; Yapici et al., 2008; Häsemeyer et al., 2009; Rezával et al., 2008).

MIP has been shown to express in neurons that arborise extensively in the brain and VNC (Kim et al., 2010; Min et al., 2016). Despite evidence that SP can enter the haemolymph it has not been detected in the CNS, yet SPR is thought be expressed in the CNS in both sexes (Yapici et al., 2008). This suggests that central SPR signalling is mediated exclusively by MIP. Though MIP is the ancestral ligand of SPR (with SP likely evolving relatively recently in a few *Drosophilae* sub-species, specifically to trigger the PMR) (Kim et al., 2010) it was identified as a ligand more recently. Consequently, the functions of MIP-SPR signalling have been the

subject of fewer studies. A role for MIP in the modulation of aspects of female reproductive behaviour has been evidenced, however. One study found that the peptide mediates changes in the female's preference for certain nutrients that occurs after mating (Hussain et al., 2016). Another study provided evidence that MIP release in the female Abg may occur downstream of SP-SPR signalling in the uterus, to mediate the PMR (Jang et al., 2017). Whether MIP-SPR modulates aspects of male reproductive behaviour has yet to be investigated. MIP also appears to mediate other functions in the CNS, as suppression of MIP signalling was shown to affect feeding behaviour in both sexes. Interestingly, however, this function was found to be independent of SPR, suggesting the existence of an additional receptor for the peptide (Min et al.,2016).

1.3.3.4.2 SIFamide

AYRKPPFNGSLFamide, or SIFamide, was originally purified from the grey fresh fly, *Neobellieria bullata*, and found to be a highly conserved neuropeptide among various arthropod species (Janssen et al., 1996; Verleyen et al., 2004). A receptor for SIFamide, SIFaR, was subsequently identified, which is also highly conserved among arthropods and appears to be activated exclusively by SIFamide (Jørgensen et al., 2006). In the *Drosophila* larval nervous system, SIFamide is expressed in just 4 neurons in the pars-intercerebralis (PI) that project the entire length of the CNS (Verleyen et al.,2004; Terhzaz et al., 2007). In the adult CNS, these 4 neurons showing dense and complex innervation of most regions of the neuropil, suggesting that they arborise significantly during development (Terhzaz et al.,2007). Ablation of SIFamide neurons leads to intriguing reproductive behaviour phenotypes. SIFamide neuron-ablated males produce courtship behaviour when placed with wild-type males, as well as females. Females lacking SIFamide neurons copulate with wild-type males more quickly; an indicator of increased sexual receptivity. The same phenotypes were observed when SIFamide expression was knocked down by RNAi, suggesting that it is loss of SIFamide *per se* that underlies these phenotypes. Based on the 'hypersexual' behaviour of

SIFamide perturbed animals, it has been suggested that the normal function of SIFamide is to suppress reproductive behaviours (Terhzaz et al., 2007). A follow up study found that indiscriminate courtship in males is also observed when the receptor for SIFamide, SIFaR, is knocked down by RNAi, though the penetrance of this phenotype was notably weaker. Expression of *SIFaR-RNAi* under the control of *fru-GAL4* also resulted in male-male courtship (Sellami & Veenstra, 2015). It should be noted, however, that the presence (or extent) of SIFaR expression in *fru* neurons was not investigated. Indeed, the distribution of SIFaR in the CNS in general is, to date, unclear. More recent studies of SIFamide have identified additional functions for this peptide. For instance, manipulation of SIFamide signalling disrupts both sleep and appetitive feeding behaviour (Park et al.,2014; Martelli et al.,2017; Huang et al.,2021).

1.3.3.4.3 Corazonin

Corazonin (Crz) was first discovered as a modulator of cardiac function in the cockroach (Veenstra, 1989), and subsequently shown regulate a range of physiological processes in different insect species (Tawfik et al., 1999; Wise et al., 2002; Kim et al., 2004). In the adult *Drosophila* brain, Crz is expressed in approximately 8 cells in each hemisphere of the central brain (Zhao et al., 2010). In addition, the peptide is found in a male-specific cluster of 4 neurons in the Abg (Tayler et al.,2012). Close association of Crz neurons with areas of neurohemal secretion suggests that, as in other species, Crz may serve hormonal functions in *Drosophila.* Central neuromodulatory functions have also been evidenced, however, including the regulation of aspects of reproductive behaviour. Males in which Crz neurons have been silenced show normal courtship behaviour but copulate for a dramatically increased period of time (increased from ~20 minutes to ~100), during which time they fail to transfer sperm or seminal fluid. Conversely, thermogenetic activation of Crz neurons, or injection of the peptide, triggers male ejaculation. Male ejaculation still occurs in decapitated males and is absent in males in which only Crz brain neurons are activated, suggesting that the function of Crz in regulating male ejaculation is mediated by neurons in the Abg (Tayler et al., 2012). The use of

a *CrzR-GAL4* line, generated in the same study, evidenced *CrzR* expression in Abg serotongergic neurons innervating the male reproductive tract, activation of which also triggers ejaculation. Crz signalling in the nervous system has also been shown to mediate behavioural responses to nutritional stress in *Drosophila* (Kapan et al.,2012; Zhao et al., 2010). This function is common to both sexes, however some sex-differences in this function have been reported (Zhao et al.,2010).

1.3.3.4.4 Neuropeptide F

Neuropeptide F (NPF) is sometimes referred to as dNPF (*Drosophila* NPF), or *long*-NPF, to distinguish it from short NPF (sNPF); a distinct peptide encoded by a different gene. The nutritional environment of *Drosophila* larvae is thought to be encoded by NPF signalling, with the peptide modulating foraging and social behaviours accordingly (Shen et al., 2001; Wu et al.,2003). In the adult, NPF signalling through its receptor, NPFR, regulates ethanol sensitivity and intake (Wen et al.,2005; Shohat-Ophir et al.,2012); aggression (Dierick et al.,2007); locomotor rhythms (He et al., 2013), and (independently) feeding and sleep-wake behaviours (Zhang et al.,2013; Chung et al., 2017). The role of NPF in the modulation of reproductive behaviour is less well explored. However, one study found that the peptide influences the modulation of mating duration normally displayed by males in response to social context (Kim et al., 2013). In addition, neuronal expression of NPF has been shown to be regulated by *fru* (Lee et al., 2006). As a result, expression of the peptide is sexually dimorphic, with NPF found in 26 neurons in the male brain and just 20 in females. Feminisation of male-specific NPF neurons was shown to decrease levels of male courtship, while ablation of the same neurons impaired evening anticipation: the increase in locomotor activity of *Drosophila* males usually shown during the light-dark transition (Lee et al.,2006).

1.3.3.4.5 Tachykinin

Via differential splicing, the *Drosophila Tachykinin* (*Tk*) gene gives rise to 6 distinct peptides (DTK1-6). Five of these are expressed in the adult CNS, in a population of approximately 150 neurons (Winter et al., 2003; Nässel & Winther, 2010). Like Crz, some of these neurons innervate areas of the brain in which secretion into the haemolymph occurs, implying a hormonal function for the peptide. However, like Crz, neuromodulation of behaviour has been shown to occur via Tk signalling in the adult CNS. Tk was first implicated in the control of aggression in *Drosophila* males, via its release from a small population of *fru* neurons in the male brain (Asahina et al., 2014). Male courtship and aggression are thought to rely upon some shared neural circuit elements. However, manipulations of Tk neurons performed in this study left courtship behaviours intact, suggesting these two behaviours can be modulated separately, and implicating Tk in the control of aggression specifically. However, a subsequent study found that Tk expression in the brain is required for the detection of anti-aphrodisiac gustatory pheromones that inhibit courtship, with males lacking the peptide failing to suppress courtship in response to these signals (Shankar et al., 2015). Thus, in certain contexts, Tk appears to modulate male courtship behaviour, as well as aggression.

1.4 Study objectives

In understanding how neuropeptides perform their diverse functions, knowledge of where they act is essential. The availability of effective antisera has allowed the distribution of many neuropeptides to be mapped via immunohistochemistry. However, due to the ability of peptides to diffuse over long distances, the extensive axonal projections of many peptidergic neurons, and the possibility of neurohemal secretion, a neuropeptide's distribution does not necessarily define its loci of action. As such, knowledge of the distribution of neuropeptide receptors (NPR) is crucial to understanding where neuropeptides act, yet represents a significant gap in our understanding (Nässel & Zandalawa, 2019). Thus, the first key objective of this study was to map the distribution of neuropeptide receptors in the nervous system (Chapter 3). Receptors for neuropeptides implicated in the modulation sex-specific and reproductive behaviours (discussed in section 1.3.3.4) were selected for this investigation. These 'candidate' *NPR*s were the sex peptide receptor (SPR); SIFamide receptor (SIFaR); corazonin receptor (CrzR); neuropeptide F receptor (NPFR) and Tachykinin receptor (Tkr86C). To understand how these neuropeptides function specifically in the modulation of reproductive behaviour a key question is whether their receptors are expressed in the neural circuits governing these behaviours. Thus, the expression of candidate NPRs in *fru* and *dsx* neurons was investigated (Chapter 4). Owing to the particularly intriguing reproductive behaviour phenotypes associated with SIFamide perturbation, subsequent investigation focused upon this neuropeptide. Behavioural assays were conducted to examine the extent to which signalling through SIFaR, both globally and in *dsx* neural circuits, is important for wild-type reproductive behaviours. Physiological experiments were also piloted to ascertain how SIFamide might modulate the properties of these neural circuits *in-vivo* (Chapter 5). Finally, the contexts in which SIFamide might modulate reproductive behaviour were examined, and an analysis of SIFamide neuron morphology and connectivity was performed to shed light upon how it performs its integrative functions (Chapter 6).

Chapter 2: Materials and Methods

2.1 *Drosophila* husbandry & behavioural assays

2.1.1 *Drosophila* stocks

Drosophila stocks were maintained at room temperature (~22°C). Flies used in experiments were housed under 'standard conditions': 25°C, 50% relative humidity, on a 12-hr light/12-hr dark cycle. All flies were housed in plastic vials or bottles containing food medium consisting of: 7.2g agar, 25g yeast, 47.2 g cornmeal, 100g dextrose, 2.2g tegosept and 8.4 ml ethanol per litre of water. *Drosophila* stocks used in this study are listed in Table 1. Before use in experiments, stocks were isogenised by crossing to Canton-S flies for at least 6 generations. To genotype flies, DNA was isolated using the DNeasy Blood & Tissue Kit (QIAGEN, cat# 69504), following the manufacturers recommendations. Polymerase chain reactions (PCR) were carried out on a Surecycler 8800 using MyTaq™ Red Mix (Bioline, cat# 25043). PCR primers were designed with Primer3 in Geneious 20.1.2 software.

2.1.2 Behavioural assays

Flies used in behavioural experiments were collected within 4 hours of eclosure. After collection, virgin males were housed individually, for 5-7 days, and virgin females in groups of no more than 3, for 3-5 days, under standard conditions. For each sex, only individuals collected on the same day were used in any single experiment.

Courtship assays: unless otherwise stated, all courtship assays were performed under standard conditions in custom made acrylic courtship arenas (https://southernacrylics.co.uk/). Each arena comprised 20 individual circular courtship chambers (20 mm diameter, 3.2 mm height) with a retractable divider, in a 4x5 arrangement. After transferring individuals to the chambers, the arena was mounted on a custom 3D printed support, backlit by a FLFLSi200-IR infrared backlight (http://www.falcon-illumination.com/productdetail_FLFL.php). Videos were recorded using a Basler ace A2440-75um camera (cat# 35927) with a 35mm VIS-NIR fixed focal length lens (cat# 67716) and a UV/VIS cut-off filter (cat# 89834) from Edmund

Optics (https://www.edmundoptics.co.uk/). The uncompressed AVI videos were recorded at a resolution of 2400x1600 pixels, 8-bit grayscale, 25 frames per second, using LabVIEW software (https://www.ni.com/en-gb/shop/labview.html). *Ad libitum* feeding assays were performed as per Martelli et al., 2017.

2-choice/courtship preference assays: males were placed in a courtship chamber with a single wild-type virgin female and a droplet of water containing: 5% sucrose, 3% yeast, 0.25 % Agar and 2% activated charcoal (the latter to allow clear visualisation of the droplet on IR back-lit videos). Behaviours were scored for 10 minutes. Unlike standard courtship assays, behaviours were scored manually. Courtship index was calculated according to the proportion of the time the male was engaged in courtship behaviours (following, tapping, wing extension, attempted copulation). Feeding index was determined as the proportion of the time the male was orientated to the food droplet and engaged in feeding behaviour (proboscis extension). Courtship preference index was calculated as: time courting – time feeding / time courting.

2.1.3 Automated behaviour tracking

All courtship assay videos were subjected to automated tracking. The uncompressed AVI video files were converted to mFMF video files using 'any2ufmf' software (http://ctrax.sourceforge.net/any2ufmf.html). The mFMF video files were tracked with Caltech FlyTracker (http://www.vision.caltech.edu/Tools/FlyTracker/) 69 with the following settings: num_chunks = 1, num_cores = 1, max_minutes = 15, save_JAABA = 1. Courtship behaviours were annotated using the Janelia Automatic Animal Behavior Annotator, JAABA (http://jaaba.sourceforge.net/)70.

	Affected chromosome(s)	Origin	Stock reference	Description/notes
Isogenised strains				
Canton-S (CS)				Isogenised 'wild-type' strain
Balancer chromosome strains				
w⁻ ; if/Cyo; sb/TM3 $_{ser}$	II & III	Goodwin lab		
w⁺;;TM3 $_{sb}$/TM6 $_{tb,Hu}$	III	Goodwin lab		
'Driver' lines				
w⁺;lfa-GAL4	II	BDSC	84690	T2A-GAL4 inserted at or near the 3' end of *SIFa*
elav-GAL4	X	BDSC	458	
w⁺;;Actin5C-GAL4/TM6 $_{tb,Hu}$	III	BDSC	3954	
w⁺;;dsx-GAL4	III	Goodwin lab		GAL4 coding sequence inserted in *dsx* by ends-in homologous recombination
SPR-GAL4/FM7a	X	Goodwin lab		GAL4 coding sequence inserted in *SPR* by CRISPR-Cas9
w⁺;;CrzR $^{p65AD[MI03082]}$/TM3 $_{sb}$	III	BestGene	40188 (BDSC)	*y[1] w[*]; Mi{y[+mDint2]=MIC}CrzR $^{[MI03082]}$* stock injected with
w⁺;;NPFR $^{p65AD[MI15273]}$/TM3 $_{sb}$	III	BestGene	61009 (BDSC)	*y[1] w[*]; Mi{y[+mDint2]=MIC}NPFR $^{[MI15273]}$* stock injected with p65AD, [+] orientation
w⁺;;SIFaR $^{p65AD[MI05376]}$/TM3 $_{sb}$	III	BestGene	40221 (BDSC)	*y[1] w[*]; Mi{y[+mDint2]=MIC}SIFaR $^{[MI05376]}$* stock injected with p65AD, [-] orientation
w⁺;;SIFaRLexA[MI05376]/TM3 $_{sb}$	III	BestGene	84435 (BDSC)	*y[1] w[*]; Mi{y[+mDint2]=MIC}SIFaR[$^{MI05376]}$* stock injected withLexA, [-] orientation
w⁺,SPR $^{p65AD[CRISPR]}$/FM7a	X	Goodwin lab		p65AD coding sequence inserted in *SPR* by CRISPR-Cas9
w⁺;;TkR-86C $^{p65AD[MI05788]}$/TM3 $_{sb}$	III	BestGene	42116 (BDSC)	y[1] w[*]; Mi{y[+mDint2]=MIC}TkR86C[MI05788] injected with p65AD(2), -/- orientation.
W⁺;;fru ^{DBD}	III	Goodwin lab		
w⁺;;dsx ^{DBD}	III	Goodwin lab		GAL4^{DBD} inserted into *dsx* locus by ends-in homologous recombination
Effector' lines				
y ^{1},v ^{1};;UAS-dsRNA(SIFaR) 25831	III	BDSC	25831	RNAi against SIFaR (TriP). Used in *SIFaR∩dsx* LOF experiments
y ^{1},v ^{1};;UAS-dsRNA(SIFaR) 34947	III	BDSC	34947	RNAi against SIFaR (TriP)
w⁺;;UAS-RNAi(SIFaR) 1783	III	VDRC	1783	RNAi against SIFaR (GD library)
w⁺;;UAS-RNAi(SIFaR) 330283	II	VDRC	330283	RNAi against SIFaR (shRNA library)
w⁺;UASmCD8::GFP/Cyo	II	Goodwin lab		
UAS-CaLexA	II & III	BDSC	66542	*w[*]; P{w[+mC]=LexAop-CD8-GFP-2A-CD8-GFP}2; P{w[+mC]=UAS-mLexA-VP16-NFAT}H2, P{w[+mC]=lexAop-rCD2-GFP}3/TM6B $_{Tb[1]}$*
w⁺;UAS-Watermelon/Cyo; sb/TM3 $_{ser}$	II	Josh Dubnau		Membrane GFP and nuclear RFP under UAS control.
W⁻;UAS-Stinger	II	BDSC	84277	nuclear GFP under UAS control.
W⁺; lexAop-tdTomato; MKRS/TM6B $_{Tb1}$	II	BDSC	66690	nuclear RFP under LexAop control
w⁺;UAS-GcaMP(6s)/Cyo;UAS-GcaMP(6s)/TM3 $_{ser}$	II & III	Goodwin lab		Green flourescent ca^{2+} reporter under UAS control.
W⁺;UAS-SIFaR/Cyo	II	BestGene	24482 (BDSC)	*y[1] M{}ZH-2A w[*]; M{3xP3-RFP.attP'}ZH-51C* stock injected with pUAS-attB-SIFaR.
W⁺;UAS-dSTIM/Cyo	II	BDSC	41756	*Drosophila* Stromal Interaction Molecule expressed under *UAS* control
w⁺;UAS-FLP/Cyo	II	Goodwin lab		FLP recombinase expressed under UAS control
w⁻; UAS-Epac1-camps	II	BDSC	25407	Expresses a FRET sensor of cAMP under UAS control
w+;UAS-2eGFP	II	BDSC	6874	non-localised GFP, under UAS control
UAS-prepro ANF-EMD	X	BDSC	7001	Rat neuropeptide fused to an emerald variant of GFP, under UAS control
LOF insertion/deletion lines				
w⁺;;SIFaR ^{B322}	III	BDSC	16202	pBAC{5HPw} inserted in *SIFaR* (21,288,749 [-]).
W⁺;;Df(3R) $^{(ED10838)}$/TM2	III	BDSC	9485	Chromosomal deletion 3R: 21,134,314 – 21,296,499. ~25 deleted genes, including *SIFaR*
w⁻;;P{RS5} $^{5-SZ-3483}$	III	DGRC	126089	DrosDel collection. P{RS5} insertion upstream of *SIFaR* (3R:21,296,365 [+])
w⁻; Pbac{RB} ^{CG3301e03242}	III	BDSC	18137	Exelixis collection. Pbac{RB} insertion in *CG3301* , downstream of *SIFaR* (3R:21,275,133 [-])
w⁺;;SIFaR-Deletion/TM6 $_{tb,hu}$	III	Goodwin lab		Deletion from: 3R:21,275,133 – 3R:21,296,365. Made by hs-FLP mediated recombination between: w⁻;;P{RS5}$^{5-SZ-3483}$ and Pbac{RB}^{CG3301e03242}
w⁺;;SIFaR $^{FlpSTOP[-]}$/TM3 $_{sb}$	III	BestGene	40221 (BDSC)	*y[1] w[*]; Mi{y[+mDint2]=MIC}SIFaR $^{[MI05376]}$* injected with pFlpStop-attB-UAS-2.1-tdTom[-] 'disrupting' insertion orientation
w⁺;;SIFaR $^{FlpSTOP[+]}$/TM3 $_{sb}$	III	BestGene	40221 (BDSC)	*y[1] w[*]; Mi{y[+mDint2]=MIC}SIFaR $^{[MI05376]}$* injected with pFlpStop-attB-UAS-2.1-tdTom[+] 'non-disrupting' insertion orientation

Table 1: Genotypes and descriptions of *Drosophila* strains. BDSC = Bloomington *Drosophila* Stock centre, Indiana. VDRC = Vienna *Drosophila* Resource Centre, Vienna. KYOTO = KYOTO stock centre, Kyoto

2.2 Generation of novel *Drosophila* lines

2.2.1 *SPR-GAL4* and *SPR-p65AD* lines

These were designed generated in the Goodwin lab by Ammerins de Haan and Megan Neville. The pDsRed-attP vector (a gift from Melissa Harrison & Kate O'Connor-Giles & Jill Wildonger; Addgene plasmid # 51019) was modified, inserting the GAL4/p65AD coding sequence and hsp70 3'UTR between regions where homology arms are inserted, creating pDsRed-attP-GAL4/p65AD. Two fragments with homology to the SPR genomic regions on either side of the first codon (ATG) were cloned using Gibson Assembly Master Mix (New England Biolabs) into pDsRed-attP-Gal4 digested with XhoI and NotI. gRNA-expressing constructs pCFD3-SPR5 and pCFD3-SPR3 were generated in the vector pCFD3-dU6:3gRNA (a gift from Simon Bullock; Addgene plasmid #49410) against guide regions SPR5 (CCACGAGATGGACAACTATACGG) and SPR3 (CCAATTCGTATCGTGCCACATGC).

Constructs were co-injected into the strain *y[1]sc[1] v[1]; {y[+t7.7] v[+t1.8]=nanos-Cas9}attp40*, progeny were screened for $DsRed^{+}$ expression in the eye. Multiple independent lines were identified, all of which mapped to the X chromosome. PCR and sequencing were used to confirm insertions.

2.2.2 *UAS-SIFaR* line

Primers to amplify the ~2.3 Kbp CDS of the *SIFaR* gene were designed with Primer3 in Geneious 20.1.2 software with the following sequences:

SIFaR Forward: TCTGAATAGGGAATTGGGAATTCATGATGGCAGCCAGCGGG

SIFaR Reverse: GATCCTCTAGAGGTACCCTCGAGCTATTCGTAGAGATCCG.

PCR was performed on cDNA, generated from Canton-S head RNA extracts, over 35 cycles, with a 65°C annealing temperature for 20 seconds and a 60 second extension time. Reaction product(s) were run on a 1.3% Agarose gel, from which the target sequence was excised and

purified using QIAquick Gel Extraction and PCR & Gel Cleanup Kits (QIAGEN®, cat. # 28704,28706, 28506 and 28115). The purified sequence was inserted into the pUAST-attB-LT3-Dam plasmid, pre-digested with EcoR1 and Xho1 restriction enzymes, using the NEBuilder® HiFi DNA Assembly Cloning Kit (NEB, cat. # E5520S). Transformation of the assembled plasmid into *E. coli* was performed using the High Efficiency Transformation Protocol (NEB cat. # C2987H/C2987I). Plasmid DNA was purified using the NucleoSpin Plasmid QuickPure™ Kit (Macherey-Nagel™, cat. # 11902422). To confirm successful integration of the target sequence, the purified plasmid was digested with *Xho1* and the resulting product(s) were separated by Agarose gel electrophoresis, yielding two fragments corresponding to products of the predicted sizes. Finally, the purified plasmid was sequenced using primers targeting the predicted *SIFaR* CDS:

Forward 1: CATGCCCATTCCGACCATGGC

Forward 2: GCATCATGATAATCGGCATA

Forward 3: GTCATCGCACTACCACCAGA

Reverse 1: GCCGAGATCAAGCTGGTGGT

Reverse 2: AGGTGAAGGTCATCAAGATG

Reverse 3: TACGAACTTCTGCCGCTTATC

In-silico alignment of the plasmid sequence with the predicted *SIFaR* CDS was performed using Geneious 20.1.2 software. The *pUAS-SIFaR* construct was integrated into the genome by PhiC31 integrase-mediated transgenesis, with injections of the *pUAS-attB-SIFaR* plasmid into BDSC stocks # 24482 (*ZH51-C* docking site) and # 8622 (*attp2* docking site), facilitating integration on the second and third chromosomes respectively. Injections and screening for transgenic markers were carried out by BestGene.

2.2.3 *SlFaR*^{deletion} allele

Two transposable-element (TE) insertion lines, P{RS}(5-SZ-3483) and PBac{RB}(CG3301), were selected from an *in-silico* screen of TE's flanking the *SlFaR* locus. Larvae expressing both TE's and a heat sensitive FLP recombinase were incubated at 37°C for 1 hour to induce recombination. Recombinants, identified by mosaic loss of eye colour at the adult stage, were isolated and crossed individually to a balancer chromosome line. To confirm deletion of the target region, DNA extracted from the progeny was PCR amplified using a mixture of published and custom designed primers. Sanger sequencing of the PCR products was then performed to determine the complete sequence of the locus following recombination.

2.3 Western blotting

2.3.1 Protein extraction

For whole animal protein samples, 15 x 5-7-day old male flies were homogenized on ice, in a solution containing 400µl of homogenisation buffer (80% ddH2O, 10% Glycerol, 5% 1M Tris-HCl pH7.5, 3% 5M NaCl, 1% 0.5M EDTA and 1% Triton X-100) and 2.4µl of Protease-Phosphatase inhibitor cocktail (Sigma-Aldrich). For brain protein samples, ~50 brains were dissected in chilled PBS, then homogenised on ice in 30µl of homogenisation buffer and 2µl of Protease-Phosphatase inhibitor cocktail. All samples were centrifuged at 16,000 RCF for 15 minutes at 4°C. The supernatant was then stored at -80°C. Protein concentration for each sample was determined using the Pierce™ BCA Protein Assay Kit (Thermo Scientific™).

2.3.2 SDS-PAGE and electrophoretic transfer

Protein samples were diluted in Bolt™ LDS Sample Buffer (Invitrogen™), Bolt™ Sample Reducing Agent (Invitrogen™) and ddH_2O to concentration of ~2ug/µl, incubated at 99°C for 5 minutes then placed on ice. 15 µl of sample was loaded into each well of a Bolt™ 4-12%,

Bis-Tris, 1.0 mm, Mini Protein Gel. SeeBlue™ Plus2 Pre-stained Protein Standard (Invitrogen™) was used as a visible marker. SDS PAGE was performed at 145 Volts for 45 minutes at room temperature in NuPAGE™ MOPS SDS Running Buffer (Invitrogen™). Electrophoretic transfer of protein to the nitrocellulose membrane was performed at 100V for 1.5 hours in TowBin Buffer (25mM Tris, 192mM glycine) with 10% methanol.

2.3.3 Immunolabelling and detection

Membranes was stained briefly with Ponceau S Solution (Sigma-Aldrich) to confirm protein transfer, trimmed to size, then de-stained in 0.1M NaOH. After several washes in TBST (Tris Buffered saline with 0.1% Tween-20), membranes were incubated in blocking buffer (TBST with 5% Skim-milk powder) overnight at 4°C. Membranes were incubated in primary antibody (rabbit anti SIFaR (proteintech), 1:1000, or mouse anti- alpha tubulin (Sigma-Aldrich), 1:4000) for 1.5 hours at room temperature, washed in TBST, then incubated in secondary antibody (Goat anti-mouse or Goat anti-rabbit HRP (Invitrogen™), 1:10,000) for 45 minutes at room temperature, followed by several washes in TBST. Immunodetection was performed with the Syngene chemiluminescent system, using Pierce™ ECL Western Blotting Substrate.

2.4 RT-qPCR

2.4.1 RNA extraction & cDNA synthesis

For each genotype analysed, RNA was extracted from four biological replicates, each with 20 (10 male, 10 virgin female) 5-7 day whole adult flies, using TRIzol™ (ThermoFisher, cat# 15596026) following the manufacturer's instructions. Genomic DNA was removed from the samples using ezDNase™ and cDNA was synthesised using the SuperScript™ IV First-Strand Snthesis System (ThermoFisher# 18091050), following the manufacturer's instructions, with 1ug input RNA primed with oligo(dT).

2.4.2 qPCR

Primers for qPCR were designed using Primer3 in GeneiousPrime 20.1.2 software and evaluated for off-target priming with Primer-Blast (Ye et al., 2012). Potential for amplicon secondary structure formation was assessed using the DNAfold tool in GeneiousPrime 20.1.2. Optimal annealing temperatures were determined by gradient PCR with pooled cDNA, while primer specificity was evaluated by extraction and Sanger sequencing of the resulting fragments. qPCR was carried out using the RocheTM LightCycler® 480 with the LightCycler® 480 SYBR Green I Master kit (cat# 04707516001). Cycling conditions were performed following manufacturer's recommendations with 60°C - 70°C annealing temperature for 10 seconds, and a 10 second extension time. Data analysis was performed with qbase+ software, version 3.2 (Biogazelle, Zwijnaarde, Belgium - www.qbaseplus.com). Primer efficiencies were calculated from amplification of serial dilutions of a pooled cDNA template. Three technical replicates were run for each biological replicate across each primer set. Technical replicates with a difference in Cq value of greater than 0.5 were discarded. Reference gene stability was evaluated with geNorm (Vandesompele et al., 2002).

2.5 Immunohistochemistry

2.5.1 Tissue dissection and fixation

3-7-day old flies were briefly anaesthetised with CO_2 and kept on ice to maintain anaesthesia, before dissection in a 35mm dissecting dish containing ice cold 1X Phosphate Buffered Saline (PBS). Dissected samples were transferred immediately to a 24-well tissue culture plate containing PBS and kept on ice. Samples were fixed within 1 hour of dissection in PBS with 4% Paraformaldehyde (PFA) for 20-30 minutes at room temperature. Fixative was washed off in several changes of PBT (PBS with 0.3% Triton X-100).

2.5.2 Antibody staining

Samples were incubated in blocking solution (PBT with 5% Normal Goat Serum (NGS)) overnight at 4°C, then in primary antibody, diluted in blocking solution, for 2-3 hours at room temperature and an additional 2-3 days at 4°C. Primary antibodies used in this study are listed in table 2. Samples were washed in 5-6 changes of PBT for ~4 hours at room temperature, then incubated in secondary antibody (Alexa Fluor, Thermofisher) diluted 1:300 in blocking solution. overnight at 4°C. After repeating the washing steps, samples were incubated in PBT with 70% glycerol for 1-2 days at 4°C, then transferred to lysine-coated glass slides. Samples were submerged in Vectashield (Vecta laboratories) and covered with a 25mm x 25mm glass coverslip, sealed with nail polish.

Name	Target Antigen	Host species	Concentration	Origin
GFP	GFP	Chicken	1:1200	Abcam (ab13970)
nc82	Bruchpilot	Mouse	1:50	DSHB
mCherry	mCherry	Rat	1:1000	Thermofisher (M11217)
tdc2	tdc2	Rabbit	1:1000	Abcam (ab128225)
elav	elav	Rat	1:500	DHSB
repo	repo	Mouse	1:1000	DHSB
phalloidin	F-Actin		1:200	Abcam (ab176757)
SIFaR	SIFaR	Rabbit	1:1000	Proteintech
SPR	SPR	Rabbit	1:1000	Gift from Young-Joon Kim (GIST, Gwangju)
SIFa	SIFa	Rabbit	1:2000	Gift from Jan Veenstra (CNRS, Bordeaux)

Table 2: List of primary antibodies.

2.6 Fluorescence imaging

2.6.1 Confocal microscopy

Most confocal imaging experiments were carried out with a Leica DM500 CS upright confocal microscope with Leica TCS SP5 X laser system, using Leica LAS AF software to control acquisition parameters. Remaining confocal imaging experiments were carried out on a Zeiss 880 upright confocal microscope, using ZEN Black software to control acquisition parameters. In both cases, Z-stacks were generated by sequentially scanning optical sections at each wavelength at 200-400hz, with an interval of ~0.5 µm and a pinhole aperture of 1 airy-unit. Images were processed in Imaris 8.2.1 (Bitplane Scientific, AG, Zürich) and Fiji (ImageJ).

2.6.2 FRET imaging

Brains were dissected in ice-cold Ca^{2+} and Mg^{2+} free Adult haemolymph like saline (AHLS) and imaged Immediately. For imaging, brains were adhered to the bottom of a 35mm petri dish (Corning Falcon) and submerged in 1.8ml of AHLS containing Ca^{2+} and Mg^{2+}. 200µl of Test

compounds (dissolved in DMSO and diluted in AHLS) were added dropwise from a Gilson micropipette over a ~30 second time frame. Confocal imaging of FRET was carried out on a Axio Examiner upright Zeiss 880 fitted with Airyscan detector, using a 40x Apchromat water immersion objective (MICRON, Department of Biochemistry). CFP was excited with an argon laser, and CFP & YFP emission were scanned simultaneously on to separate detectors. Images were captured every 3 seconds. Analysis of confocal FRET data was carried out in Fiji (ImageJ). For the ROI selected, at each time point, mean YFP intensity was divided by mean CFP intensity. Background intensities were subtracted from the calculated YFP/CFP value at each time point. Widefield FRET imaging experiments were carried out in the Zaccolo lab, using a wide-field fluorescence microscope. All analysis of wide-field FRET data was performed by Andreas Koschinksi (Zacollo group).

2.6.3 *In-vivo* 2-photon calcium imaging

In-vivo calcium imaging was performed at 5.92 frames per second, 256 X 256 pixels resolution, using a 2-photon microscope (Scientifica) controlled by ScanImage 3.8 software. Excitation was via a Ti-Sapphire laser (Chameleon Ultra II, Coherent) at ~140 fs pulses, 80 MHz repetition rate, cantered on 910 nm. 5-7-day old adult virgin males were anaesthetized on ice and mounted in a custom 3D-printed recording chamber using melted wax. The head capsule was opened using forceps and the brain was bathed in carbogenated (95% O_2, 5% CO_2) AHLS. 'Target' males and females were anaesthetized on ice then mounted on a custom-built manipulator. Images were corrected for X and Y movement using image stabilizer in Fiji (http://www.cs.cmu.edu/~kangli/code/Image_Stabilizer.html). All experiments showing substantial movement in the Z plane were discarded. Regions of interest (ROI) were selected manually. Analysis was carried out using custom scripts in MATLAB, written by Annika Rings (Goodwin group). To obtain average ΔF/F traces, average fluorescence in the ROI was measured against the average baseline fluorescence from 10 s to 200 ms before the onset of the stimulus.

2.6.4 *Ex-vivo* wide-field calcium imaging

Brains or VNCs of 5-7-day old flies were dissected in ice-cold Ca^{2+} and Mg^{2+} carbogenated (95% O_2, 5% CO_2) AHLS, then adhered to polylysine-coated covers slips and submerged in AHLS. Micropipettes (glass capillaries, Harvard Apparatus, 300092) were connected to a picospritzer III (Intracell). A custom-built manipulator was used to position the tip of the microelectrode close to the target neurons. A custom LED was set up to provide a trigger signal 1 second before activation of the picospritzer. The air pressure delivered by the picospritzer was adjusted to the minimum required to eject solution from the microelectrode. Texas-red dye was included in the solution to allow ejection to be visualised by imaging test pulses in the red-fluorescence channel. Samples were illuminated with an X-Cite 120PC Q arc lamp. Images were acquired using a using a Scientifica Slicescope with a 403, 0.8 NA water-immersion objective and Andor Zyla sCMOS camera with Andor Solis software. Time series were acquired at 25 Hz (512 x 512 pixels,16 bit)

Chapter 3: Investigating the Distribution of Neuropeptide Receptors

3.1 Introduction

To determine where neuropeptides act, cells expressing the relevant receptor(s) must be identified. The nature of neuropeptide receptor (NPR) proteins; almost exclusively 7 transmembrane GPCRs, makes the production of antisera far more challenging than it is for their peptide ligands. Accordingly, there are few examples of NPR antibodies of sufficient affinity and specificity to accurately reveal protein distribution (Michel et al., 2009). As an alternative to the direct detection of NPR proteins, transgenic tools can be used to map the expression of NPR-encoding genes by labelling the cells that express them. However, depending on how these tools are generated, it can be unclear whether the expression pattern of the target gene is fully recapitulated. As such, knowledge of NPR distribution represents a significant gap in our understanding of peptidergic signalling systems (Nässel & Zandalawa, 2019).

Acknowledging the limitations of any single method for detecting gene or protein distribution, three independent approaches were adopted in this study to investigate the distribution of NPRs. First, a transgenic approach was used to generate *GAL4, split-GAL4* and *LexA NPR* 'driver' lines for the 5 'candidate' *NPR*s: *CrzR, NPFR, SIFaR, SPR* and TkR86C. Second, analysis of single cell RNA sequencing (scRNAseq) data, generated by the Goodwin lab, was performed, with a focus on these candidate *NPR*s. Finally, a novel antibody against SIFaR; the focus of functional investigations in this study, was designed and validated. This chapter discusses each of these approaches in turn.

3.2 Transgenic approaches to map neuropeptide receptor expression

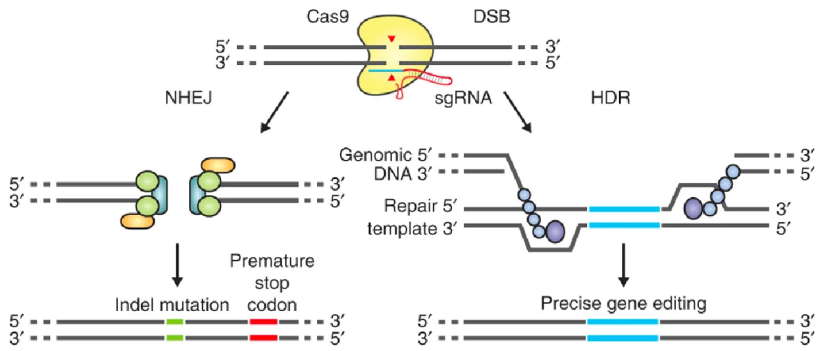
Transgenic approaches, whereby foreign DNA is inserted into the genome, offer a powerful means of visualising the expression of endogenous genes. In *Drosophila*, this is most often achieved via the use of binary expression systems. In this case, DNA encoding a visually detectable product, such as a fluorescent protein, is used as the 'responder' element,

expression of which is restricted to cells in which the 'driver' element is present. Coupling of driver expression to an endogenous target gene can be achieved by incorporating cis-regulatory elements of the target gene into the driver construct. In this way the transgene is, in theory, subject to the same transcriptional regulation as the target gene, irrespective of where it is inserted in the genome. In practice, however, the inability to incorporate all regulatory elements into the driver construct, coupled with positional effects, means expression of the target gene may not be fully recapitulated. To remedy this, several approaches permit insertion of driver constructs directly into target genes, ruling out positional effects and ensuring the entire complement of regulatory elements for the endogenous gene is present. To produce driver lines for candidate NPRs, two such approaches were adopted in the current study: CRISPR-Cas9 - used to generate SPR driver lines, and the Trojan MiMIC system (Diao et al., 2015) - used to generate driver lines for the remaining candidate NPRs.

3.2.1 Generation of *SPR-GAL4* and *SPR-p65AD* lines by CRISPR-Cas9

CRISPR-Cas9 is a method of precise and targeted genome editing. The technology utilises a naturally occurring endonuclease, cas9, which can generate a double stranded (DSB) in DNA. Using an engineered guide RNA (gRNA) sequence, with complementarity to a specific genomic locus, cas9 activity is guided to a specific site, where a DSB is induced. The DSB is repaired by one of two repair pathways: non-homologous end- joining (NHEJ) or Homology directed repair (HDR). HDR utilises a homologous DNA sequence as a template for repair. By supplying a template sequence with regions of homology (homology arms) flanking an exogenous sequence, a transgene (encoded by the exogenous sequence) can be inserted at the break point (Gratz et al., 2015) (fig.1A). Using this method, GAL4 and p65AD encoding sequences were inserted within an exon present in all *SPR* transcripts (fig.1B).

A



B

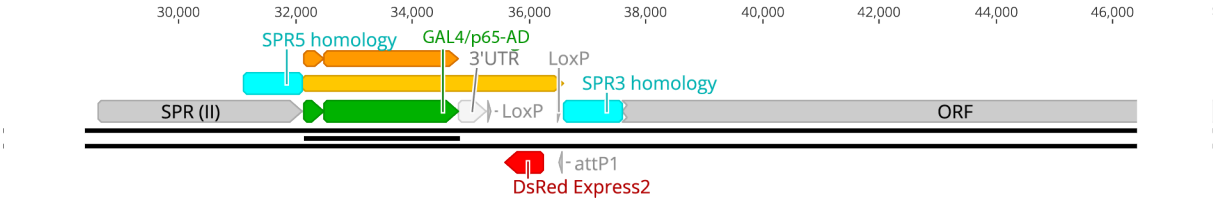


Figure 1: CRISPR-Cas9 genome editing. **(A)** Schematic overview of technology, taken from Ran et al., 2013. The cas9 endonuclease (yellow) is guided to the target site by a single guide RNA (sgRNA), with complementarity to the target site, where it induces a double stranded break (DSB). The DSB is repaired either by the NHEJ pathway: endogenous repair machinery (green and orange) re-ligates the DNA but with mutagenic errors, or by the HDR pathway, in which a repair template is used, which can be designed to include exogenous DNA (light blue). **(B)** *SPR* locus**,** following insertion of GAL4/p65AD encoding sequence via CRISPR-Cas9 HDR.

3.2.2 Generation of *split-GAL4* and *LexA* lines using the Trojan-MiMIC system

Like CRISPR-Cas9, the Trojan MiMIC system allows integration of DNA constructs directly into endogenous target genes. The technique utilises a collection of lines generated by insertion of the engineered Minos mediated integration cassette (MiMIC) at several thousand sites across the genome, many of which are located within the coding sequence of annotated genes. Inverted bacteriophage ΦC31 attP sites within the MiMIC transposon allow it to be replaced by a construct bearing complimentary attB recombination sites, via recombination- mediated cassette exchange (RMCE) (Venken et al., 2011b). Thus, to couple the expression of a transgene, like *GAL4*, to an endogenous target gene, a construct encoding the transgene,

preceded by a splice acceptor (SA), and flanked by attB recombination sites, is introduced into flies bearing a MiMIC in that gene. Initially, MiMIC insertions were restricted to 5' untranslated regions (UTRs), since co-translation of transgenes inserted in introns produces non-functional fusion proteins (Diao et al., 2015). However, with the development of 'Trojan' cassettes, which include a viral T2A-peptide sequence to promote synthesis of multiple protein products from a single transcript (Tang et al., 2009), it became possible to target MiMICs in coding introns, greatly increasing the number of potential gene targets (Diao et al., 2015) (fig.2A,B).

MiMIC sites within the coding introns of the 5 *NPR* candidates were identified by analysis of the annotated sequences in FlyBase (fig.2C). With a view to not only map the general distribution of these receptors but investigate their expression in sub-sets of neurons relevant to reproductive behaviour, *split-GAL4* lines were generated by insertion of a *T2A-p65AD-polyA* 'Trojan' exon into these MiMIC sites, via ΦC31 mediated RMCE (Injections performed by the BestGene). Given the particular focus on SIFaR in this study, a *SIFaR-LexA* line was also produced by insertion of *T2A-LexA-polyA* Trojan exon, via the same method. A summary of these lines can be found in Table 1.

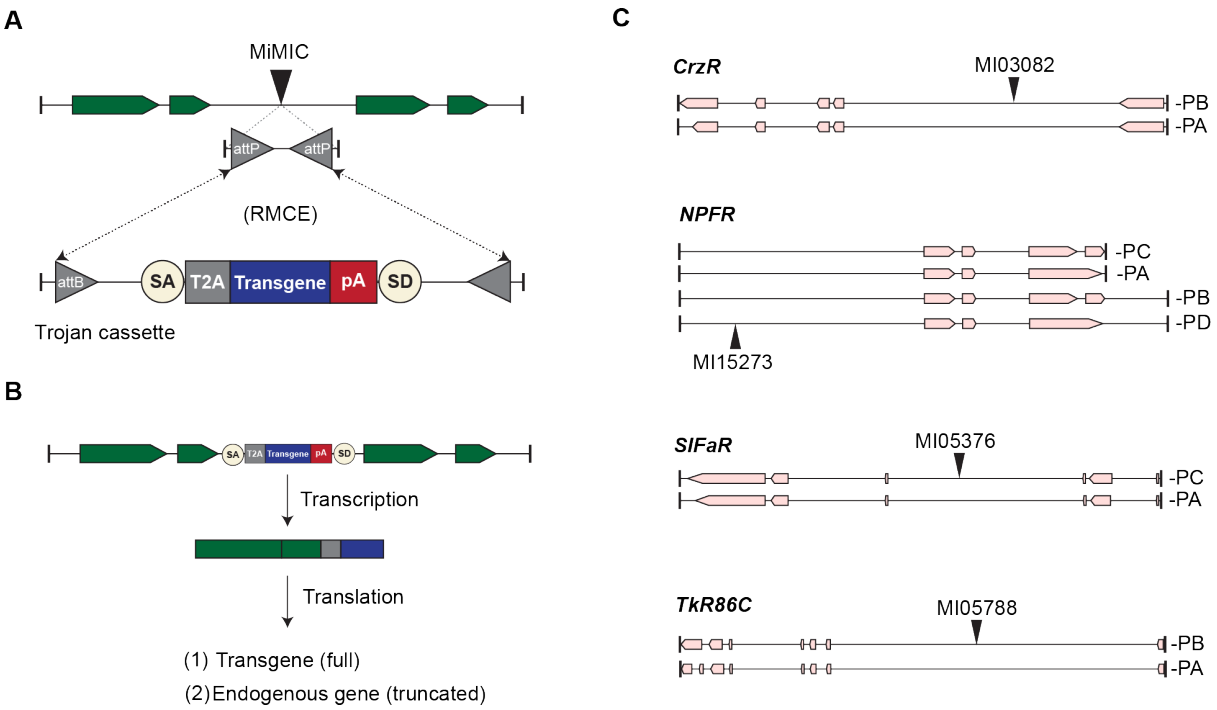


Figure 2. The Trojan MiMIC system. (A) Schematic showing insertion of a Trojan exon at an intronic MiMIC site, via Recombination Medicated Cassette Exchange (RMCE). SA/SD = Splice acceptor/Splice donor, pA = polyadenylation sequence**. (B)** Schematic showing incorporation of the Trojan exon into the endogenous transcript, which gives rise to two proteins - (1) & (2). **(C)** Isoforms of candidate neuropeptide receptor encoding genes, showing the locations of intronic MiMIC sites.

Target *NPR*	Transgenic Approach	Insertion Site	Inserted Construct(s)	Line(s) Generated
CrzR	Trojan MiMIC	MI03082 (3L:12,345,530 [+])	*T2A-p65AD-polyA*	*CrzR-p65AD* [MiMIC]
NPFR	Trojan MiMIC	MI15273 (3R:6,189,906 [-])	*T2A-p65AD-polyA*	*NPFR-p65AD* [MiMIC]
SIFaR	Trojan MiMIC	MI05376 (3R:21,286,336 [+])	*T2A-p65AD-polyA*	*SIFaR-p65AD* [MiMIC]
			T2A-LexA-polyA	*SIFaR-LexA* [MiMIC]
SPR	CRISPR-Cas9	(SPR5) X:5,478,764 (SPR3) X:5,479,214	*pDsRed-attP-GAL4*	*SPR-GAL4* [CRISPR]
			pDsRed-attP-p65AD	*SPR-p65AD* [CRISPR]
TkR86C	Trojan MiMIC	MI05788 (3R:10,745,595 [+])	*T2A-p65AD-polyA*	*TkR86C-p65AD* [MiMIC]

Table 1: Summary of transgenic *NPR* driver lines

3.2.3 Applying transgenic tools to map neuropeptide receptor distribution

3.2.3.1 *SPR-GAL4*$^{[CRISPR]}$ and *SPR-p65AD*$^{[CRISPR]}$ expression in the CNS

Existing data on SPR expression in the CNS comes from the use of SPR antisera, which reveals broad expression the brain and VNC (Yapici et al., 2008). However, given the propensity of GPCR antibodies to produce non-specific staining (reviewed in Michel et al., 2009), further evidence is needed to establish the distribution of this receptor. The *SPR-GAL4*$^{[CRISPR]}$ line was combined with a range of fluorescent reporters to characterise its expression pattern. The use of membrane-localising GFP variants (mGFP), such as *mCD8::GFP*, can provide insights into the morphology of expressing cells. However, in the CNS of *SPR-GAL4;UAS-mCD8::GFP* flies, a diffuse GFP signal covering much of surface of the brain and VNC was observed, in which individual cells were indistinguishable. To resolve single cells, an alternative reporter, *UAS-stinger*, was used, which encodes a GFP that localises to the cell nucleus (nGFP). In this case, GFP^{+} nuclei of a large number of cells were observed, covering most of the CNS surface. This was the case in both sexes (fig.3A). The diffuse GFP distribution produced by the membrane-localising GFP variant, together with the pervasive expression, suggested that *SPR-GAL4* may be expressed in glial cells, as well as neurons. To investigate this possibility, the nervous systems of *SPR-GAL4;UAS-stinger* flies were stained with *repo*, an antibody that marks the nuclei of all glia. In these samples, co-localisation of GFP and repo was evident in nuclei across the CNS, suggesting that *SPR* is expressed in glia (fig.3B). However, among all GFP^{+} cells in this line, a greater proportion did not co-stain with *repo*, indicating that the majority of *SPR*-expressing cells in the CNS are neurons. Consistent with this, combining the *SPR-p65AD*$^{[CRISPR]}$ *split-GAL4* line with the *elav-DBD* hemidriver, in which the DNA binding domain (DBD) of GAL4 is expressed only in neurons, produced GFP expression in large numbers of cells across the CNS in both sexes, indicating extensive neuronal expression of *SPR* (fig.3C).

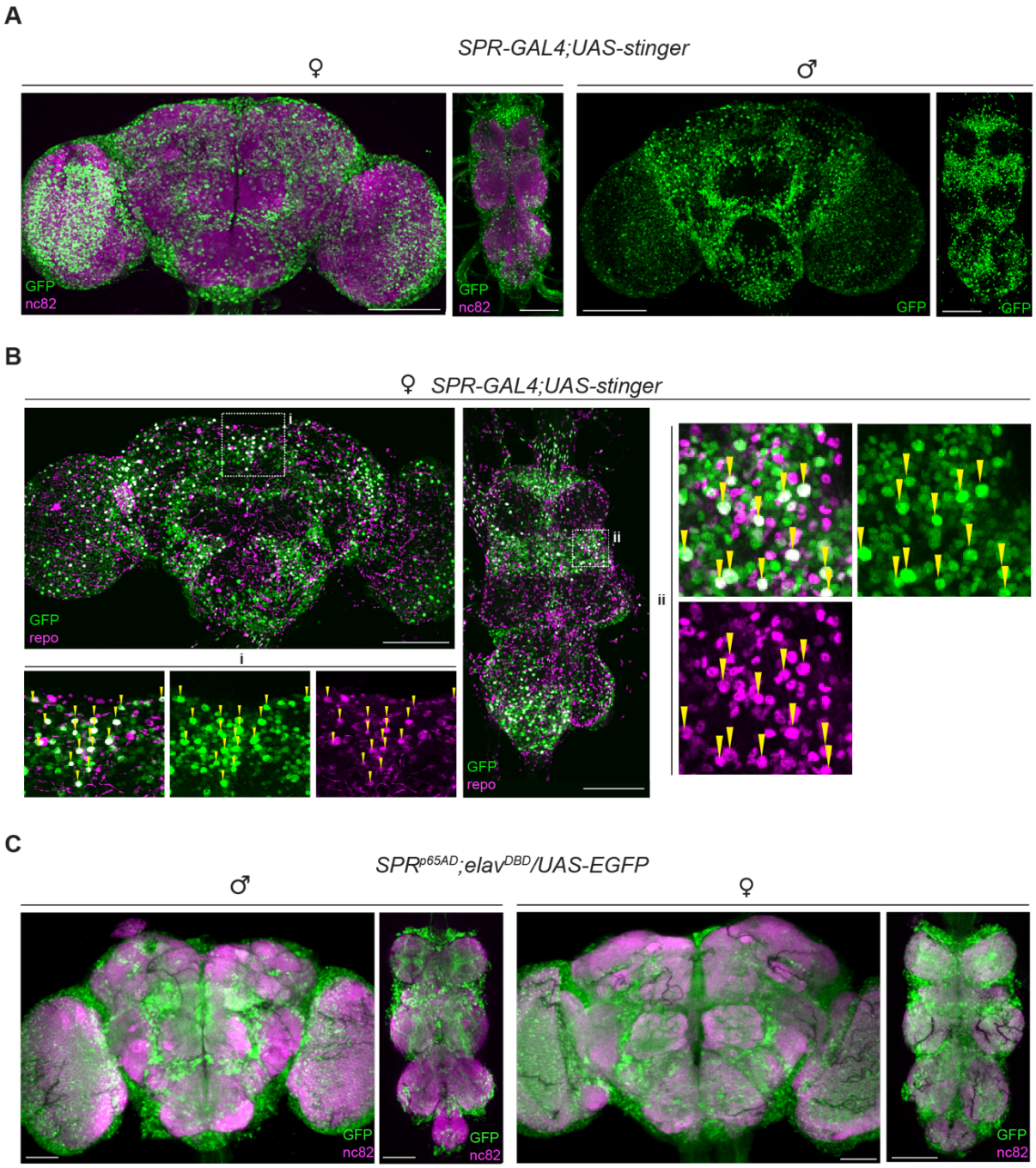


Figure 3. SPR expression in the CNS. (**A**) Maximal Z-projections, showing expression of nuclear GFP, driven by *SPR-GAL4*, in brain and VNC of adult females (left) and males (right). Female samplesstained with nc82, to mark the neuropil (magenta), and GFP (green). Male samples stained with GFP only. (**B**) Maximal Z-projections of adult female brain and VNC; same genotype as in (A), co-stained with GFP (green) and repo (glial nuclei; magenta), showing co-localisation of GFP and repo in the brain;panel (i) (below) and VNC; panel (ii) (right)(**C**) Neuronal expression of *SPR*. Maximal Z-projections, showing expression of enhanced GFP, driven by *SPR-p65AD* and *elav-DBD* hemidrivers, in brain and VNC of adult males (left) and females (right), stained with nc82, to mark the neuropil (magenta), and GFP (green). All scale bars = 50µm.

3.2.3.2 *SPR-GAL4[CRISPR]* expression in peripheral tissues

Outside the CNS, *SPR-GAL4[CRISPR]* expression was observed in the gut: in cells of the foregut, cardia, anterior midgut, and Malpighian tubules (fig.4A-C). *SPR* expression has not previously been described here, and further investigation will be required to elucidate the function of SPR in these cells. The expression and function of SPR in the female reproductive tract (FRT) has been explored in several studies investigating the mechanisms underlying the female post- mating response (PMR), whereby females undergo behavioural changes after mating. Early studies identified a component of the male seminal fluid, sex-peptide (SP), as the trigger of the PMR (Kubli, 2003) (Liu & Kubli, 2003) (Chapman et al., 2003), with SPR subsequently identified as the receptor through which SP acts (Yapici et al., 2008). Antisera provided evidence that, consistent with this ligand-receptor interaction, SPR is expressed in the FRT (Yapici et al., 2008). Subsequent studies revealed that the PMR is mediated by SPR signalling in just 3 bilateral sensory neurons in the FRT, which express *ppk*, a proprioceptive neuron gene, and both *fru* and *dsx* (Hasemeyer et al., 2009) (Yang et al., 2009) (Rezával et al., 2012). It was, therefore, of interest to examine *SPR-GAL4[CRISPR]* expression in the FRT, to compare with data from the use of antisera and, potentially, to provide direct evidence of SPR expression in the *ppk/fru/dsx* neuronal cluster. Using *SPR-GAL4[CRISPR]* to drive expression of UAS-stinger, GFP-labelled nuclei were observed in the lower region of the common oviduct, the spermathecae, seminal receptacle, uterus, and ovipositor (fig.4E-G), suggesting that SPR is expressed by many cells throughout the FRT. mGFP expression in these cells produced a seamless distribution of GFP throughout these regions (fig.4H-J), preventing the morphology of individual cells being distinguished. Further investigation is therefore needed to identify these cell types, and to determine the function of SPR in non-neuronal tissues of the FRT. By co-staining the reproductive tract with elav, a marker of neurons, the *ppk* sensory neuron cluster was identified. Within this bilateral cluster of 7-8 neurons, 3 were GFP-labelled in the *SPR- GAL4;UAS-stinger* line (fig.4F). Though it is already known that the PMR depends on SPR expression in 3 *ppk/fru/dsx* neurons, this finding

provides evidence that *SPR* is expressed exclusively in these 3 cells among the cluster. Interestingly, GFP expression was also observed in a small number of neurons innervating the lower part of the uterus. There is little information available as to the function of these neurons, and *SPR* expression has not previously been described here. It would be of interest to determine the function of these neurons, whether SP interacts with the SPR in this region and, if so, what the functional consequences of this interaction are.

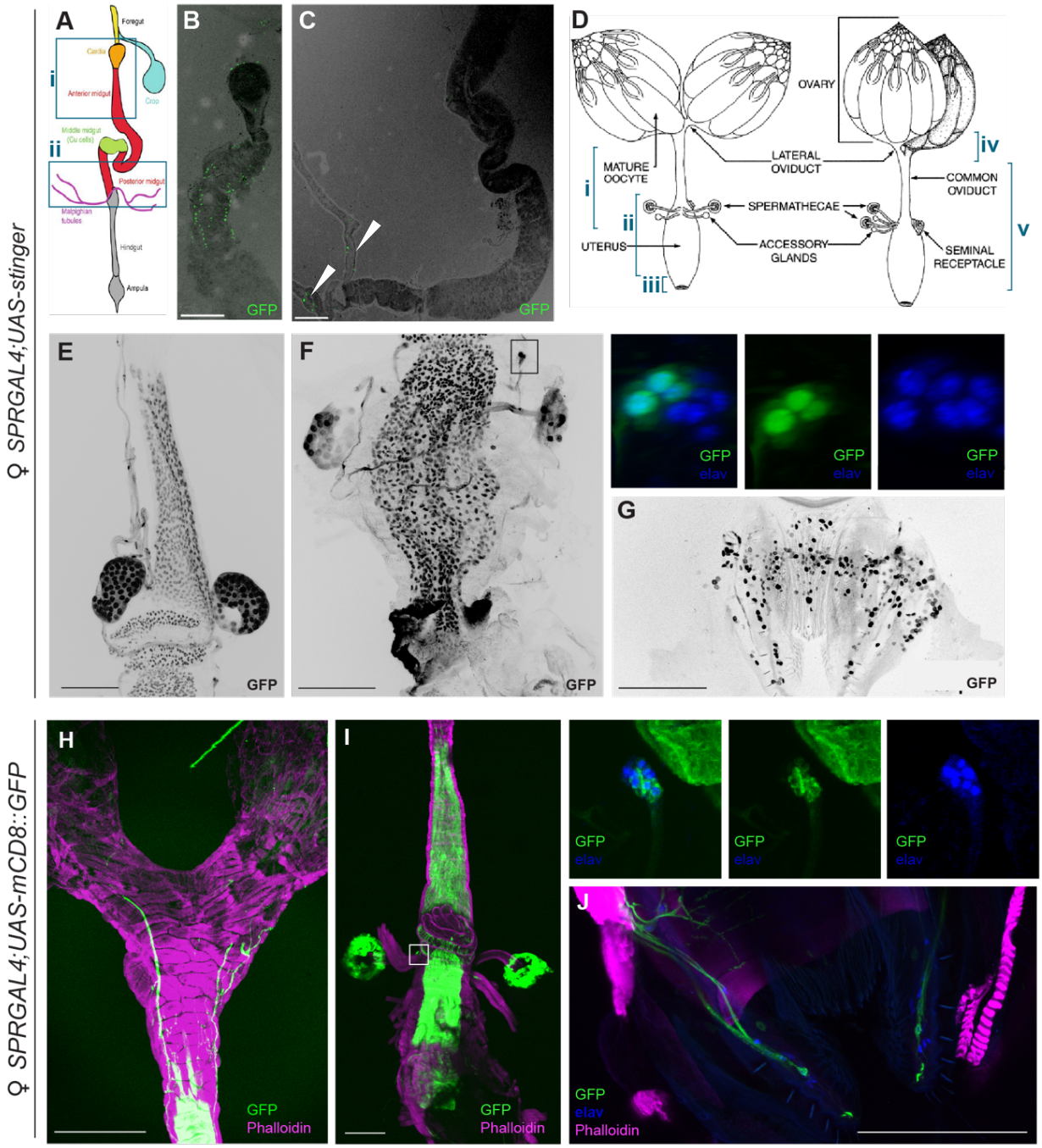


Figure 4: *SPR* expression in peripheral tissues. (**A**) Schematic of the *Drosophila* gut - adapted from Sadaqat et al.,2021. (**B**) Anterior region of gut (region (i) in (A)), from female expressing nGFP (green) driven by *SPR-GAL4*. (**C**) Posterior region of gut (region (ii) in (A)), from female of same genotype. White triangles indicate locations of GFP^{+} nuclei (**D**) Schematic of the female reproductive tract (FRT), adapted from Bloch-Qazi et al., 2003. (**E-G**) nGFP (black) driven by *SPR-GAL4* in the FRT; regions (i), (ii) & (iii) in (D), respectively. Right panel: co-localisation of GFP (green) and *elav* (blue) in sensory neurons (boxed region in (F)). (**H-J**) mGFP (green) driven by *SPR-GAL4*, in the FRT, co-stained with phalloidin (magenta) (H = region (iv) in (D), I = region (v) in (D), J = region (iii) in (D)). Right panel: expanded view of boxed region in (I), showing GFP (green) and *elav* (blue) staining. All confocal images are maximal Z-projections. All scale bars = 100μm.

3.2.3.3 Expression of neuropeptide receptor-MiMIC lines in the nervous system

Though the distribution of SIFamide has been described previously (Terhzaz et al., 2007) (Martelli et al., 2017), the expression pattern of the receptor remains uncertain. A *SIFaR-GAL4* line was produced, which drives reporter expression broadly in the nervous system (Sellami & Veenstra, 2015) (Wong et al., 2019). However, as conceded by the authors of the former study, a small promotor region of the *SIFaR* gene was used, raising a question mark over the extent to which this line recapitulates *SIFaR* expression. Combining the *SIFaR-LexA*$^{[MiMIC]}$ line with *LexAop-Tdtomato*, a red-fluorescent reporter that localises to the nucleus, revealed that, like *SPR, SIFaR* is expressed in a large number of cells distributed across much of the brain and VNC surface (fig.5A). This broad expression pattern was also observed when using the *SIFaR-p65AD*$^{[MiMIC]}$ line in conjunction with the ubiquitously expressed *tubulin-DBD* to drive GFP expression (fig.5A). To examine the neuronal expression of *SIFaR-p65AD*$^{[MiMIC]}$, the *elav-DBD* line was used, which expresses in all neurons. With this hemi-driver combination, GFP expression was evident in many cells throughout the CNS, suggesting broad neuronal expression of *SIFaR.* This was the case in males and females (fig.5B,C). Examination of peripheral tissues in this line revealed expression in neurons in all 3 sets of legs in both sexes (fig.5B,C), as well as innervation of the crop and ampulla regions of the gut. Expression was also seen in the reproductive tract of females, but not males. In the FRT, expression was restricted to a subset of the *ppk* sensory neuron cluster, as well as neurons innervating the vaginal plate. An identical expression pattern was observed when intersecting with the *tubulin-DBD* line, suggesting that, unlike *SPR, SIFaR* is expressed exclusively in neuronal cells in the FRT. Whilst the functional significance of SPR expression in FRT sensory neurons has been established, it is not clear why these cells might express *SIFaR*. It would be of interest to determine whether selective perturbation of SIFaR in these neurons affects aspects of female reproductive behaviour. In addition, while it is known that sex-peptide enters the FRT during copulation, it is unclear whether SIFamide is present in peripheral tissues. If it is, this is presumably due to neurohemal secretion of the peptide, since SIFamide neurons arborise

exclusively within the CNS (Terhzaz et al., 2007). It should be noted, however, that expression of *SIFaR* in the FRT and the legs is restricted to neurons that arborise within the VNC. Therefore, it is possible that SIFamide acts on these peripheral neurons by signalling through SIFaR expressed on their VNC neurites. Methods of detecting the cellular distribution of the SIFaR protein will be necessary to investigate these alternative hypotheses.

The *elav-DBD* hemi-driver line was used to characterise the neuronal expression pattern of the 3 remaining candidate NPR *split-GAL4* lines: *CrzR-p65AD*$^{[MiMIC]}$*, NPFR-p65AD*$^{[MiMIC]}$ *and Tkr86C-p65AD*$^{[MiMIC].}$ In all cases, GFP expression was seen in many neurons across the CNS, in both sexes (fig.6). This suggests that these receptors are expressed by large numbers of neurons across distinct functional compartments of the CNS.

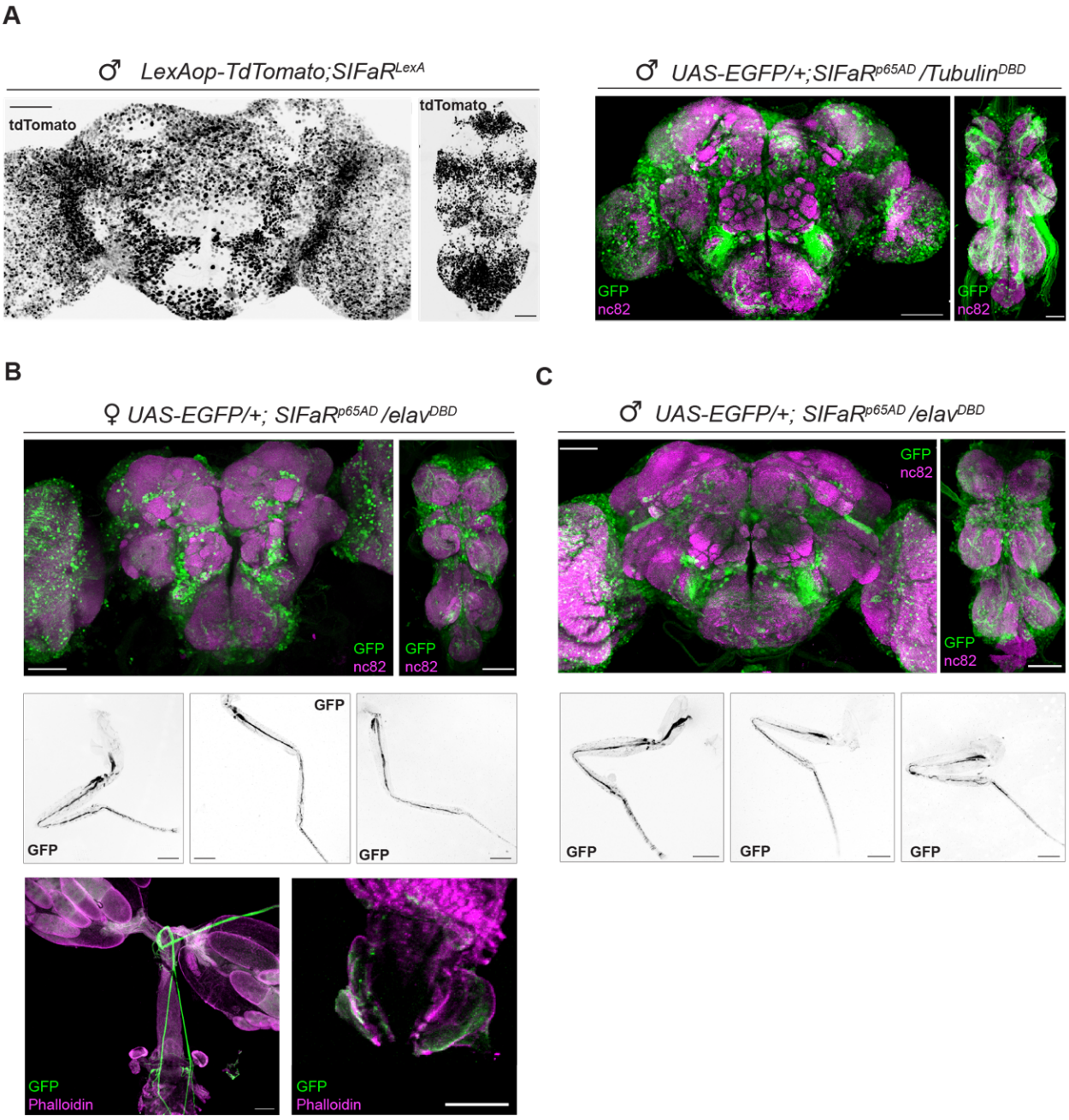


Figure 5: SIFaR expression in the nervous system. (**A**) Left: Male brain and VNC with nRFP (TdTomato; black) driven by *SIFaR-LexA*. Right: Male brain and VNC with EGFP driven by *SIFaR-p65AD* and *Tubulin-DBD* hemidrivers, stained with nc82 (magenta) and GFP (green). (**B**) Neuronal *SIFaR* expression in females. Top: female brain and VNC with E-GFP driven by *SIFaR-p65AD* and *elav-DBD* hemidrivers stained with nc82 (magenta) and GFP (green). Middle: foreleg (left), midleg (middle) and hindleg (right) from female of same genotype, with GFP in black. Bottom: Upper (left) and lower (right) reproductive tract from female of same genotype, stained with phalloidin (magenta) and GFP (green). (**C**) Neuronal *SIFaR* expression in males. Top: male brain and VNC with EGFP driven by *SIFaR-p65AD* and *elav-DBD* hemidrivers stained with nc82 (magenta) and GFP (green). Middle: foreleg (left), midleg (middle) and hindleg (right) from male of same genotype, with GFP in black. All images show maximal Z-projections. All scale bars = 50μm.

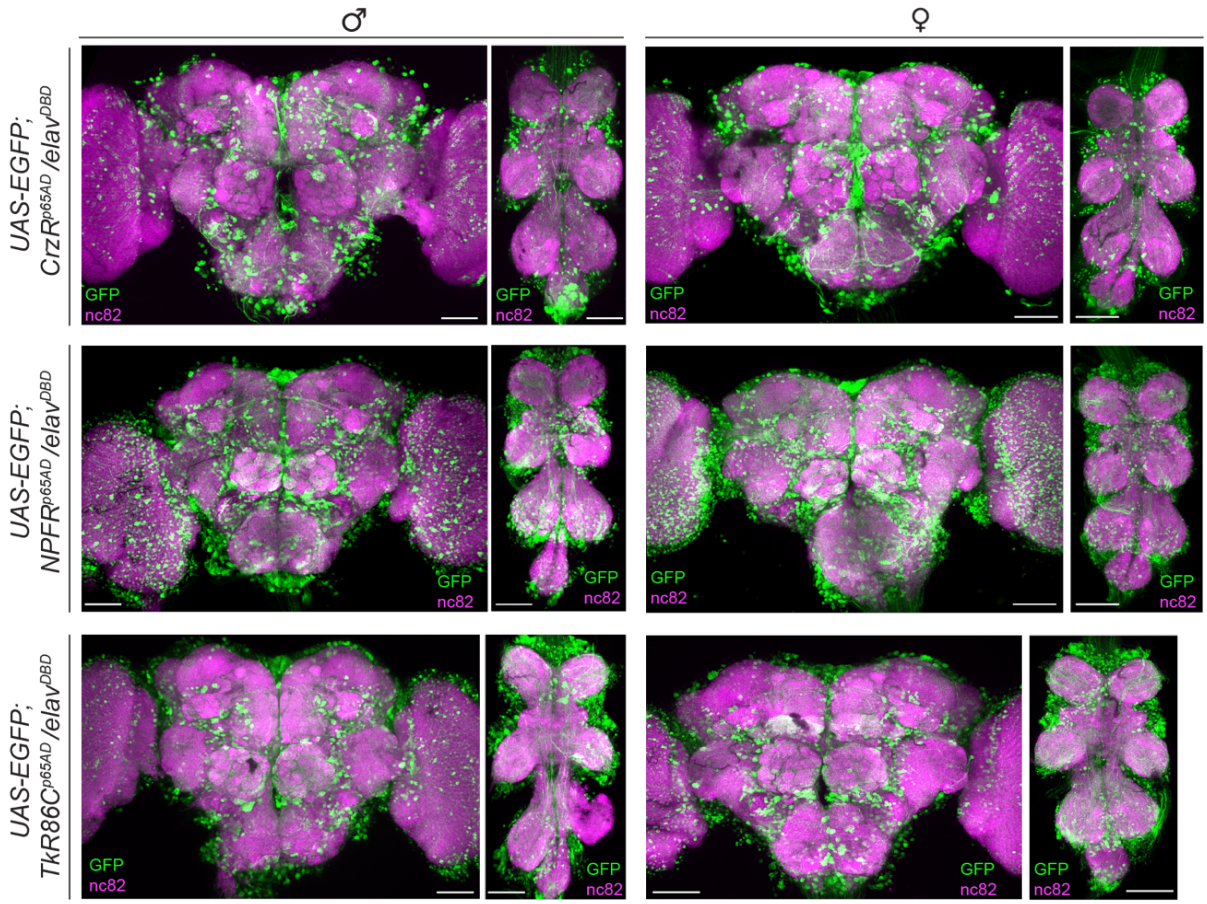


Figure 6. *CrzR, NPFR, and TkR86C*: neuronal expression in the CNS. Top panel: E-GFP driven by *CrzR-p65AD* and *elav-DBD* hemidrivers in brain and VNC of male (left) and female (right). Middle panel: E-GFP driven by *NPFR-p65AD* and *elav-DBD* hemidrivers in brain and VNC of male (left) and female (right). Bottom panel: E-GFP driven by *TkR86C-p65AD and elav-DBD* hemidrivers in brain and VNC of male (left) and female (right). All samples stained with nc82 (magenta) and GFP (green). All images show maximal Z-projections. All scale bars = 50µm.

3.3 Mapping neuropeptide receptor expression using single-cell RNA sequencing

Using single-cell RNA sequencing (scRNAseq) methods, the transcriptional profile of thousands of cells can be analysed in parallel, at single cell resolution. The Goodwin lab recently used scRNAseq to generate a 'transcriptomic atlas' of the adult VNS (Allen et al., 2019) and brain. The approach utilised the Chromium system from 10x Genomics, in which microfluidic technology is used to encapsulate single cells from a dissociated tissue sample

into individually barcoded nanolitre droplets. Attachment of this DNA barcode to transcripts from the cell ensures that the cellular origin of pooled transcripts is maintained (Macosko et al., 2015). The further addition of short unique molecular identifier (UMI) sequences during reverse transcription distinguishes copies of identical transcripts, allowing the expression level of each gene to be quantified in single cells. Using this approach, expression data was generated for 12,923 genes in 26,000 cells from the VNS, and for 13,377 genes in 70,552 cells from the brain. In the current study, these data sets were analysed to investigate the expression of the 5 candidate *NPRs* for which transgenic driver lines had been generated. Preliminary analysis was conducted to approximate the number of cells that express these genes (breadth of expression), and the number of transcripts present in these cells (expression level). *NPR* expression was then considered in the context of the broader transcriptional profile of expressing cells, providing an overview of *NPR* expression across cell types, identifying co-expressed genes of interest, and highlighting expression in selected cell types.

3.3.1 Quantifying neuropeptide receptor expression in the CNS

The expression patterns of the *NPR* driver lines generated in this study are indicative of broad *NPR* expression in the nervous system. It was of interest to determine whether this conclusion is supported by the scRNAseq VNC and brain data. For each candidate *NPR*, brain and VNC data sets were analysed to determine the number of cells from which the relevant transcripts were captured. For all 5 *NPRs*, transcripts were captured from cells numbering in the thousands across the CNS (Table 2). The cell-counts suggest that *SPR* expression is particularly broad, with transcripts captured from nearly a third of cells across both data sets (31,440 of 97,320 cells). Indeed, among ~25 *NPRs*, representing most identified *NPRs* in the genome, *SPR* transcripts were found in the greatest number of cells by far: more than double the number for the next most broadly expressed *NPR, FMRFaR (FMRFaR* = 13,456, *SPR* = 31,440). Published analysis of the VNS data identified a median of 5 GPCRs expressed per cell (Allen et al., 2019). In light of this, it was of interest to determine whether any cells

express more than one of these candidate *NPRs*. Analysis of co-expression revealed the existence of a number of cells, in both data sets, expressing 2 or more candidate *NPRs* (fig. 7). Further analysis of these cells may reveal how the expression of multiple *NPRs* relates to their function.

	Numbers of Cells		
	Brain	VNS	Total
Total	**70,552**	**26,000**	**96,552**
CrzR	3,128	412	3,540
NPFR	4,549	643	5,192
SIFaR	1,693	355	2,048
SPR	25,584	5,856	31,440
TkR86C	5,034	850	5,884

Table 2: Numbers of *NPR*-expressing cells in scRNAseq data sets.

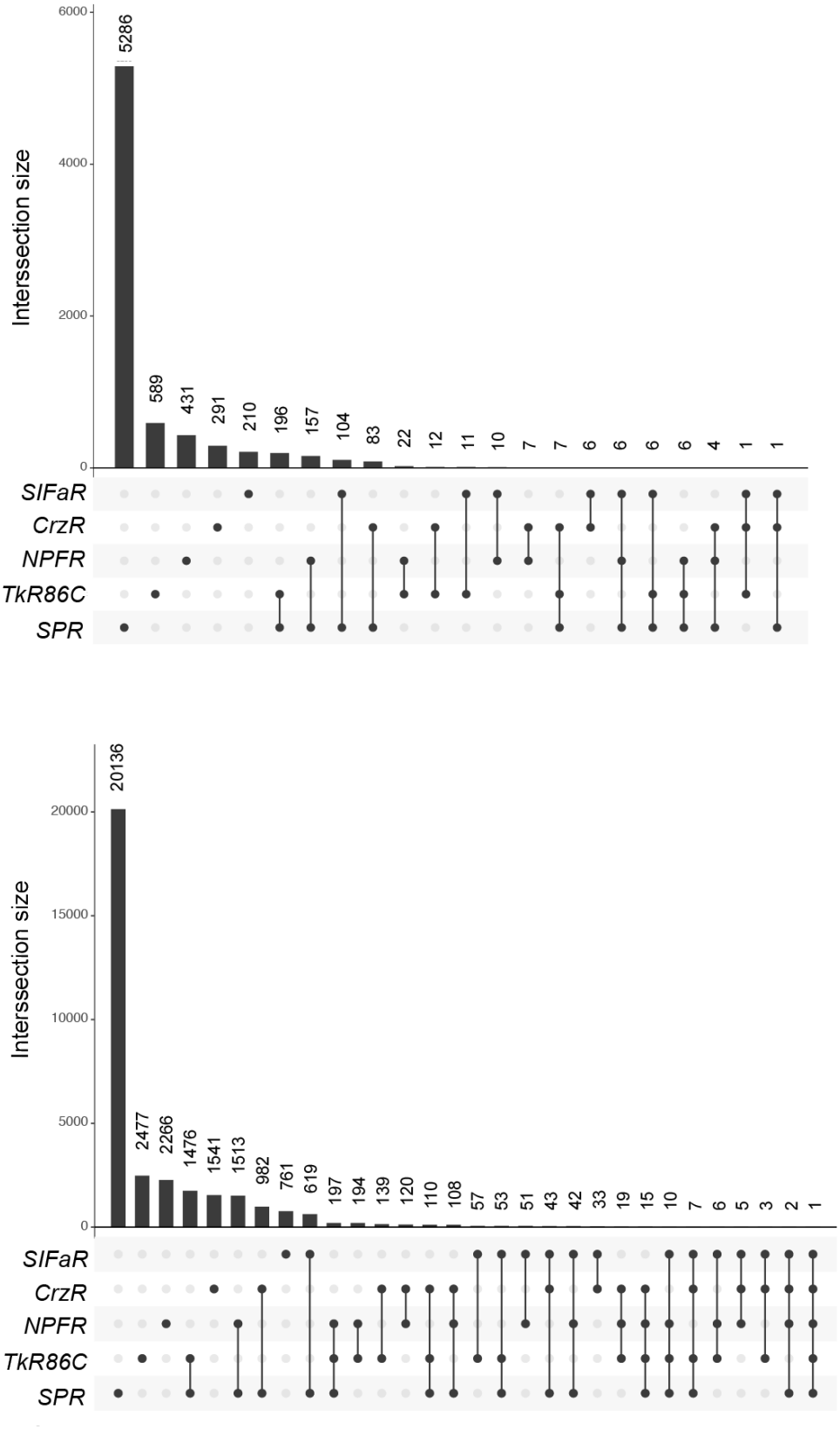


Figure 7: Co-expression of candidate *NPRs*. 'Upset' plots, showing co-expression of NPRs in cells of the VNC (top) and brain (bottom). Numbers above the bars give the total number of cells expressingthe combination of NPRs specified in the corresponding section of the panel below.

When interpreting this data, it should be noted that scRNAseq approaches invariably fail to capture all cell transcripts: a phenomenon known as dropout. For genes expressed at low levels, i.e., one or a few transcripts per cell, expression will be missed entirely in some cells due to dropout, ultimately leading to underestimation of the number of cells expressing that gene. Interestingly, genes expressed at very high levels, such as neuropeptides, are likely to be over-represented in scRNAseq data (Allen et al., 2019). Considering these points, it was of interest to determine the expression *level* of *NPRs*, not only to provide insights into their transcriptional regulation, but to ascertain how this may influence their representation in the data. To do this, non-normalised numbers of transcripts were analysed in cells containing at least one transcript of the gene. In the VNC, the maximum expression level of an average gene is 11 transcripts per cell, while for *CrzR, NPFR, SIFaR* and *TkR86C* the highest expression was between 3 and 5 transcripts (fig. 8). For each of these NPRs, the average expressing cell contained just 1 transcript (mean = 1.1 transcripts). Similarly low expression of these 4 *NPRs* was evident in the brain data, where the maximum expression level was between 6 and 8 transcripts, seen in only a small fraction of expressing cells, and most cells (mean = 86%) contained just one transcript (fig. 8). Transcripts of *SPR*, which the data suggests is expressed more broadly than other *NPR*s, were detected at much higher levels in some cells. For example, in the VNS, 64 cells contained more than 11 transcripts (the maximum of a 'typical' gene) and the highest expressing cell contained 58 (fig. 8). However, the vast majority (91%) contained 4 transcripts or fewer, and the mean expression level was 1.95 transcripts. Therefore, while a few cells contain large numbers of *SPR* transcripts, most express the gene at the low levels typical of most other *NPR*s. A similar distribution of *SPR* expression level was evident in the brain data (fig. 8). That NPRs appear to be expressed at relatively low levels is consistent with other lines of evidence that suggest GPCRs are abundant in the genome but sparsely represented in the transcriptome (Fredriksson & Schiöth, 2005). Crucially, the low expression of *NPR*s makes them more susceptible to dropout in scRNAseq, meaning the cell counts presented in this section may well underestimate the numbers of cells that express these *NPR*s in the CNS.

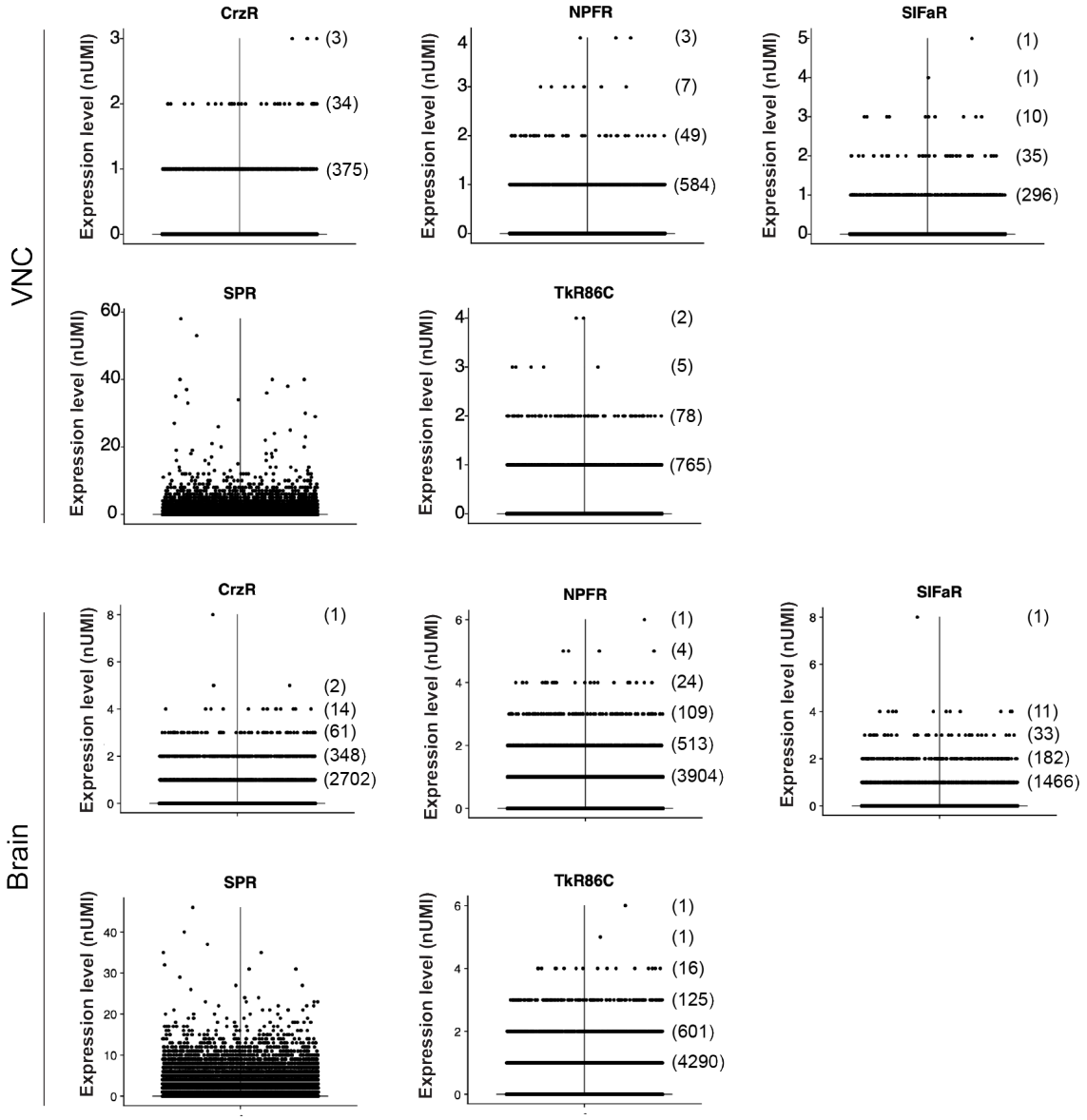


Figure 8: NPR expression levels in the nervous system. Violin plots, showing the number of transcripts (nUMI) in individual cells (dots) for each NPR gene, in the VNC (top) and brain (bottom) data. Numbers in brackets give the number of cells in each expression level group.

3.3.2 Expression of neuropeptide receptors across cell types in the CNS

By processing the raw data generated by a scRNAseq experiment, cells can be 'clustered' based on their transcriptomic profile, using dimensionality reduction tools, such as tSNE. For instance, processing of the VNS data set identified 119 transcriptionally distinct clusters (fig.

9)(Allen et a, 2019), while similar processing of the brain data set resolved 137 clusters. A cell's identity is defined by the compliment of genes it expresses and, as such, clusters in the scRNAseq data represent the range of cell types found in the tissue. To get an overview of the expression of candidate *NPR*s across cell types, data points (single cells) on tSNE plots of brain and VNS data were coloured according to expression of each *NPR* gene. Expressing cells can be seen distributed across most clusters, suggesting that *NPRs* are expressed in a range of cell types within the brain and VNC (figs.10 & 11, respectively). To quantify the breadth of *NPR* expression across cell types, the number of expressing clusters (clusters in which >1% of cells express the gene) was determined for each *NPR*. Candidate *NPR*s were expressed in an average of 83 (70%) clusters in the VNC (fig. 10) and 127 (93%) in the brain (fig.11). The most broadly expressed was *SPR;* found in all brain and VNC clusters, while *SIFaR* was expressed in smallest proportion of clusters in both data sets: 50 (42%) in the VNC and 105 (77%) in the brain. Plotting the proportion of expressing cells within each cluster revealed that, though candidate *NPR*s are expressed broadly, there is a degree of relative enrichment in certain clusters (fig.12). To determine the top 3 expressing clusters for each *NPR*, clusters were ranked by the proportion of expressing cells within the cluster (table 2). Further investigation into these enriched clusters may shed light on the types of cells that are particularly subject to modulation by these peptides. For example, in the VNC, *SIFaR* showed its highest enrichment in cluster 84, and several other *NPR*s were expressed by a relatively high proportion of cells in this cluster (compared to the average expressing cluster). Analysis of the transcriptomic profile of cluster 84 cells has revealed them to be neurosecretory cells that are specially adapted to release peptides (Allen et al, 2019). That the expression of *NPRs* appears prevalent in this cluster, provides evidence of an interaction between peptidergic signalling systems in the VNC. However, even in the highest expressing clusters, mos*t NPR*s were expressed in a relatively low proportion of cells: between ~10% and ~20% of cells in the top 3 VNC clusters, and between 20% and 35% in the brain. The same analysis of ~25 further *NPR*s annotated in the data sets revealed this to be true in most cases. This suggests that,

rather than defining cellular identity, expression of an *NPR* adds a level of specialisation to a range of cell types.

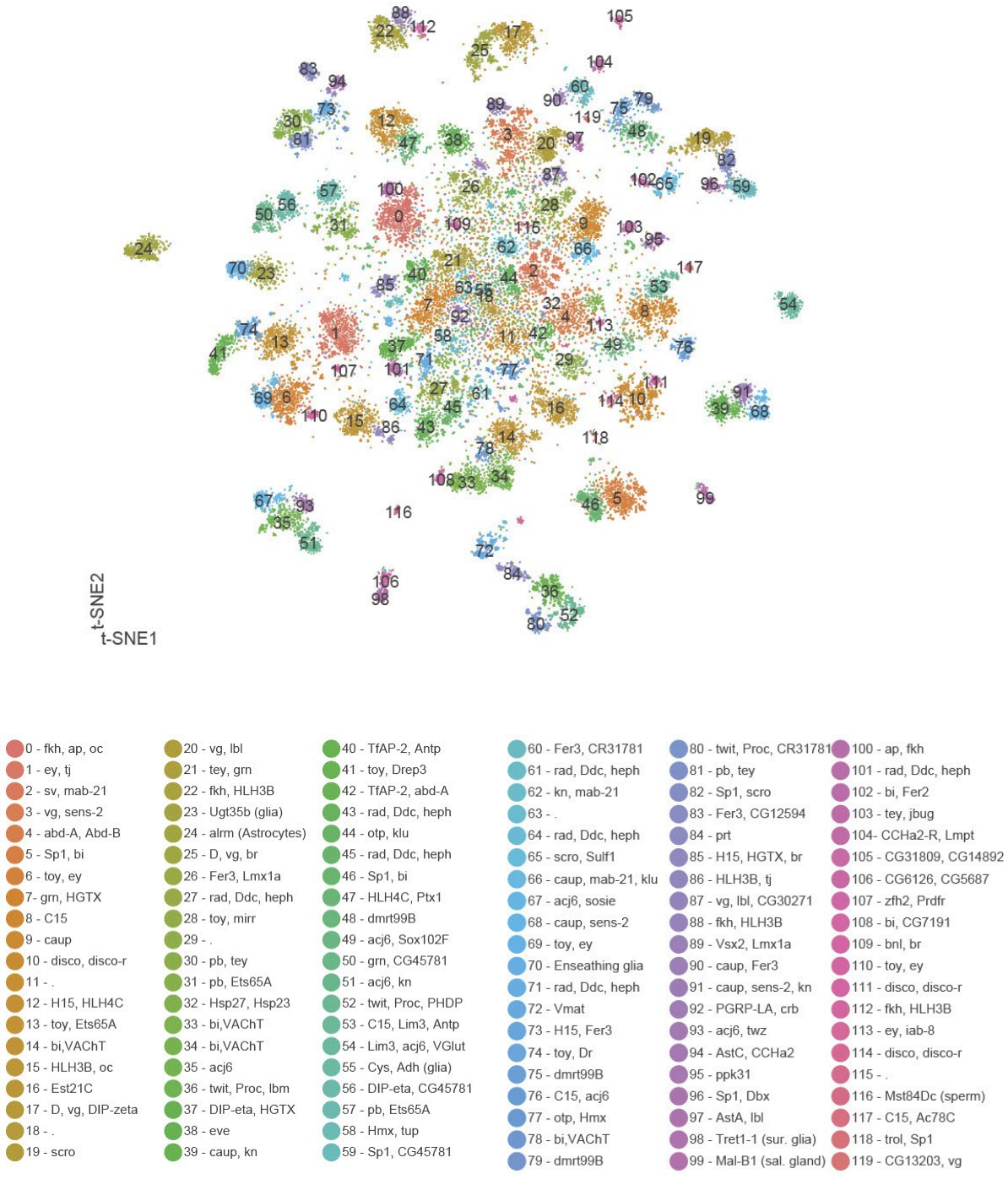
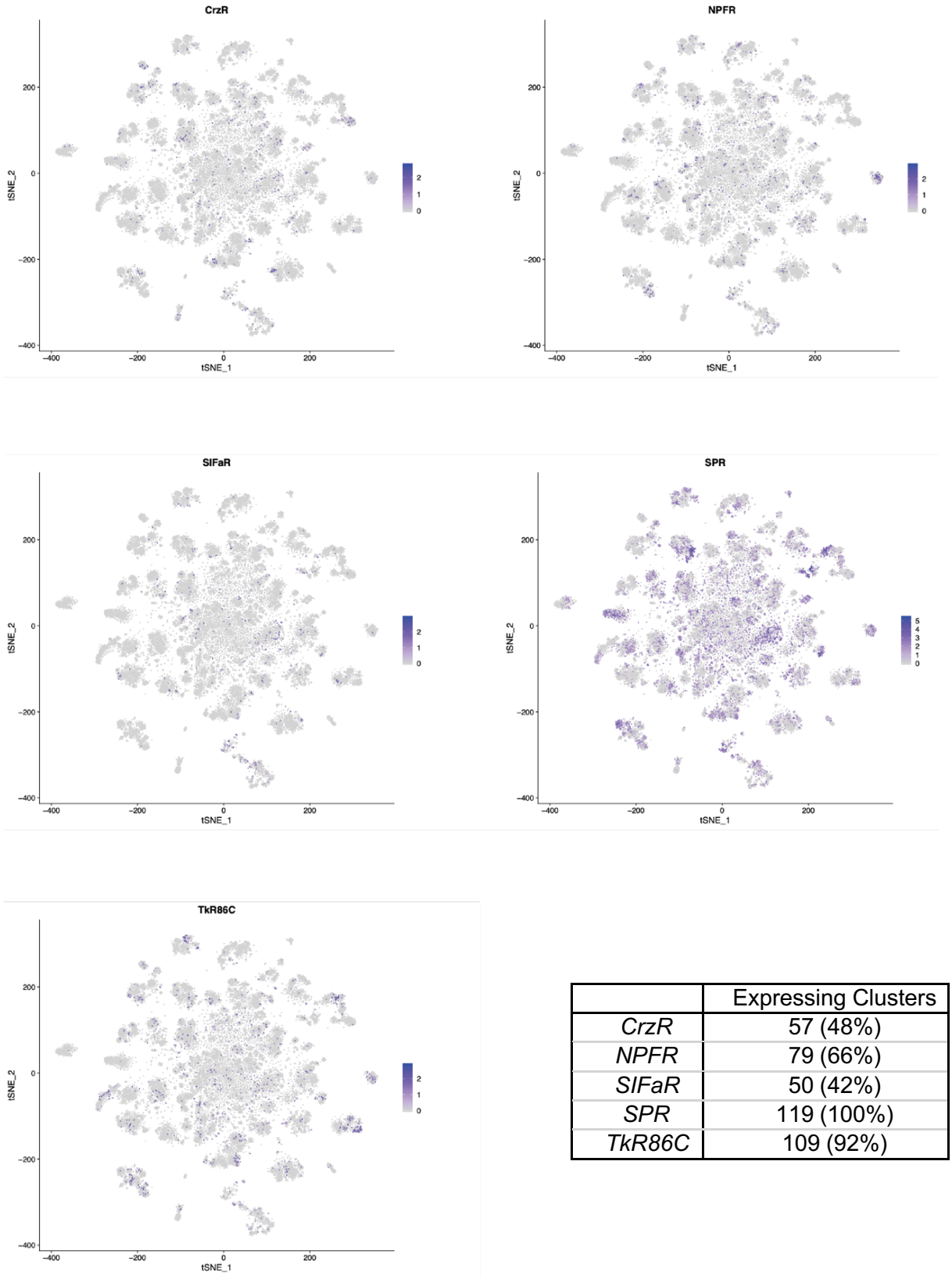
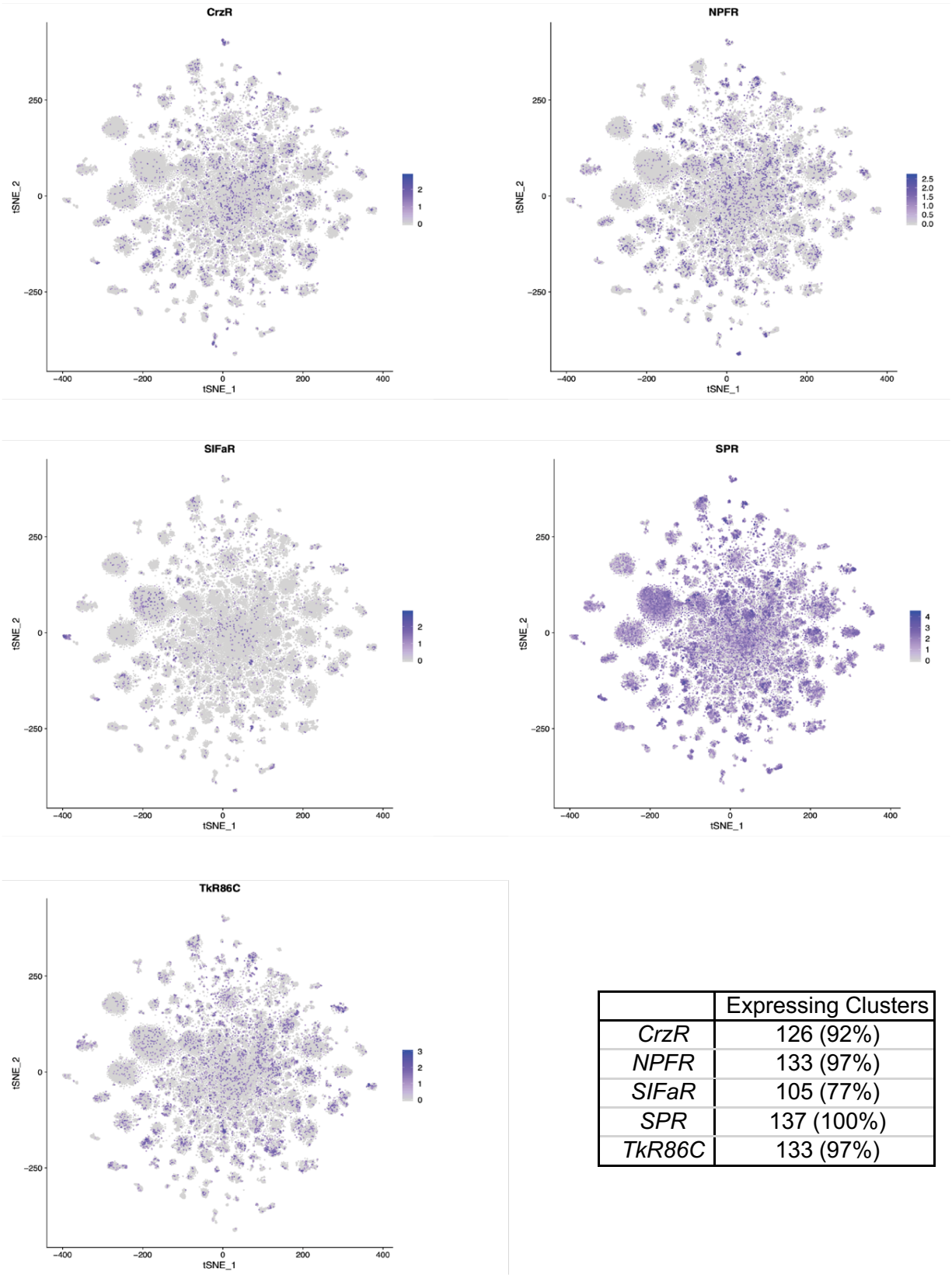


Figure 9: Clustering of cells in scRNAseq VNC data (Adapted from Allen et al., 2019). Top: 2-dimensional (t-SNE) plot of VNC cells, grouped into 119 clusters, with dots (cells) coloured according to cluster identity. Bottom: list of cluster marker genes.



	Expressing Clusters
CrzR	57 (48%)
NPFR	79 (66%)
SIFaR	50 (42%)
SPR	119 (100%)
TkR86C	109 (92%)

Figure 10: Distribution of *NPR*-expressing cells across transcriptionally distinct clusters in the VNC. t-SNE plot of VNC cells, with dots (cells) coloured, according to their expression of candidate *NPR*s, from light grey (no expression) to purple (highest expression level). Table (bottom right) gives the number of clusters in which more than 1% of cells express the *NPR* gene.



	Expressing Clusters
CrzR	126 (92%)
NPFR	133 (97%)
SIFaR	105 (77%)
SPR	137 (100%)
TkR86C	133 (97%)

Figure 11: Distribution of *NPR*-expressing cells across transcriptionally distinct clusters in the brain. t-SNE plot of brain cells, with dots (cells) coloured, according to their expression of candidate NPRs, from light grey (no expression) to purple (highest expression level). Table (bottom right) gives the number of clusters in which more than 1% of cells express the *NPR* gene.

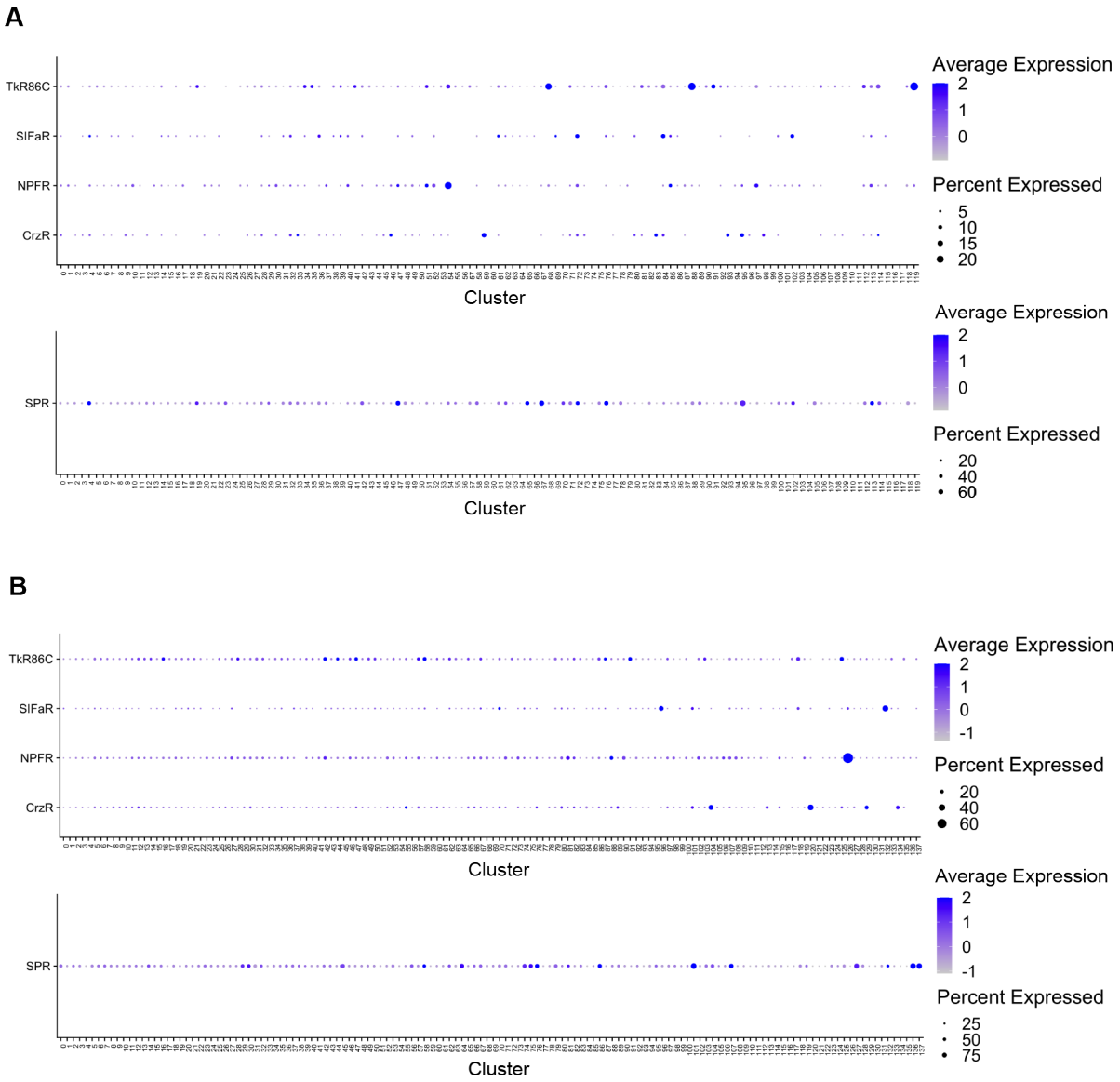


Fig 12: Candidate *NPR* cluster enrichment. (**A**) VNC data. Dotplot showing expression of NPRs across clusters. Size of the dots represents the percentage of cells in each cluster that expresses the NPR gene. Dot colour represents the average, scaled expression level in that cluster from light grey (lowest) to blue (highest) (**B**) Brain data showing same as in (A),

	Top 3 expressing clusters (VNC)	Top 3 expressing clusters (Brain)
CrzR	**59** (12%), **95** (10%), **83** (8%)	**120** (34%), **104** (29%), **129** (20%)
NPFR	**54** (20%), **97** (10%), **51** (9%)	**126** (66%), **81** (19%), **88** (19%)
SIFaR	**84** (11%), **72** (10%), **102** (10%)	**132** (35%), **96** (27%), **70** (12%)
SPR	**95** (70%), **67** (63%), **47** (56%)	**136** (90%), **137** (87%), **101** (92%)
TkR86C	**119** (23%), **88** (21%), **68** (18%)	**125** (23%), **91** (20%), **58** (20%)

Table 2: Top expressing clusters for candidate NPR genes in brain and VNC data. Cluster numbers in bold. Percentage of expressing cells in that cluster shown in brackets.

3.3.3 Expression of neuropeptide receptors in specific cell types

3.3.3.1 *NPR* expression in fast-acting neurotransmitter neuronal subtypes

Neurons can be grouped according to the type of FAN they release. In the *Drosophila* CNS, these are Ach, glutamate and GABA. By analysing the expression of genes implicated in FAN pathways, FAN identity was assigned to cells in the VNC dataset (Allen at al., 2019). Analysis of the expression of candidate NPRs in FAN-annotated cells suggests they are expressed by neurons utilising all 3 FAN types, with only small biases to a particular type in each case. Indeed, this was the case for ~25 additional *NPRs* (fig.13), suggesting that the cognate neuropeptides modulate similar proportions of cholinergic, glutamatergic and GABAergic neurons (fig. 13).

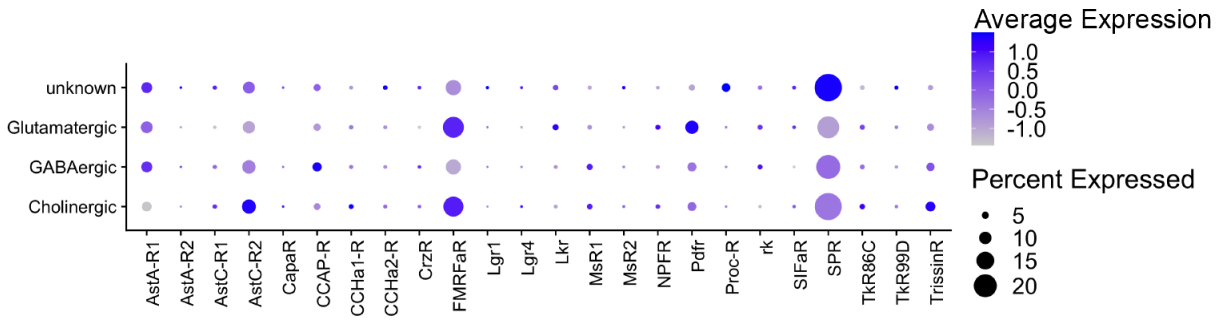


Figure 13: Expression of *NPRs* across FAN neuronal sub-types in the VNC. Dotplot, showing expression of NPR genes in cells belonging to each FAN sub-type. Size of the dots represents the percentage of cells of each sub-type that expresses the NPR gene. Dot colour represents the average, scaled expression level in those cells, from light grey (lowest) to blue (highest)

3.3.3.2 *SPR* expression in Glia

Several studies in mammals show that glial cells are modulated by neuropeptides (Van den Pol, 2012). In the current study, overlap between the expression of GFP and the glial marker *repo* in the *SPR-GAL4*$^{[CRISPR]}$ line suggested that *SPR* may be expressed in glial cells. Glial clusters were identified in the VNC data set by the expression of established glial cell markers. *SPR* transcripts were detected in ~24% of cells in these clusters, supporting the conclusion

that *SPR* is expressed by glia. Analysing the expression of *SPR* in sub-sets of glial cells, expressing markers for various glial sub-types, revealed *SPR* expression in 52 % of cortex glia, 36 % of ensheathing glia, 19 % of astrocytes and 1% of surface glia. Thus, *SPR* expression shows some bias towards certain glial sub-types, the functions of which could provide clues as to the role of non-neuronal SPR signalling; a process yet to be investigated.

3.4 Mapping neuropeptide receptor protein distribution

In revealing which cells express *NPR* genes, transgenic and scRNAseq approaches offer key insights into where neuropeptides act. However, they do not provide information about the sub-cellular distribution of these receptors. Given that neurons can arborise extensively and express receptors across cellular compartments, this information is essential in identifying the precise loci of NPR signalling. One way of visualising a protein is by tagging it. This relies upon a transgenic approach, whereby DNA encoding the tag is inserted into the coding sequence of an endogenous target gene. The tag is then incorporated into the gene transcript, giving rise to an internally tagged protein. For example, an intronic MiMIC can be exchanged for a construct encoding EGFP (Nagarkar-Jaiswal et al., 2015) (Bellen et al., 2015). A potential issue with internal tagging is that the tag may influence endogenous protein function. This is particularly the case when using fluorescent protein tags, due to their large size. EGFP, for instance, is ~27 KDa. Even if the internally tagged protein retains sufficient functionality (Nagarkar-Jaiswal et al., found this to be true of ~75% of proteins, when using the MiMIC based EGFP tagging system), small changes in structure could still affect its distribution. This consideration is especially pertinent for NPRs, which, as GPCRs, undergo complex folding, trafficking to disparate and, often, distant sub-cellular compartments and integration into cell membranes (Böhme & Beck-Sickinger, 2009). In this respect, the long-established method of detecting proteins in fixed tissues, using antibodies, has an advantage. With this in mind, an antibody was produced in the current study to investigate the distribution of SIFaR, the receptor for the neuropeptide SIFamide, the distribution of which remains elusive.

3.4.1 Generation of a novel antibody to investigate SIFaR distribution

There are significant challenges in producing antibodies that will bind selectively to GPCRs Firstly, once folded and integrated into the membrane, only a relatively small part of the protein is exposed. It is, therefore, important to analyse how GPCR primary structure relates to secondary and tertiary structural motifs. Secondly, GPCRs within a family, and even across families, can show significant sequence homology. As a result, antibodies with affinity for a target GPCR may also have affinity for other GPCRs present in the tissue (Michel et al., 2009). To minimise this risk, an epitope with as little homology to other known sequences as possible should be selected. Appropriate control experiments are still vital to rule out non-specific immunoreactivity, however (Saper et al., 2003, Hewitt et al., 2014).

This section describes the design of an anti-SIFaR antibody, based on analysis of its structure and homology to other proteins. It then provides an overview of the methodology used to produce the antibody and, finally, presents data from experiments designed to validate its affinity and specificity.

3.4.1.1 SIFaR structure analysis and epitope selection

SIFaR is a 758 amino acid (aa) rhodopsin-like 7-transmembrane (TM) GPCR (UniProt). Structural predictions for the sequence suggest that the N-terminal 104 aa's correspond to the extracellular domain of the receptor and the C-terminal 300 aa's to the intracellular domain, with the intervening sequence forming the 7-TM region (fig.14A,B). Using pBLAST ®, the sequence of SIFaR was compared with other *Drosophila* proteins in the refseq protein database. As expected, SIFaR shows a degree of homology with other *Drosophila* GPCRs, the most similar being the sequence encoding the RYamide receptor, with ~35 % sequence similarity. However, aligning SIFaR with the 100 sequences ranked highest for similarity reveals that this homology exists almost exclusively within the 7-TM region (fig.14C). Consistent with this, pBLAST ® analysis of the sequence corresponding to the extracellular domain identified no sequences with significant alignment. Based on this finding, the

extracellular domain was selected as the target region for antibody production.

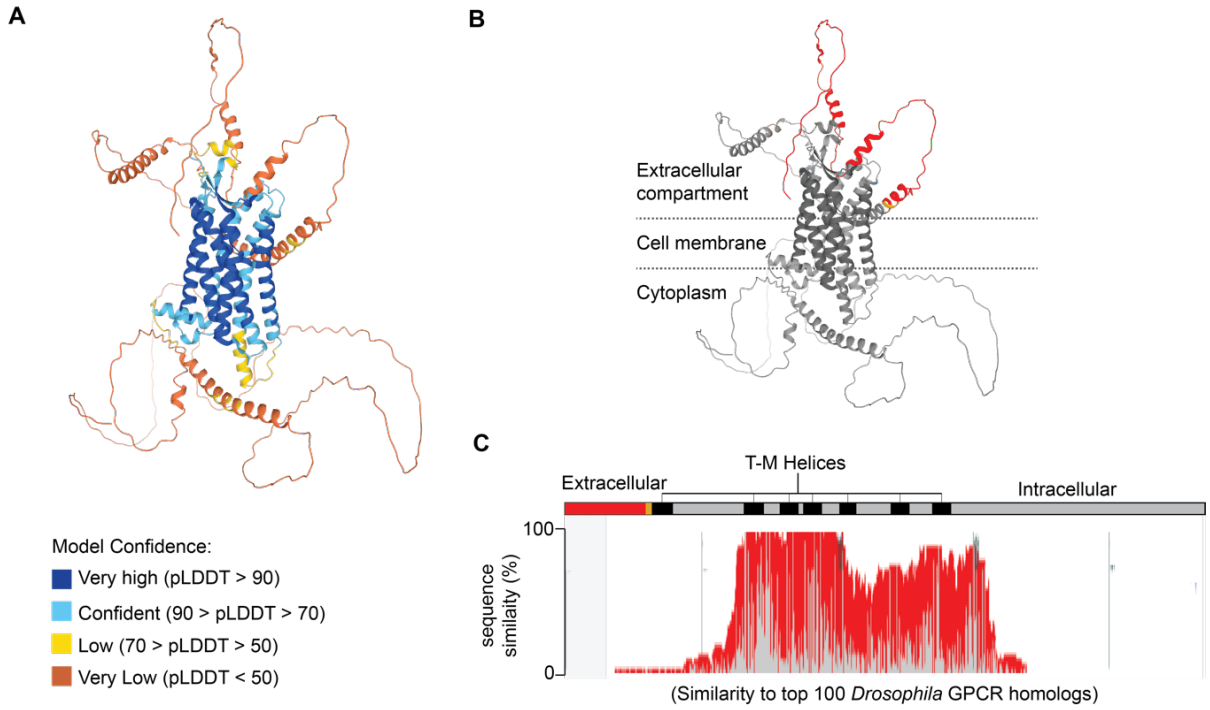


Figure 14: SIFaR protein structure analysis and epitope selection. (**A**) Model of SIFaR protein structure, based on AlphaFold web tool (Jumper et al., 2021), coloured according to model confidence. (B) AlphaFold SIFaR model, showing approximate position of protein in cell membrane. Coloured region (red) shows the 104 amino acid extracellular domain, selected as the epitope for antibody production. (C)Graph showing regions of homology to other *Drosophila* GPCRs.

3.4.1.2 Antibody production

Following selection of a target epitope, production of a rabbit polyclonal antibody was performed by the PTG lab (Proteintech®). In brief, 2 animals were immunised with a non-fixed fusion protein, consisting of the N terminal 104aa of SIFaR tagged with Glutathione S-transferase (GST). Affinity purification was carried out using a 6*HIS-fusion protein to produce purified antibodies from each animal, both with a concentration of ~150ug/ml. When comparing the immunoreactivity of the final bleed serum (pre-affinity purified antibody) in adult brain tissue with that of the pre-immune serum, using immunohistochemistry (IHC), a visible difference in fluorescence intensity was not apparent (figs.15A&B.) This suggests that molecules in the pre-immune serum contribute significantly to the detected signal, illustrating the importance of the affinity purification step.

3.4.1.3 Antibody validation in immunohistochemical assays

Detection of proteins in tissue can be achieved using various assays, two of the most common being IHC and Western Blotting. Since the primary purpose of generating a SIFaR antibody in the current study was to study the distribution of the receptor *in vivo*, the antibody was first validated using IHC. Based on scRNAseq data, and the expression pattern of *SIFaR* driver lines, SIFaR is expressed broadly in the adult brain. Therefore, antibody staining was analysed in this tissue as a positive control. Brains were incubated in an optimal concentration of anti-SIFaR (determined by serial dilution), followed by fluorescently conjugated secondary antibody. Fluorescence was evident throughout the brain, consistent with the predicted broad expression of the receptor (fig.15C). This was not the case in the absence of secondary antibody, ruling out the possibility of autofluorescence from the tissue or other reagents, or in the absence of primary antibody, confirming the specificity of the secondary antibody (figs.15 D,E). When analysing immunoreactivity more closely, in different regions of the tissue, two types of distribution were observed. On the anterior and posterior surfaces of the brain, where neuronal cell bodies are located, immunoreactivity was distributed throughout the cytoplasm of individual cells. In the neuropil region, where neurites are located, the distribution appeared more punctate (figs.15F,G). Both staining types have been described in the few studies in which anti-NPR antibodies have been tested via IHC in *Drosophila* and other insect nervous systems (Johnson et al., 2003; Bowser & Tobe, 2005; Johard et al., 2001). It should be noted that diffuse ‘background’ staining, which did not resemble either a cytoplasmic or punctate distribution, was evident throughout tissue, the presence of which could not be removed by modifications of the IHC protocol.

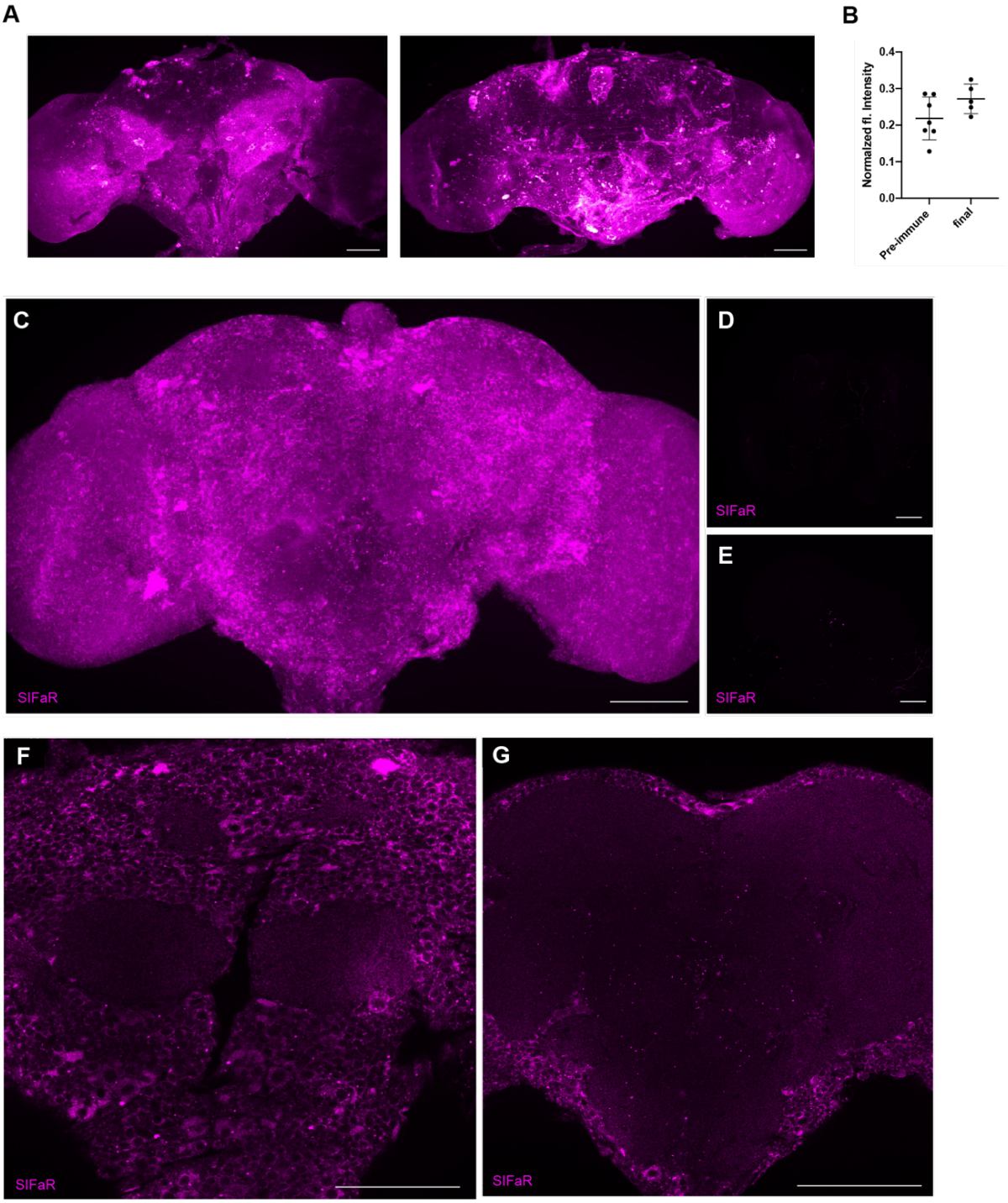


Figure 15: SIFaR antibody staining in the adult brain. (**A**) Maximal Z projections, showing staining after incubation in pre-immune serum; left (magenta) and final bleed serum; right (magenta). (**B**) Graph showing normalised fluorescence intensity of tissue samples in (A)(no statistically significant difference). (**C**) Maximal Z projection of anti-SIFaR antibody staining in the brain (magenta). (**D**) No secondary antibody control (**E**) No primary antibody control. (**F**) Single Z section from anterior surface of the brain, stained with anti-SIFaR (magenta) (example of 'cytoplasmic' staining pattern). (**G**) Single Z section from mid (neuropil) region of the brain, stained with anti-SIFaR (magenta) (example of 'punctate' staining in neuropil). All scale bars = 50μm.

Immunoreactivity in a positive control sample is an important indicator of antibody affinity, but it does not provide evidence of specificity. Specificity for the target epitope can be demonstrated using an absorption control experiment. Briefly, tissue in which immunoreactivity has been observed using the primary antibody is incubated in a solution containing the primary antibody, pre-incubated with the epitope used for immunisation. In this solution, binding sites for the epitope are inactivated, preventing interaction with the epitope in the tissue. After the addition of secondary antibodies and detection reagents, this should prevent staining, unless non- specific binding has occurred. Following this protocol, adult brains were incubated in anti-SIFaR, pre-absorbed with the GST fusion protein used for immunisation. Though staining intensity appears reduced in these samples (vs. those incubated in primary antibody alone) some staining was still evident (fig.16A). This was the case when using [10x] or [50x] of the GST-fusion protein for absorption. Staining was absent in samples incubated with GST fusion protein alone, ruling out the possibility that this staining reflects immunoreactivity of the GST fusion protein. Though a reduction in staining is consistent with some degree of specific binding, failure to abolish it entirely suggests the antibody may be binding to other molecules in the tissue.

Though an absorption control can provide evidence of antibody specificity for the target epitope, it does not demonstrate specificity for the target protein *in situ* (Hewitt et al, 2014). For this, negative controls are required. Ideally, immunoreactivity is analysed in tissue the protein is normally found in, from an animal in which its expression has been selectively abolished, for example, via perturbation of the gene. An issue pervading several aspects of the current study is that SIFaR appears to be required in early development, precluding the possibility of generating a viable *SIFaR* mutant. Though the *SIFaR* MiMIC insertion alleles generated in this study likely produce a truncated protein that does not function, predictions based on the location of the MiMIC insertion suggest the antibody target epitope is present within this truncated protein. Consistent with this, antibody staining did not differ in *SIFaR-p65AD*$^{[MI05376]}$ homozygotes versus controls. In the absence of a viable *SIFaR* mutant, staining was performed in samples expressing dsRNA against SIFaR. That this RNAi construct perturbs

SIFaR function is evidenced by the phenotype of males expressing it pan-neuronally, under the control of *elav-GAL4*, (see section 5.3.1), which closely mimics that associated with SIFamide perturbation (Terhzaz, 2007). However, SIFaR immunoreactivity was not abolished in the brains of *elavGAL4;;UAS-dsRNA(SIFaR*) flies (fig.16B). This suggests either that the RNAi construct does not reduce SIFaR levels sufficiently for detection by IHC, or that the antibody lacks specificity in IHC.

3.4.1.4 Antibody validation in western blot assays

Data produced from IHC assays suggested the SIFaR antibody may lack specificity when used in IHC, precluding its use in examining the distribution of anti-SIFaR *in situ*. Though a Western Blot (WB) assay cannot reveal protein distribution, it allows the expression level in different tissues or conditions to be compared. It was, therefore, of interest to determine whether the antibody might function better under the differing conditions of a WB. In a WB assay, proteins from the tissue are separated by electrophoresis, based on size, before incubation in primary antibody. Specificity of the antibody is, therefore, evidenced by the detection of a single 'band' with a molecular weight matching that of the target protein. For example, SIFaR has an estimated size of ~83 KDa, so an antibody specific to SIFaR should produce a single band of approximately this size. A WB assay was first performed on whole-adult protein extracts from wild-type animals (see chapter 2, Materials and Methods for full protocol). A clear band ~62 KDa was evident, suggesting strong affinity of the antibody for a protein of this size (fig.16C). Denaturation temperatures of 99°C (standard), 70°C and 37°C were tested, since lower temperatures can be preferable for the detection of GPCRs. Immunoreactivity was indistinguishable at the two higher temperatures, but no band was resolved when samples were denatured at 37°C (fig.16D). Very faint bands were present at other molecular weights; however, their signal intensity was negligible relative to the primary band, and they were observed only when using relatively high concentrations of the SIFaR antibody (1:50 or 1:100).

The difference between the size of the observed band and the predicted size of SIFaR could be due to features of the endogenous protein that affect its migration during electrophoresis, such as post-translational modifications, residue charges, or elements of secondary and tertiary structure not abolished entirely by denaturation. The alternative is that the antibody has strong affinity for a protein other than SIFaR. Assuming the latter to be true, it was possible this unintended target could be a protein expressed outside the brain. To test this hypothesis, protein extracts from whole animal and brain tissue were run side-by-side. The band at ~62 KDa was evident in both tissue samples (fig.16E), suggesting that, if this band corresponds to a protein other than SIFaR, it is a protein also expressed in the adult brain. Interestingly, the small degree of non-specific staining was slightly reduced in brain samples, suggesting the SIFaR antibody may have some non-specific affinity for a somatic protein.

As is the case in IHC, demonstrating antibody specificity in a WB assay requires a negative control. The *SIFaR*-MiMIC insertion alleles could not be used for this purpose in IHC because the truncated protein they produce contains the antibodies' target epitope. In a WB assay, however, this should be easily identified by a reduction in the size of the detected band. Based on the position of *Mi{MIC}SIFaRMI05376*, the truncated protein produced in *SIFaRp65AD*$^{[MI05376]}$ homozygotes should produce a band of ~35 KDa. However, the ~62 KDa band observed in wild type flies was also present in *SIFaRp65AD*$^{[MI05376]}$ homozygotes, with no reduction in signal intensity, and no band at a lower molecular weight present, suggesting that this band is produced by interaction of the antibody with an unintended target (fig.16F). To rule out the possibility that the *SIFaRp65AD*$^{[MI05376]}$ allele does not in fact produce a reduction in the level of full-length SIFaR, staining was also carried out on whole animal protein extracts from flies harbouring copies of the FLP-STOP(D) construct, inserted at *SIFaR*$^{[MI05376]}$, which was designed specifically to produce a strong loss of target gene expression (Fisher et al., 2017). There was no difference in the size or intensity of the band detected in these samples compared to controls (fig.16F). Taken together, these findings suggest strongly that the antibody produced in the current study does not bind specifically to SIFaR in a WB assay.

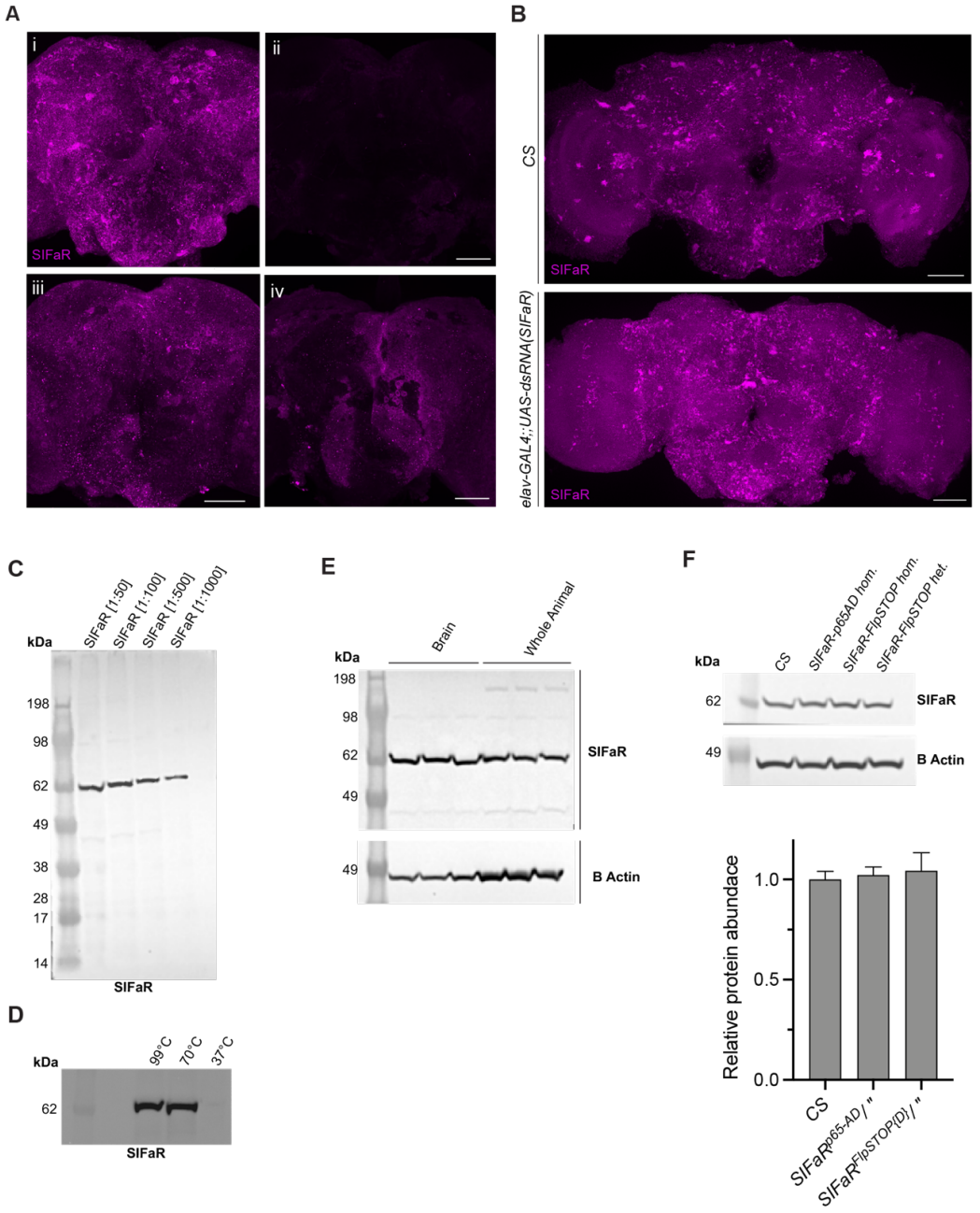


Figure 16: Validation of SIFaR antibody in IHC and WB assays (**A**) Absorption control assay. Max. Z-projections of brains incubated in SIFaR antibody (non-absorbed) (magenta) (i), GST-fusion protein (magenta) (ii), SIFaR antibody pre-absorbed with [10x] GST-fusion protein (magenta) (iii), SIFaR antibody pre-absorbed with [50x] GST-fusion protein (magenta) (iv). (**B**) Negative control assay: Max. Z projections of brains incubated in SIFaR antibody (magenta) from wild type animal (top) and animal expressing RNAi against SIFaR pan-neuronally (bottom). (**C-F**) Western Blot (WB) assays: (**C**) anti-SIFaR immunoreactivity at different target concentrations (shown above blot) (**D**) Comparison of anti-SIFaR immunoreactivity using different protein denaturation temperatures (shown above blot). (**E**) anti-SIFaR and anti-B-Actin immunoreactivity using brain and whole animal protein extracts. (**F**) Negative control assay: anti- SIFaR and B-Actin immunoreactivity using whole animal protein extracts from wild-type (CS) and SIFaR mutant animals (*SIFaRp65AD*$^{[MiMIC]}$ *SIFaRFlpSTOP*$^{[MiMIC]}$ homozygotes). Graph shows normalised fluorescence intensity of primary band relative to control (CS) group

3.5 Summary and discussion

Knowledge of the distribution of neuropeptide receptors (NPRs) in the nervous system represents a significant gap in our understanding of peptidergic signalling systems (Nässel & Zandalawa, 2019). For this reason, transgenic lines were generated in the current study to map the expression of selected *NPRs: CrzR, NPFR, SIFaR, SPR and TkR86C.* Though SPR-*GAL4* and *SIFaR-LexA* lines were produced, application of the split-GAL4 system was the primary focus, with split-GAL4 lines made for each of these receptors. Targeted insertional methods (CRISPR-Cas9 and the Trojan MiMIC system) were used to recapitulate, as fully as possible, the expression of the target genes. The general distribution of *NPRs* was established using ubiquitously expressed hemi-drivers: *elav-DBD*; to assess neuronal expression, and *tubulin-DBD*; to investigate expression across cell types*. NPR∩elav*-driven reporter expression suggests these *NPRs* are expressed in large numbers of neurons distributed throughout functional compartments of the CNS. This lends support to the idea that these neuropeptides modulate disparate functions via signalling in the discrete neural circuits subserving these functions. In the case of *SPR* and *SIFaR*, expression was also evidenced in neurons in the periphery. This could reflect signalling that occurs in peripheral tissues: SP-SPR interaction is known to occur in the female reproductive tract, for instance. Alternatively, the receptor proteins could localise to the centrally located neurites of these peripheral neurons. This could well be the case for SIFaR, since there is currently no evidence for any action of SIFamide outside the CNS. In addition, publicly available scRNAseq data indicates *SIFaR* is much more prevalent in the CNS than in peripheral tissues (http://scope.aertslab.org). However, whether SIFamide is secreted into the haemolymph has not been investigated.

Protein tagging and/or effective antibodies for these receptors are needed to determine the sub-cellular localisation of these receptors in peripheral neurons. Another interesting finding was the prevalence of *SPR* in CNS glia, as revealed by co-staining with glial cell markers. Though glial neuropeptide signalling has been investigated to some degree in vertebrate systems (Van de Pol, 2012), and neuropeptides have been identified in glia in *Drosophila*

(Sousa-Nunes et al., 2011), the functional significance of glial NPR expression and signalling in *Drosophila* remains largely unexplored.

To complement transgenic approaches, and provide additional insights into *NPR* expression, analysis of scRNAseq data generated in the Goodwin lab was conducted. The results of this analysis broadly support the data obtained via transgenic approaches: *NPR* transcripts were detected in significant numbers of cells, across neuronal cell types. That the nervous system tissue used in scRNAseq experiments was derived from 5-day old flies, further suggests that *NPR* transcription persists in large number of cells in the adult nervous system. This supports an ongoing role for signalling through each of these receptors in disparate neuronal populations. However, the possibility that post-transcriptional regulation of these receptors is employed to restrict 'effective' expression to a smaller population of cells cannot be ruled out. The scRNAseq data was further able to reveal the expression level of these receptors in expressing cells. Consistent with what has been hypothesised, *NPR*s appear to be expressed at low levels (just a few transcripts per cell), and there is only modest heterogeneity across cells in terms of this expression level. Again, the extent to which these transcripts are translated to produce functional proteins in different cell types is unclear. Nonetheless, the finding suggests that signalling specificity is not achieved by gross differences in the level of transcription of these NPRs across cell types.

To address questions beyond the scope of these methods, for instance the sub-cellular distribution of NPR proteins, and the potential for post-transcriptional regulation, a novel antibody for SIFaR was produced. Despite promising sequence specificity of the epitope, a thorough analysis of the antibodies' performance in IHC and WB assays suggested a lack of specificity for SIFaR. Thus, alternative methods of visualising the distribution of this protein should be explored, for instance gene tagging strategies (Kanca et al., 2017).

Chapter 4: Expression of Neuropeptide Receptors in the Neural Circuits of Reproductive Behaviour

4.1 Introduction

Evidence acquired thus far in the current study, suggests that the receptors for neuropeptides implicated in the modulation of reproductive behaviour are expressed broadly in the nervous system. In uncovering the mechanisms by which these neuropeptides modulate reproductive behaviour specifically, a key question is whether their receptors are expressed in the neural circuits governing these behaviours. Currently, there is no direct evidence of this. An absence of such expression would suggest that reproductive behaviours are influenced only indirectly by the neuropeptide in question, whereas expression of the *NPR* would provide evidence of a direct modulatory mechanism. Since *dsx* neurons form a more restricted neuronal population important for reproductive behaviour in both sexes (fig1.A), expression of *NPR*s in these cells was the primary focus. However, an examination of expression in fru neurons was also conducted in some cases.

The *NPR* driver and hemi-driver lines generated in the current study can be combined with other reagents to analyse their expression in sub-sets of cells. In the case of the *SPR-GAL4* and *SIFaR-LexA* driver lines, this intersection can be achieved via the use of an additional binary expression system in the same line (for example, GAL4/UAS used in conjunction with the LexA/LexAop system) or, in the case of GAL4/UAS, by expressing: i) Flp recombinase (Flp) under the control of a separate regulatory element, and ii) a UAS-responder element with an intervening stop cassette that requires removal by Flp. Genetic intersection using the *split-GAL4* hemidriver lines, generated for all candidate NPRs, is more straightforward: the expression of these *p65AD* lines in *dsx* and *fru* cells could be examined using *dsx-DBD* and *fru-DBD* lines, respectively (fig.1B).

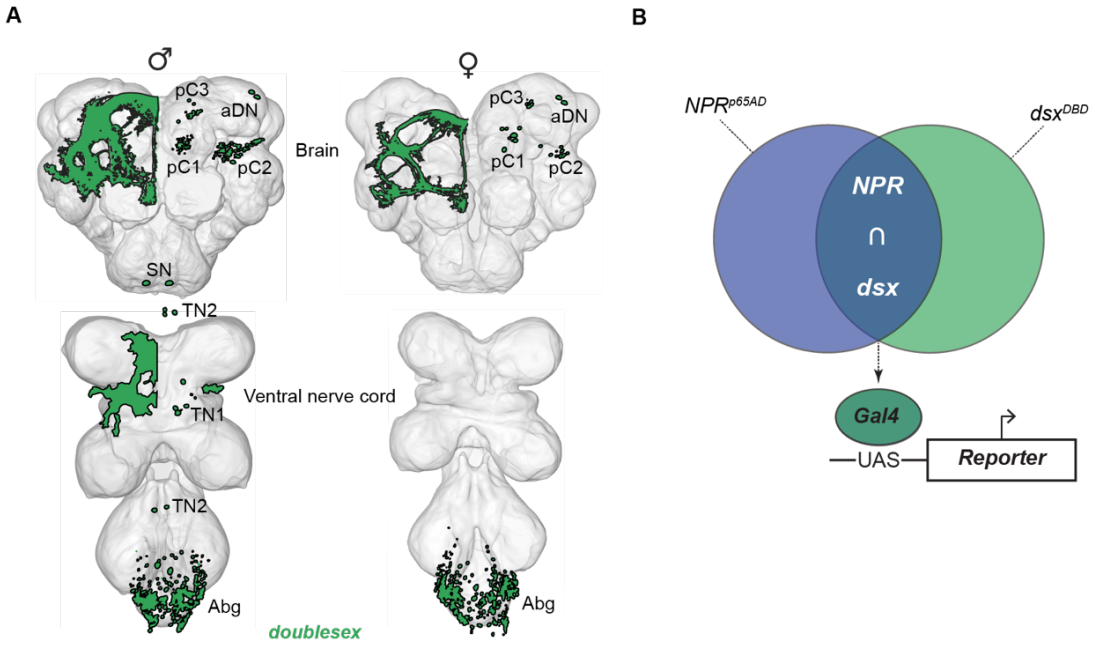


Figure 1: Intersecting *dsx* neurons in the nervous system. (A) Schematic showing *dsx* neural circuitry in the adult male (left) and female (right) CNS (green). Soma of *dsx* neurons in anatomically defined clusters (labelled) are shown on the right hemisphere, neurites on the left hemisphere. **(B)** Schematic illustrating split-GAL4 strategy for intersecting *NPR-p65AD*∩*dsx-DBD* cells and driving the expression of a fluorescent reporter to label those cells.

4.2.1 *SIFaR*

4.2.1.1 *SIFaR* expression in *dsx* cells.

To determine whether *SIFaR* is expressed in *dsx* cells in the nervous system, *SIFaR-p65AD* and *dsx-DBD* (Pavlou et al., 2016) lines were combined to drive expression of enhanced GFP (EGFP). In males, cells in the brain and VNC were GFP-labelled, the positions and morphologies of which are consistent with *dsx* neurons (fig.2A). Using cell count data from the *dsx-GAL4* line (Rideout et al., 2007) for reference, these GFP-labelled cells constituted ~28% of *dsx* neurons in the brain and ~30% of those in the VNC (table 1). Approximately 57 GFP-positive cells were observed in the tarsus of the foreleg in males (fig.2B), where there are thought to be ~72 *dsx*-expressing sensory neurons (Rideout et al., 2010). A significant proportion of these *dsx* neurons are thought to be gustatory receptor neurons (GRNs), which are involved in the detection of nonvolatile pheromones that guide male courtship (Rideout et al., 2010). GFP-labelled cells were also present in the CNS of females (fig.2C), constituting

~82% of *dsx* neurons in the female brain and 35% of those in the VNC (Table 1). In addition, approximately 30 cells were GFP-labelled in the foreleg tarsi of females (fig.2D): ~50% of the *dsx* neuronal population found here (Table 1). Less is known about the function of *dsx* sensory neurons in foreleg tarsus in females. However, female receptivity to courtship is thought to be regulated by a population of GRNs that respond to non-volatile pheromones.

The *SIFaR-p65AD* and *elav-DBD* intersection had already revealed *SIFaR* expression in the bilateral sensory neuron clusters adjacent to the spermathecae in the female reproductive tract (FRT). These neurons were previously shown to be dsx^{+} (Rezával et al., 2012). Therefore, it was of interest to determine whether these cells would be labelled *SIFaR-p65AD∩dsx-DBD* line. In this line, all ~11 neurons in the bilateral anterior and posterior clusters were GFP-labelled (fig.2F) (table 1), suggesting that *SIFaR* is expressed in all dsx^{+} neurons in these clusters. As mentioned previously, the function of *SIFaR* expression in these cells is unknown. Cells that appear to be neurons innervating the vulva (the opening of the FRT) were also GFP-labelled in the *SIFaR-p65AD∩dsx-DBD* line (fig.2F). Neither *dsx* nor *SIFaR* expression has previously been reported here. Taken together, these findings suggest tha*t SIFaR* is expressed in neurons at all levels of *dsx* circuit architecture in both sexes, both in the periphery and in all *dsx* clusters in the CNS.

SIFaR expression in *dsx* neurons was also investigated by combining *SIFaR-LexA* and *dsx-GAL4* in a single line to drive the expression of *UAS-TdTomato* (nRFP) and *UAS-stinger* (nGFP), respectively. In both sexes, overlap of nRFP and nGFP expression was evident in cells in the brain and VNC in (fig.3A-C), lending further support to the conclusion that *SIFaR* is expressed in *dsx* neurons in the CNS.

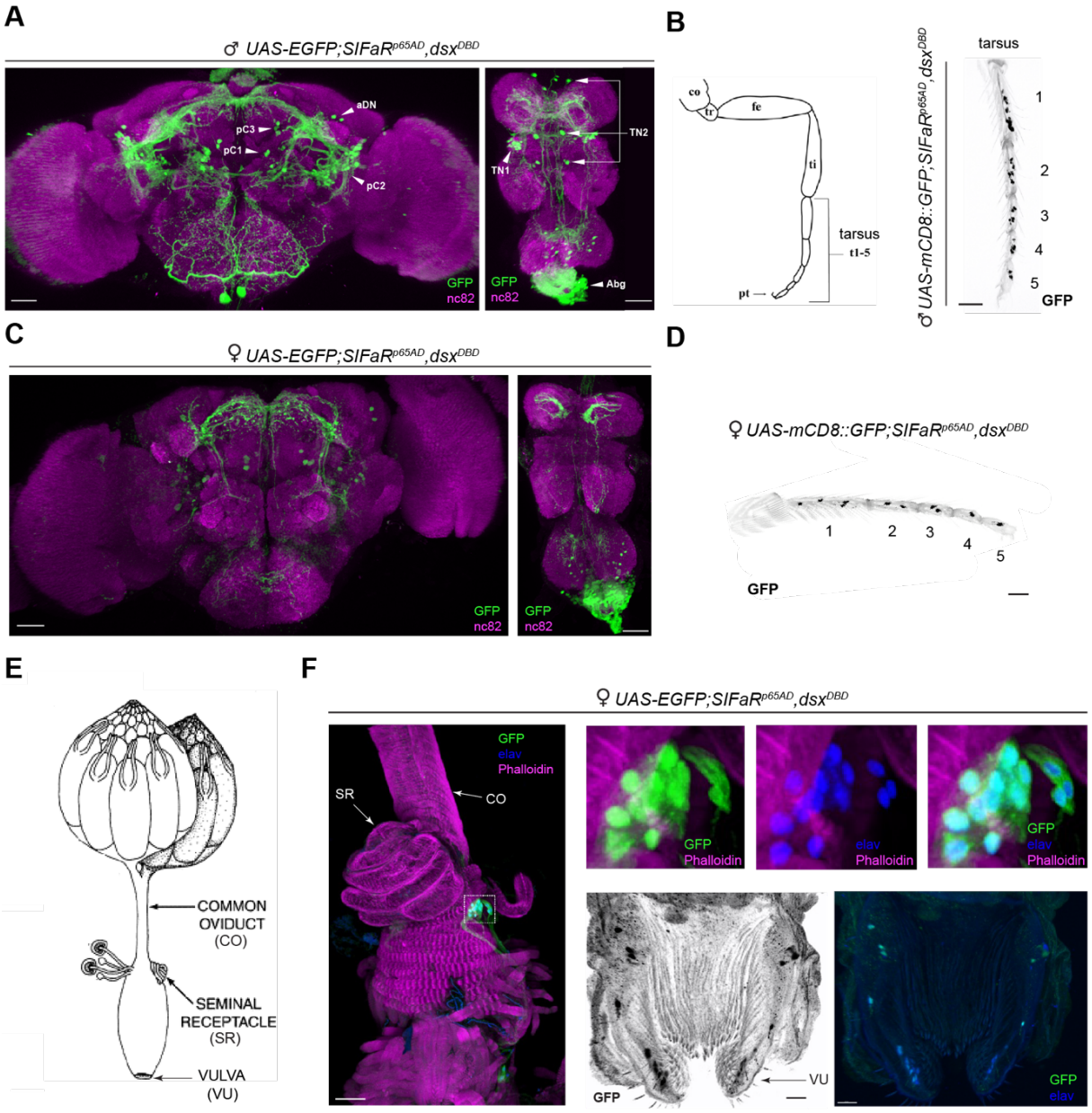


Figure 2: *SIFaR∩dsx* expression in CNS and PNS (A) Male brain (left) and VNC (right), stained with nc82 (magenta), showing the neuropil structure, and GFP (green), showing intersected *SIFaR∩dsx* cells. *SIFaR∩dsx* cells are evident throughout *dsx* CNS clusters (annotated). **(B)** Left: schematic of male foreleg (adapted from Sewell, 2006). Right: *SIFaR∩dsx* driven GFP expression (black) in the tarsus **(C)** Female brain (left) and VNC (right) showing *SIFaR∩dsx* driven GFP expression (green) **(D)** Female foreleg tarsus with *SIFaR∩dsx* driven GFP expression (black) **(E)** Schematic of female reproductive tract (FRT), adapted from Bloch Qazi et al., 2003. **(F)** *SIFaR∩dsx* driven GFP expression in the FRT. Panel; top right: close-up of boxed region, showing GFP overlap with *elav,* a neuronal marker. Bottom right: *SIFaR∩dsx* driven GFP expression in cells around the Vulva. All confocal images are maximal projections of z-stacks. Scale bars in the CNS and FRT = 50μm, in the legs = 100μm.

	Male			**Female**		
	Dsx+	GFP+	**(% dsx)**	Dsx+	GFP+	**(% dsx)**
Brain	151	43 ± 2	**28**	28	23 ± 1	**82**
VNC	305	102 ± 5	**30**	300	106 ± 4	**35**
Foreleg tarsus	72	57 ± 3	**80**	58	30 ± 3	**52**
Reproductive tract	-	-	**-**	?	40 ± 5	**?**

Table 1: Proportions of *dsx* neurons labelled by *SIFaR-p65AD* and *dsx-DBD* hemi-drivers. 'Dsx^{+}' cell counts are from Rideout et al., 2007, while 'GFP^{+}' cell counts refer to cells intersected using the *SIFaR-p65AD* and *dsx-DBD* lines in the current study.

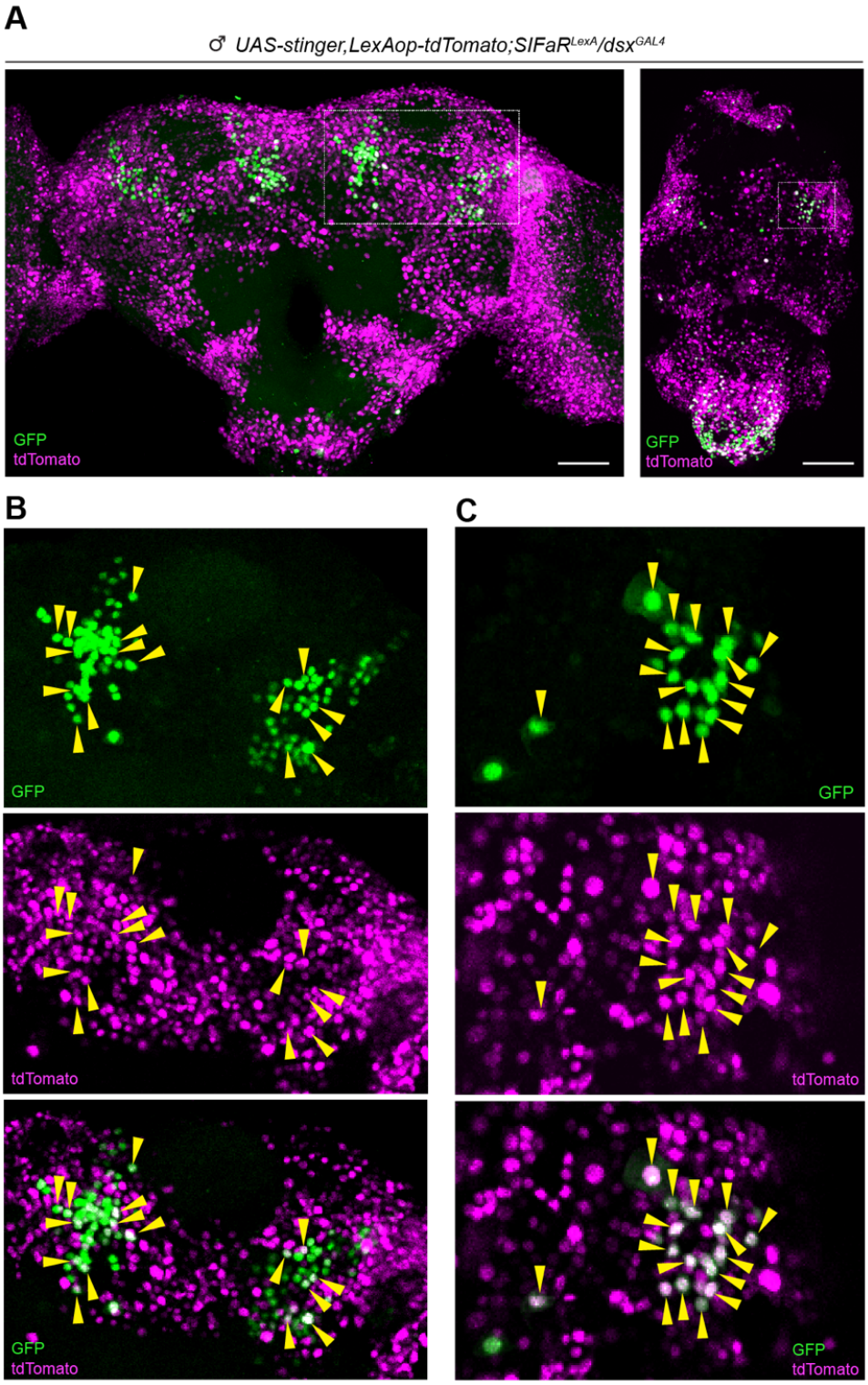


Figure 3: Overlap of *SIFaR-LexA* and *dsx-GAL4* expression in the male CNS. (A) Brain (left) and VNC (right) of males co-expressing TdTomato (magenta) under *SIFaR-LexA* control, and nGFP (green) under *dsx-GAL4* control. **(B)** Close up of boxed brain region, with yellow arrows showing co-expression of GFP and TdTomato. **(C)** Close up of boxed VNC region, with yellow arrows showing co-expression of GFP and TdTomato. Images in (A) are maximal confocal z-stacks, while overlay images (B & C) show individual z-sections. Scale bars = 50μm.

4.2.1.2 *SIFaR* expression in *fru* cells.

To investigate *SIFaR* expression in *fru* cells, the *SIFaR-p65AD* line was used in conjunction with the *fru-DBD line*; a CRISPR-MiMIC line generated by the Goodwin lab. In males, these hemidrivers drove GFP expression in large numbers of cells in the brain and VNC (fig.4A), suggesting pervasive SIFaR expression in fru cells. Examination of peripheral tissues in this line revealed GFP-labelled cells in all 3 sets of legs and in the labellum; a structure at the tip of the proboscis (fig.4C-E). Fluorescent reporters driven by co-expressed *SIFaR-LexA* and *fru-GAL4* showed overlapping expression in cells in the same regions, further supporting the conclusion that *SIFaR* is expressed in central and peripheral *fru* circuitry (fig.5). Though females do not produce functional Fru protein, transcription of fru still occurs and is captured by the *fru-GAL4* line. In females co-expressing *SIFaR-LexA* and *fru-GAL4*, overlapping reporter expression was observed in the brain, VNC and foreleg (fig.5). However, unlike in males, neither *SIFaR-LexA* nor *fru-GAL4* expressed in the antennae or proboscis (fig.6). Taken together, these findings support the conclusion that *SIFaR* is expressed in *fru*-expressing neurons in the PNS and CNS in both sexes.

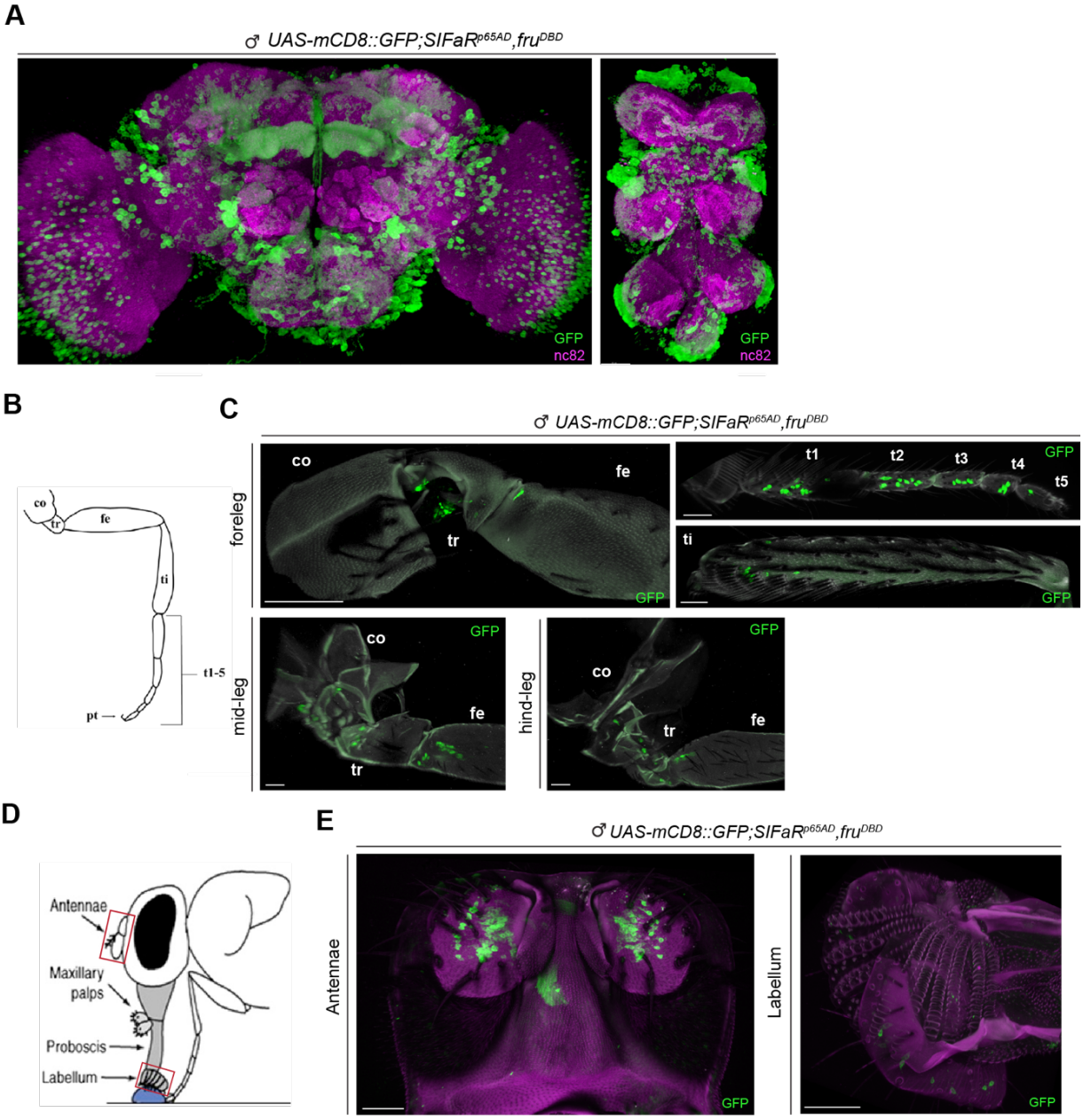


Figure 4: *SIFaR*∩*fru* expression in in the male CNS and PNS. (A) Male brain (left) and VNC (right) stained with nc82 (magenta), showing the neuropil structure, and GFP (green), showing intersected *SIFaR*∩*fru* cells**. (B)** Schematic of male foreleg, adapted from Sewell, 2006. **(C)** *SIFaR*∩*fru* cells (GFP; green) intersected in several regions of the male foreleg. **(D)** Schematic of male head showing antennae and proboscis structures (adapted from Reinhard, 2019). **(E)** Antennae (left) and Labellum (right) showing *SIFaR*∩*fru* cells (GFP; green) (tissue autofluorescence shown in magenta). Scale bars in the CNS = 50µm, in peripheral tissues = 100µm.

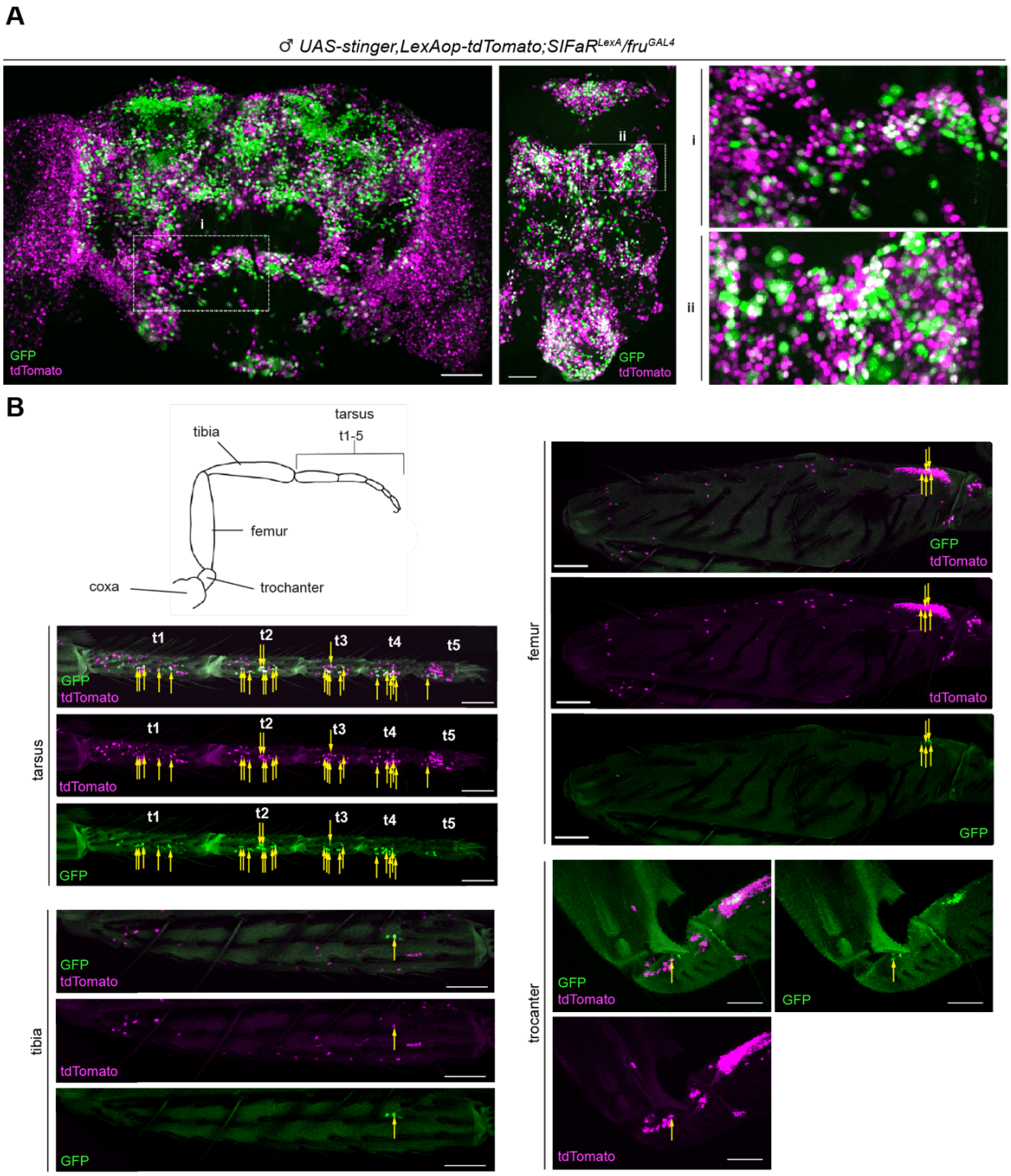


Figure 5: Overlap of *SIFaR-LexA* and f*ru-GAL4* expression in the male CNS & PNS (A) Brain (left) and VNC (right) of males co-expressing TdTomato (magenta) under *SIFaR-LexA* control, and nGFP (green) under *fru-GAL4* control. Far right: close-ups of boxed regions (i) & (ii), showing reporter co-labelling. **(B)** schematic of foreleg (adapted from Sewell, 2006) and confocal images showing GFP and tdTomato co-labelling in cells throughout regions of the foreleg. All images are maximal z projections of confocal z-stacks. Scale bars in the CNS = 50µm, in peripheral tissues = 100µm.

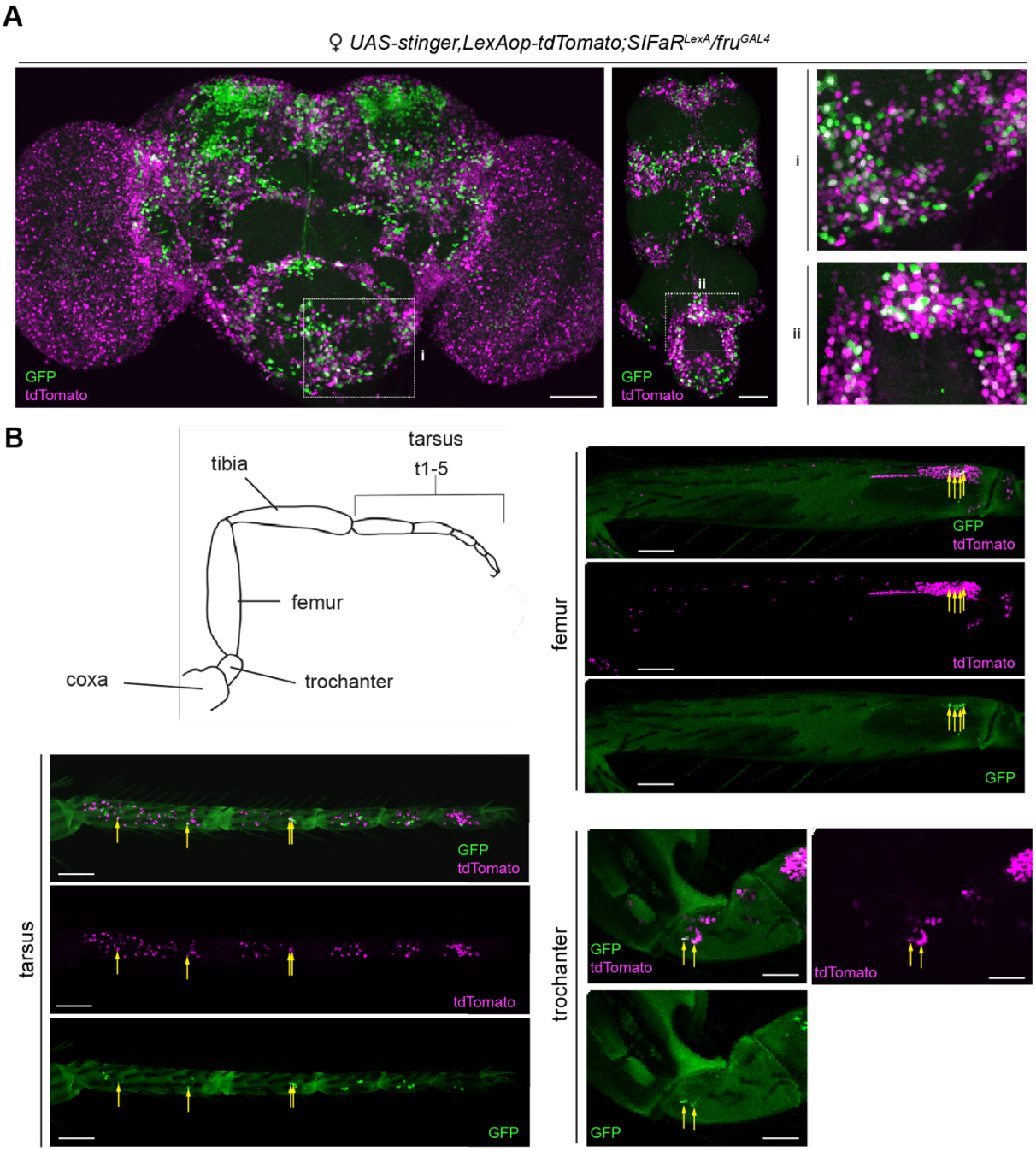


Figure 6: Overlap of *SIFaR-LexA* and f*ru-GAL4* expression in the female CNS & PNS. (A) Brain (left) and VNC (right) of females co-expressing TdTomato (magenta) under *SIFaR-LexA* control, and nGFP (green) under *fru-GAL4* control. Far right: close-ups of boxed regions (i) & (ii) showing reporter co-labelling. **(B)** schematic of foreleg (adapted from Sewell, 2006) and confocal images showing GFP and tdTomato co-expression in cells throughout regions of the foreleg. All images are maximal z-projections of confocal stacks. Scale bars in the CNS = 50μm, in peripheral tissues = 100μm.

4.2.2 *SPR* expression in *dsx* cells

With *SPR-p65AD* and *SPR-GAL4* lines generated in the current study, potential *SPR*∩*dsx*-expressing cells could be intersected using two different approaches. In the first, *SPR-GAL4* was used in conjunction with a dsx-Flp line (Rezával et al., 2014), generated by the Goodwin lab by insertion of the Flp recombinase (Flp) coding sequence into the *dsx* locus. These alleles were combined with a *UAS>STOP>GFP* line, in which reporter expression requires excision of the STOP cassette by Flp. Interestingly, two CNS expression patterns were seen in this genotype. In approximately half of the females analysed (n= 14/27), GFP was observed exclusively in neuronal processes in the abdominal ganglion (Abg) region of the VNC (fig.7A). In the remaining 13 females, the expression pattern was much broader, with cells throughout the brain and VNC GFP-labelled (fig.7B). The much broader expression pattern resembles that observed when using *SPR-GAL4* independent of the Flp/FRT system (see chapter 3, fig.3). This suggests either that Flp expression was not restricted to *dsx* cells in these animals, or that GFP expression was somehow induced in cells without excision of the STOP cassette.

Closer examination of the Abg neurons labelled in the more restricted expression pattern revealed that they project to/from the female reproductive tract (FRT) (fig.7C). Evidence from several studies supports SPR expression in FRT sensory neurons (Yapici et al., 2008) (Hasemeyer et al., 2009) (Yang et al., 2009) that are *dsx*+ (Rezával et al., 2012), and this conclusion is supported by data from the current study (see chapter 3, fig. 4). Via SPR, SP modulates the activity of these neurons to trigger the post-mating response (PMR). The GFP-labelled neurites evident in the *SPR-GAL4;UAS>STOP>GFP,dsx[Flp]* line could, originate solely from these sensory neurons. Alternatively, they could be associated with dsx neurons further down the hierarchy of the PMR circuit; namely, motor neurons that control egg-laying (Middleton et al., 2006; Rodríguez-Valentín et al., 2006; Rubenstein & Wolfner, 2013; Rezával et al., 2014).

A population of these descending Abg *dsx* neurons were shown previously to be Octopaminergic (OA) (Rezával et al., 2014), revealed by their positive staining for Tdc2; an

antibody that binds a key enzyme in the Octopamine synthesis pathway. To ascertain whether this neuronal population overlaps with that intersected by *SPR-GAL4* and *dsx[Flp], SPR-GAL4;UAS>STOP>GFP, dsx[Flp]* VNCs were co-stained with Tdc2. GFP expression and tdc2 staining did not overlap in the Abg region(fig.7C), suggesting that *SPR* is not expressed by OA*/dsx* neurons in the Abg. This suggests that, while *SPR* is found in *dsx* sensory neurons in the FRT, the receptor is not expressed by neurons elsewhere in the circuit that coordinates the female PMR.

SPR expression in *dsx* cells was also investigated using *SPR-p65AD* in conjunction with the *dsx-DBD* line. This hemidriver combination was used to drive expression of a nuclear localising GFP (nGFP), and expression was examined in the CNS in males and females. In both sexes, GFP-labelled cells were evident in the brain and VNC (fig.7D). In males, the positions of some GFP-labelled nuclei were consistent with them representing *dsx* neurons. However, there were additional cells labelled in regions of the CNS where *dsx* neurons are not thought to be located (fig.8A). *Repo* co-staining of the *SPR-GAL4* line in the current study had indicated *SPR* in glial cells (see chapter 3, fig.3) - a conclusion supported by the scRNAseq data. In addition, the scRNAseq data also posits the existence of some dsx^+ glial cells (Goodwin lab data, unpublished). Therefore, it was conceivable that these 'additional cells' could be SPR^+dsx^+ glia that were successfully intersected using these hemi-drivers. To investigate this possibility, the CNS of *SPR-p65AD;;UAS-nGFP,dsx-DBD* males were co-stained with *repo*. GFP and repo co-staining was, as expected, not observed in putative SPR^+dsx^+ neurons in the brain (fig. 8B) or VNC (fig.8E) but was observed in the majority of the remaining GFP-labelled cells, in the brain and VNC (fig.8C,F, respectively). This suggests that, in males, *SPR* is expressed in neurons throughout all levels of the *dsx* circuit hierarchy, and also in glial cells in the CNS.

Similarly, in *SPR-p65AD;;UAS-nGFP,dsx-DBD* females, the number and position of GFP-labelled cells was not consistent with expression restricted to *dsx* neurons. However, repo co-staining experiments were not performed with females from this line. Thus, it is unclear whether any or all these cells are *dsx* glia, or else represent 'leaky' expression of one or more of the transgenes. If any of these cells are, in fact, *dsx* neurons, this would be inconsistent with the

results obtained using the *SPR-GAL4;UAS>STOP>GFP,dsx[Flp]* line, which was indicative of *SPR∩dsx* expression being restricted to a population of neurons arborising the Abg. However, the inconsistent expression pattern observed using *dsxFLP* as a means of intersection itself raises a question mark over the reliability of this data. Further investigation of *SPR* expression in *dsx* cells in females is therefore needed.

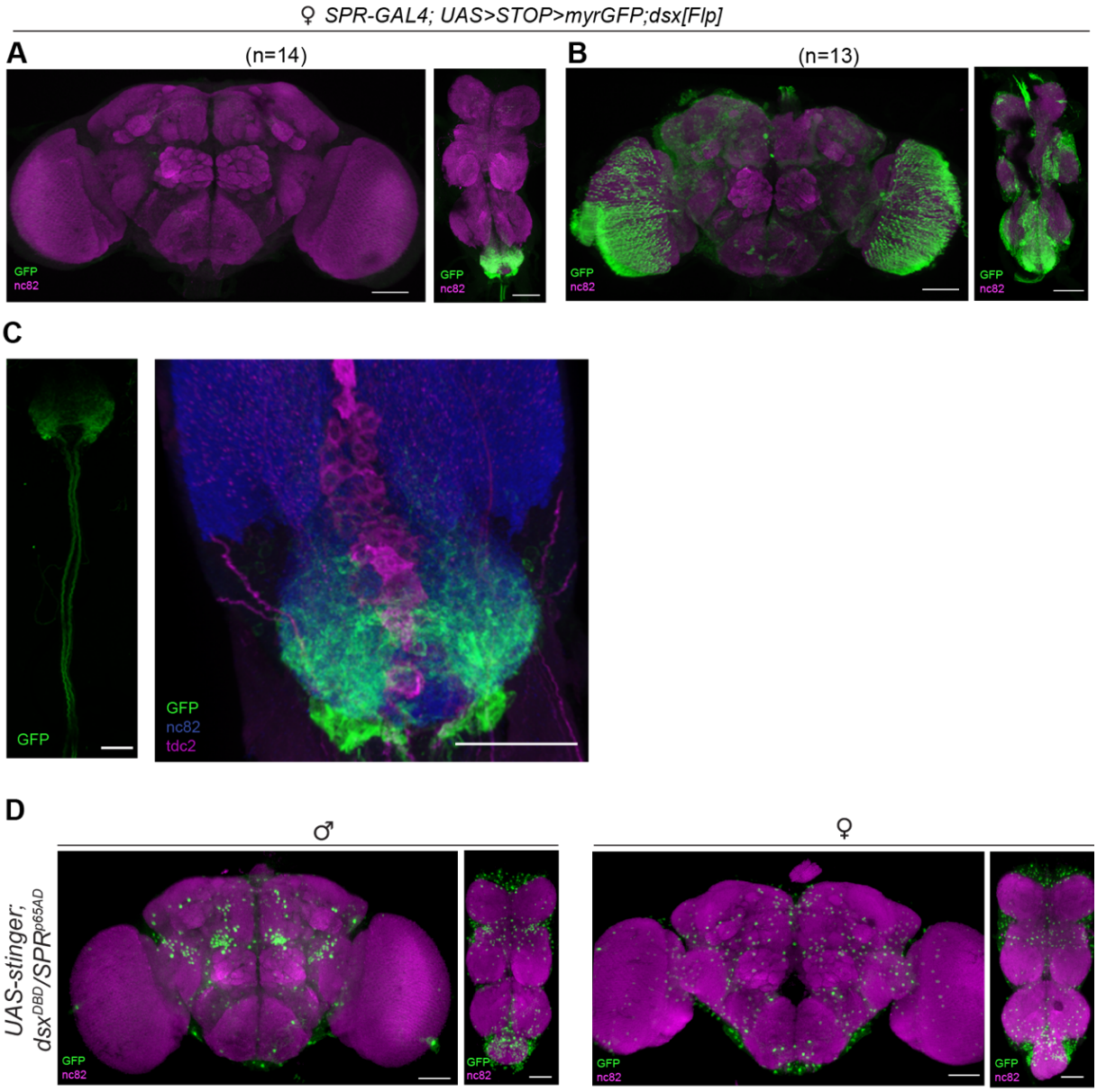


Figure 7: *SPR* expression in *dsx* and *fru* cells in the CNS. (A) *UAS>STOP>myrGFP* expression in the brain (left) and VNC (right), driven by *SPR-GAL4* and *dsx[FLP],* in 14 females of this genotype. **(B)** Expression in 13 females of the same genotype as in (A). **(C)** Left: neurites of VNC neurons labelled in (A), ascending from/descending to the female reproductive tract. Right: close-up of Abdominal ganglion region of VNC, co-stained with tdc2 (magenta). **(D)** nGFP (green) expression driven by SPR-p65ADand dsx-DBD, in the male CNS (left) and female CNS (right). All images are maximal z-projections of confocal stacks. Scale bars = 50µm.

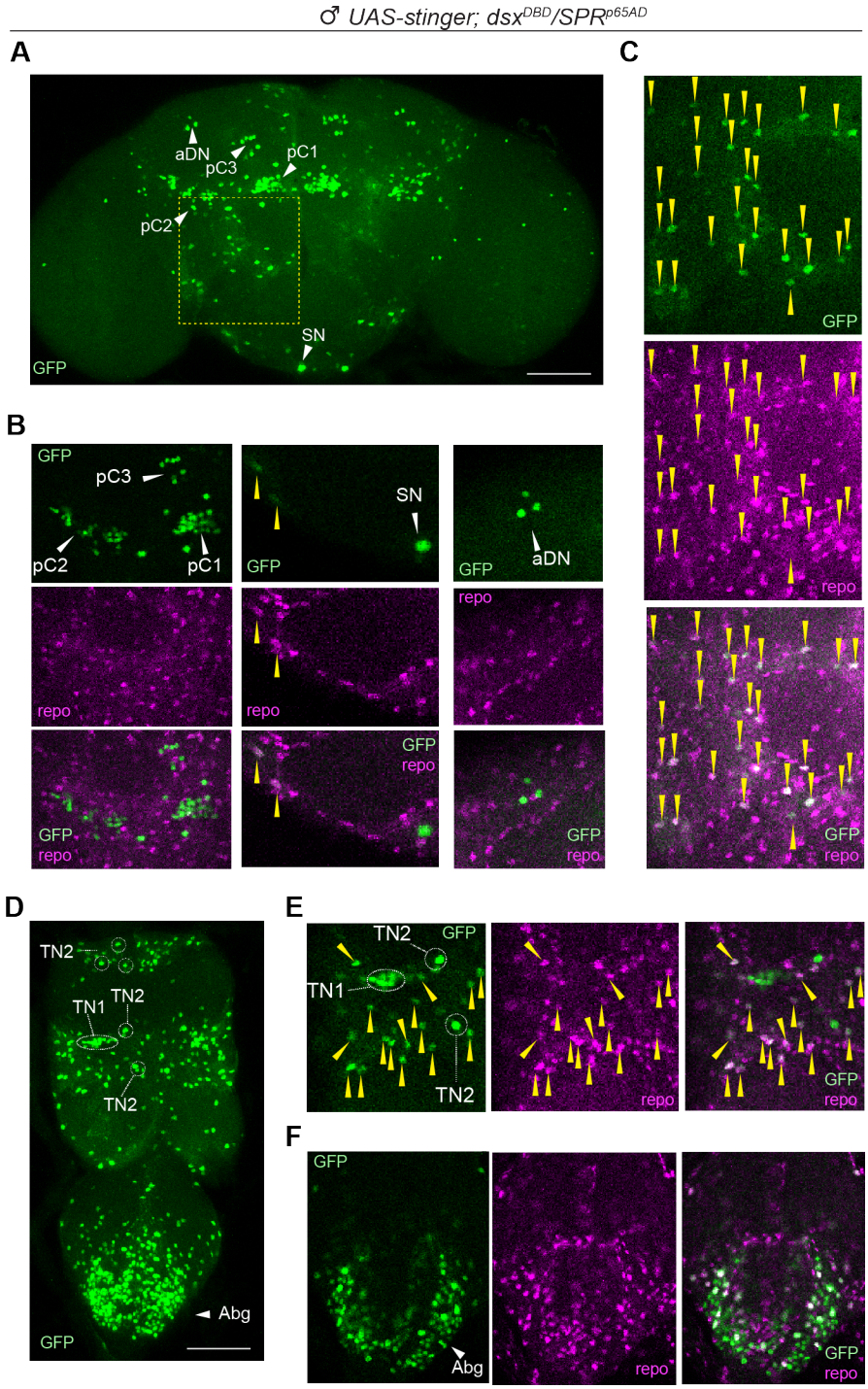


Figure 8: neuronal and glial *SPR∩dsx expression in the CNS*. (A) nGFP (green) expression driven by SPR-p65AD and dsx-DBD, in the male brain, with putative *dsx* neuronal clusters labelled. **(B)** Close-up images of nGFP-labelled putative *dsx* neurons (top panel) with *repo* staining (magenta) of same regions (middle panel) and overlay (bottom panel). **(C)** Close-up image of yellow boxed region in (A), (top) containing nGFP-labelled cells not thought to be *dsx* neurons with repo staining of same region (middle) and overlay (bottom). (**D**) nGFP (green) expression driven by *SPR-p65AD* and *dsx-DBD*, in the male VNC, with putative *dsx* neuronal clusters labelled. (*E*) Close-up images of nGFP-labelled putative *dsx* neurons (left), with *repo* staining of same regions (middle) and overlay (right). (F) Close up of Abg VNC in (D), showing nGFP-labelled cells (left), repo staining of same region (middle) and overlay (right). All images are individual z-sections of confocal stacks, except (A) and (D), which are maximal z-projections. Scale bars = 50μm,

4.2.3 *CrzR*, and *NPFR* expression in *dsx* cells

CrzR and *NPFR* expression in *dsx* cells was investigated using the MiMIC-*p65AD* lines generated for these receptors, crossed to a line expressing *UAS-stinger* (nGFP) in combination with the *dsx-DBD* hemidriver. *CrzR-p65AD∩dsxDBD* drove GFP expression in a more restricted population of cells in the CNS (fig. 9A). In the male brain, just 1 bilateral neuron was GFP-labelled, located in the subesophageal zone (SEZ). The widely arborising 'SN' neuron is the only *dsx* neuron found in this region (Nojima et al., 2021). In the VNC of males, a population of neurons in the Abg were labelled, which appear to represent a relatively small sub-set of the *dsx* neurons found in this region. In females, *CrzR-p65AD∩dsxDBD* drove GFP expression exclusively in a group of cells in the Abg of the VNC, which likely represents a sub-set of the Abg *dsx* neuronal cluster. Therefore, compared to, for example, *SIFaR*, the receptor for *Crz* appears to be expressed in a smaller proportion of *dsx* neurons.

In the male brain, *NPFR-p65AD∩dsxDBD* drove nGFP expression in a population of cells comparable in size to that intersected in the *SIFaR-p65AD∩dsxDBD* line, comprising neurons in all known *dsx* brain clusters (fig.9A). In the VNC of males, a putative subset of *dsx* Abg neurons were GFP-labelled, as were a population of *dsx* TN1 neurons, located in the Accessory Mesothoracic neuropil (AMNp)(fig.9A). *dsx* TN2 neurons, located closer to the VNC midline, which were intersected by *SIFaR-p65AD∩dsxDBD* line, were not GFP-labelled. Thus, unlike SIFaR, *NPFR* does not appear to be expressed throughout all *dsx* clusters in the male VNC. In females, *NPFR-p65AD∩dsxDBD* labelled just 3 bilateral *dsx* neurons in the brain; putatively, a single *dsx* pC1 neuron and 2 neurons in the *dsx* pC3 cluster. However, GFP expression in ~2 bilateral neurons in the SMP region, where there are thought to be no *dsx* neurons, was also observed (fig. 9A). In the VNC of females, GFP expression was evident in a significant proportion of putative *dsx* Abg neurons. Interestingly, expression was also consistently observed in a single bilateral cell in the AMNp, despite neuronal *dsx* expression being restricted to the Abg in females (fig. 9A). Expression in these cells, and in the cells identified in the SMP of the brain, could indicate either ectopic neuronal expression in this line,

or, as was shown to be the case for *SPR*, expression of *NPFR* in non-neuronal *dsx* cells. Alternative reporters and/or co-staining with neuronal or glial markers would be needed to confirm this.

It should be noted that in ~50% of the dissected animals from both the *CrzR-p65AD∩dsxDBD* and *NPFR-p65AD∩dsxDBD* lines, GFP expression was driven in a much larger population of cells in the CNS, including significant numbers of cells that could not conceivably be *dsx* neurons (fig. 9B). At present, the basis of this binary expression pattern is not clear. The identification of $dsx^{+}SPR^{+}$ glia, could indicate that the broader expression seen in these animals corresponds to glial expression. However, this does not account for the fact that this was observed in only ~50% of cases, with expression restricted to putative *dsx* neurons in the remaining animals. In addition, interrogation of the scRNAseq data set suggested that, unlike *SPR, CrzR* and *NPFR* expression is not pervasive in glial cells. This makes it unlikely that the extensive expression outside the *dsx* neuronal population reflects the expression of these receptors in *dsx* glia. As such, it is more likely that the *UAS-stinger;dsxDBD* line, to which the *NPFR-p65AD* and *CrzRp65AD* lines were crossed, conferred stochastic ‘leaky’ reporter expression.

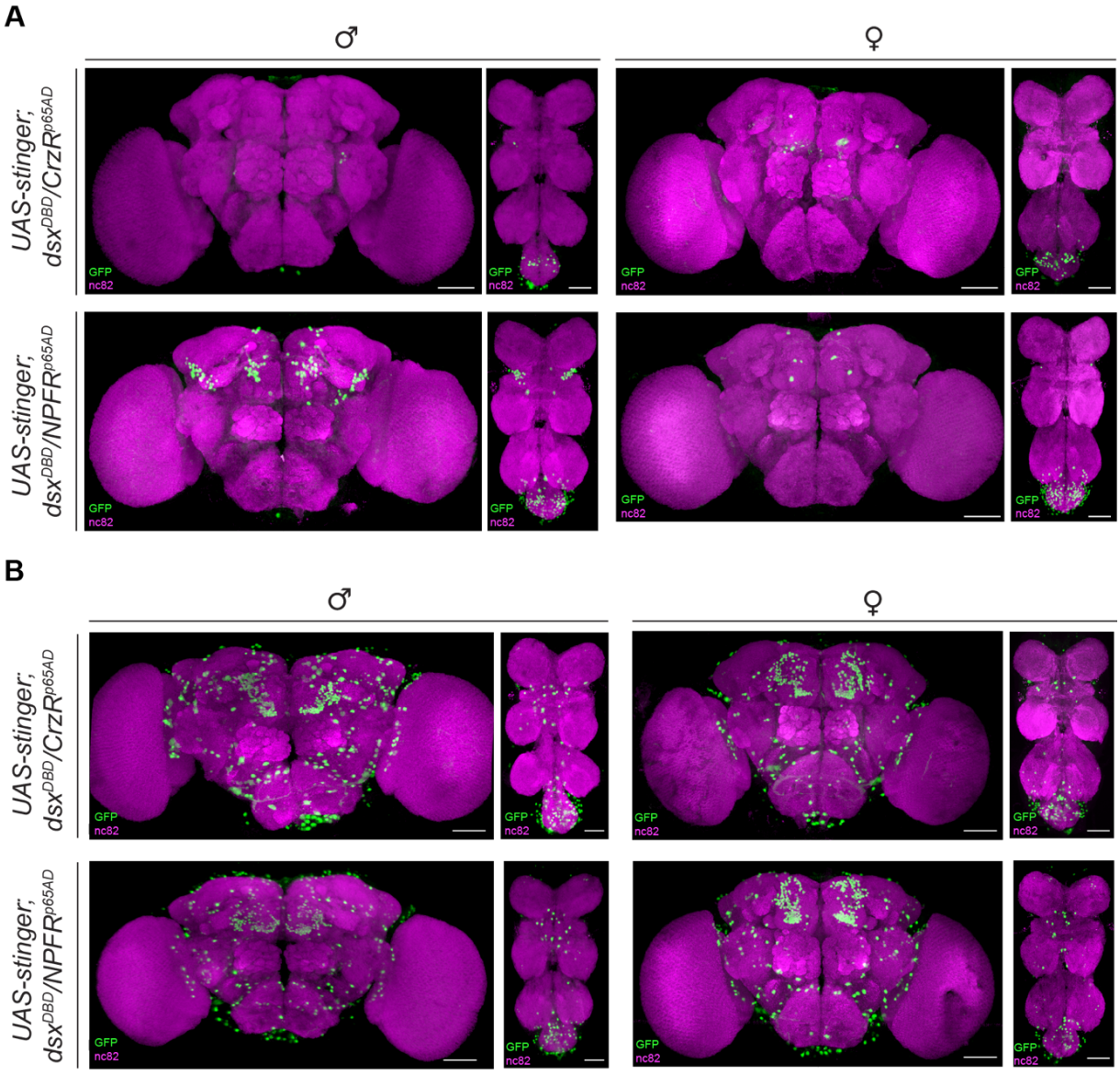


Figure 9: *CrzR*∩*dsx* and *NPFR*∩*dsx* expression in the CNS. (A) nGFP expression (green), driven by *CrzR-p65AD* (top row) and *NPFR-p65AD* (bottom row), combined with *dsx-DBD;* in ~50% of animals (n>30). **(B)** Expression in remaining ~50% of animals of the same genotypes as in (A). All images are maximal z-projections of confocal stacks. Scale bars = 50µm,

4.3 Summary and discussion

Intersectional genetic tools were used in the current study to address a simple question: are the receptors for neuropeptides implicated in the modulation of reproductive behaviour expressed in the neurons underlying these behaviours? Overall, the results of this investigation suggest that they are. This supports the possibility that the modulatory functions of these neuropeptides are mediated, at least in part, by direct signalling through the neurons in question. The evidence supporting *SIFaR* expression in sex-circuits was, arguably, the most complete and robust: expression in *dsx* and *fru* neurons in the CNS and PNS in both sexes was consistently observed, using both the *split-GAL4* and *SIFaR-LexA* lines in conjunction with *dsx* and *fru GAL4*. The functional significance of SIFaR expression throughout multiple levels of dsx and fru circuit architecture is that the SIFamide could potentially modulate the circuit at multiple levels. For example, *SIFaR* expression was evidenced in *fru*-expressing olfactory receptor neurons (ORNs) in the antennae, which respond to sex pheromones and are crucial for courtship (Stockinger et al., 2005), and also overlapped with *dsx* and *fru* in gustatory receptor neurons (GRNs) in the foreleg, which convey information about non-volatile sex-pheromones (Spieth, 1974). Thus, SIFamide could conceivably modulate courtship in males by sensitising/desensitising sensory input to central courtship circuits. Across all organisms, the initial stages of sensory processing are thought to be a key target for various neuropeptides (Carlsson et al., 2010; Bargmann, 2012), and SIFamide itself is thought to regulate feeding via modulation of gustatory and olfactory information processing (Martelli et al., 2017). However, SIFaR expression was also evidenced in *dsx* neurons in the central brain in both sexes. Thus, SIFamide could alternatively (or additionally) modulate the activity of higher order neurons that integrate multisensory information and execute behavioural responses. The potential for neuromodulators, including neuropeptides, to act at multiple levels of circuit architecture is evidenced in other invertebrate systems (Marder et al., 2014).

Evidence of *SPR* expression in *dsx* cells was, at least superficially, somewhat conflicting. In females, *SPR-GAL4* combined with *dsx[FLP]* produced expression solely in a population of

neurons innervating the FRT and Abg. The lack of tdc2 immunoreactivity in these neurons suggests they are not the descending octopaminergic *dsx* neurons that form the 'output-arm' of the post-mating response circuit (Rezával et al., 2014) but are instead likely to be the sex-peptide (SP) sensing *dsx* neurons that ascend from the FRT. In contrast, using the *split-GAl4* approach, putative *SIFaR*∩*dsx* cells were labelled throughout the CNS in females. However, evidence from *repo* co-staining in males suggests that a significant proportion of these cells may be glia. Since *dsx* expression in glia is supported by the scRNAseq data, it is conceivable that the *MiMIC split-GAL4* lines can capture this expression, while other reagents that aim to capture *dsx* expression (namely, the dsx-GAL4 and dsx-Flp lines) cannot. In males, *SIFaR*∩*dsx* cells in the CNS, intersected via *split-GAL4* lines, were confirmed to comprise *dsx* neurons and glia. Despite evidence to suggest that SP enters the haemolymph (Lung & Wolfner, 1999), there is currently no evidence of SP/SPR signalling occurring the CNS. Central SPR signalling is more likely to involve myoinhibitory peptide (MIP); the ancestral ligand for SPR (Kim et al., 2010). MIP/SPR signalling has been implicated in the modulation of female reproductive behaviour (Jang et al., 2017) (Hussain et al., 2017), but not that of males. The functional significance of *SPR* expression in male CNS *dsx* neurons is, therefore, not clear.

Though analysis of the scRNAseq data was able to provide general insights into the breadth and level of *NPR* expression, deriving meaningful information about the expression of these genes in *dsx* and *fru* cells is more challenging. However, preliminary analysis suggested that candidate *NPR*s and *dsx* expression does overlap in some cells (table 1). Information about the absolute numbers of cells is, however, of limited utility. This is due, in particular, to the propensity for 'lowly' expressed genes, like *NPRs*, to be subject to dropout, leading to underestimations of cell numbers. However, though intersections sizes in the scRNAseq data should not be directly compared to those obtained via transgenic approaches, the existence of co-expression is (despite the propensity for *NPR* dropout) is meaningful.

	Female				Male			
	VNS		Brain		VNS		Brain	
	# cells	*dsx*⁺ cells	# cells	*dsx*⁺ cells	# cells	*dsx*⁺ cells	# cells	*dsx*⁺ cells
dsx	351	-	184	-	292	-	518	-
SIFaR	138	17 (5%)	708	5 (3%)	217	23 (8%)	985	21 (4%)
CrzR	176	10 (3%)	1535	9 (5%)	236	16 (7%)	1593	26 (5%)
NPFR	302	11(3%)	2169	12 (7%)	341	16 (7%)	2380	30 (6%)
TkR86C	454	14 (4%)	2638	17 (9%)	396	8 (3%)	2396	44 (8%)
SPR	2825	201 (57%)	13002	65 (35%)	3031	86 (30%)	12582	174 (34%)

Table 2: preliminary analysis of *dsx* and *NPR* co-expression scRNAseq VNC and brain data sets
Percentages, in brackets, refer to the proportion of *dsx* cells containing *NPR* transcripts

Chapter 5: Modulation of Reproductive Behaviour by SIFamide

5.1 Introduction

A potential role for SIFamide in modulating reproductive behaviour was highlighted by SIFamide loss-of-function (LOF) experiments. Loss of the peptide produces indiscriminate courtship in males, while SIFamide-*less* virgin females become more receptive to male courtship (Terhzaz et al., 2007). Based on these findings, it has been suggested SIFamide might function be to suppress reproductive behaviour in contexts in which its expression is 'inappropriate'. *How* SIFamide carries out this function, however, remains unclear. In the current study, the functional implications of disrupted SIFamide signalling were investigated by targeting the peptide's only known receptor, SIFaR. As a tool for various applications, a *SIFaR* gene deletion was generated (section 5.2.1). Reproductive behaviours were analysed in animals harbouring a mutagenic insertion in *SIFaR*, predicted to confer strong SIFaR loss- of-function (section 5.2.2). Having already identified *SIFaR* expression in *dsx* neural circuits (see section 4.2.2.1), activity in which drives most aspects of reproductive behaviour, it was of interest to determine whether this expression is required for normal reproductive behaviour. To this end, the split GAL4 system was used to target expression of a *SIFaR-RNA*i construct to *SIFaR*∩*dsx* cells, and reproductive behaviours were analysed in this line. To complement targeted loss-of-function experiments, a *UAS-SIFaR line* was generated to selectively 'rescue' *SIFaR* expression (section 5.3). A *SIFaR* 'FLP-STOP' line was also generated to allow temporally controlled disruption of *SIFaR* in sex-circuits (Section 5.4). Finally, to understand the physiological effects of SIFaR signalling in target neurons, potential signal pathways for the receptor were investigated. Experiments facilitating visualisation of *SIFaR* signalling in sex-circuits *in-vivo* were then designed and validated (section 5.5).

5.2 Reproductive behaviour in *SIFaR* mutants

Animals in which SIFamide neurons have been genetically ablated, or in which the peptide has been knocked down by RNAi, display overt reproductive behaviour phenotypes. In males

courtship is produced towards males, as well as females, while females copulate more quickly; an indicator of increased sexual receptivity (Terhzaz et al.,2007). The following points should be noted, however. First, ablation of SIFamide neurons removes not only the peptide from the CNS, but any other influence these neurons may have. Second, though the SIFamide RNAi construct was shown to effectively knockdown the peptide (Terhzaz et al., 2007), unintended targets cannot be entirely ruled out when using this approach. Third, though a ligand/receptor relationship has been established for SIFa/SIFaR, peptides can have affinity for additional receptors. Therefore, to determine whether SIFamide signalling through SIFaR *per se* is important for the regulation of reproductive behaviour, it was of interest to analyse reproductive behaviour in *SIFaR* LOF animals. For this purpose, and for its application to other experiments (for example, to validate reagents and assays devised to measure *SIFaR* expression), a *SIFaR* gene deletion was generated (section 5.2.1). Finding this genetic deletion to be homozygous lethal, however, phenotypic analysis of SIFaR loss-of-function was conducted using the mutagenic *SIFaR-p65AD*$^{[MiMIC]}$ allele (section 5.2.2).

5.2.1 Generation and characterisation of a *SIFaR* deletion allele

Gene deletions constitute true null alleles, ensuring complete loss of the encoded protein's function. A screen of available genetic deficiency lines identified a line, *Df(3R)ED10838*, in which a region encompassing *SIFaR* is excised (fig.1A). Spanning some 160 Kbp's, however, this deletion covers ~25 genes downstream of *SIFaR* and is, unsurprisingly, homozygous lethal. A more precise *SIFaR* deletion was therefore needed. Numerous strategies can be used to produce gene deletions (summarised in Venken et al., 2011a). In the current study, a *SIFaR* deletion allele was generated by Flp-mediated recombination between transposable elements (TEs) flanking the *SIFaR* locus (see chapter 2.2.3 for detailed protocol). Flp-mediated recombination can occur only between TE's bearing Flp target 'FRT' sites in the same orientation. A screen of TEs in the vicinity of *SIFaR* identified two suitable candidates: *P{RS5}5-SZ-3483*, 'RS5'; a P-element construct from the *DrosDel* collection (Ryder et al.,

2004), located 5' to *SIFaR*, and *PBac{RB}CG3301e03242,* 'RB'; a piggy-bac construct, part of the *Exelixis* collection (Thibault et al., 2004), located 3' to *SIFaR* (fig.1B,C). The intervening region excised by recombination, thus, includes the whole *SIFaR* sequence. It should be noted, however, that the deletion implicates two additional genes*: CG17298;* unannotated at the time of the experiment and the function of which is still not described, and *CG3301;* believed to encode a steroid dehydrogenase enzyme. Unfortunately, disruption of these genes was unavoidable using this method: *CG17298* is located within the *SIFaR* locus, while *CG3301* lies just ~400 bp downstream of *SIFaR,* and no TE's lie within this region. Flies harbouring RS5 and RB insertions on homologous chromosomes and expressing Flp recombinase under the control of a temperature-sensitive promotor (hsFlp), were heat-shocked to induce recombination (fig.1B,C). The presence of *miniwhite,* a gene producing eye pigmentation, in both TE's permitted screening for recombination by loss of eye colour. Successful excision of the target region in recombinants was confirmed by PCR, using published and custom designed primers and by sanger sequencing of the deletion locus (fig. 2A,B) (see chapter 2, materials and methods for full protocol).

The newly generated SIFaR deletion allele (*SIFaR^{del})* allele was placed in *trans* to a balancer chromosome. After several generations, it became apparent that progeny lacking the balancer marker were absent from the population, suggesting the presence of both copies of *SIFaR^{del}* might not be viable. To test this hypothesis, the *SIFaR^{del}* allele was placed in *trans* to a 3rd chromosome balancer expressing a red fluorescent protein (RFP) marker, allowing visual identification of *SIFaR^{del}* homozygotes at all developmental stages. While RFP- eggs were evident in this line, only RFP^{+} eggs gave rise to larvae (n= ~1000), suggesting that the *SIFaR^{del}*/*SIFaR^{del}* genotype is embryonic lethal. This lack of viability could reflect an essential role for the other genes implicated in the deletion: *CCG3301* and *CG1729,* rather than for SIFaR. There is, however, evidence to suggest that this is not the case, and that it is loss of *SIFaR,* specifically*,* that produces embryonic lethality. First, lines in which expression of *CG3301* or *CG1729* is perturbed, either via a mutagenic insertion or expression of RNAi, are viable (according to stock centre reports), suggesting these genes aren't required for survival.

Second, lethality was subsequently observed in two other lines in which the expression of *SIFaR* is disrupted: first, the *SIFaR*^{B322} line, in which a mutagenic PBac{5HPw+} construct is inserted in *SIFaR,* and second, in animals expressing the *VDRC 1783* RNAi construct, which targets *SIFaR,* in all neurons. Considered in isolation, lethality in either context does not afford conclusive evidence that *SIFaR* is essential in development, particularly because, though *SIFaR*^{B322} is lethal, two other mutagenic insertions in *SIFaR*; *Mi{MIC}MI05376* and *Mi{MIC}MI088* (the latter of which also disrupts the *CG17298* gene) are not, and, though pan-neuronal expression of *the VDRC 1783* RNAi construct is lethal, this is not the case for 3 other RNAi constructs targeting *SIFaR*. However, taken together, and supported by the observation that the *SIFaR*^{del} allele is homozygous lethal, these findings strongly indicate that *SIFaR* serves an essential function in early development. A previous study, in which an RNAi-based mutagenic screen was conducted to identify GPCRs involved in *Drosophila* fight regulation, found that lethality in the context of pan-neuronal *VDRC 1783* RNAi expression could be rescued by expression of *Drosophila* Stromal Interaction Molecule (dSTIM) (Agrawal et al., 2013). Based on this finding, in the current study, *elav-GAL4 and UAS-dSTIM* transgenes were incorporated into the *SIFaR*^{del} line, in an attempt to rescue the lethality of the deletion. However, progeny expressing both transgenes, that were homozygous for *SIFaR*^{del}, did not survive beyond embryonic stages. Thus, while dSTIM expression may rescue the lethality of SIFaR knockdown by RNAi, it does not appear capable of doing so in the context of a true *SIFaR* null allele.

The current study focused upon the role of *SIFaR* signalling in the adult nervous system. That *SIFaR*^{del} homozygotes do not survive beyond embryonic stages presented a significant obstacle to this; precluding the possibility of analysing adult phenotypes associated with complete *SIFaR* LOF, and preventing the use of this line as a negative control in other experiments. The finding does, however, highlight an essential function for *SIFaR* in early development, which has yet to be investigated. The 2013 study of Agrawal et al. provides some preliminary clues as to this function. The authors found that ubiquitous expression of *AcGq*, which encodes a constitutively active form of Gq, the heteromeric G protein, did not rescue the

lethality produced by RNAi mediated *SIFaR* knockdown, suggesting *SIFaR* may not signal through this pathway. The lethality was, however, rescued by expression of dSTIM; a molecule that senses endoplasmic reticulum Ca^{2+} levels and triggers pathways to replenish them accordingly (Agrawal et al., 2013). This suggests *SIFaR* could play an essential role in the regulation of intracellular Ca^{2+} in neurons during early development. Further investigation would be needed to clarify this, however.

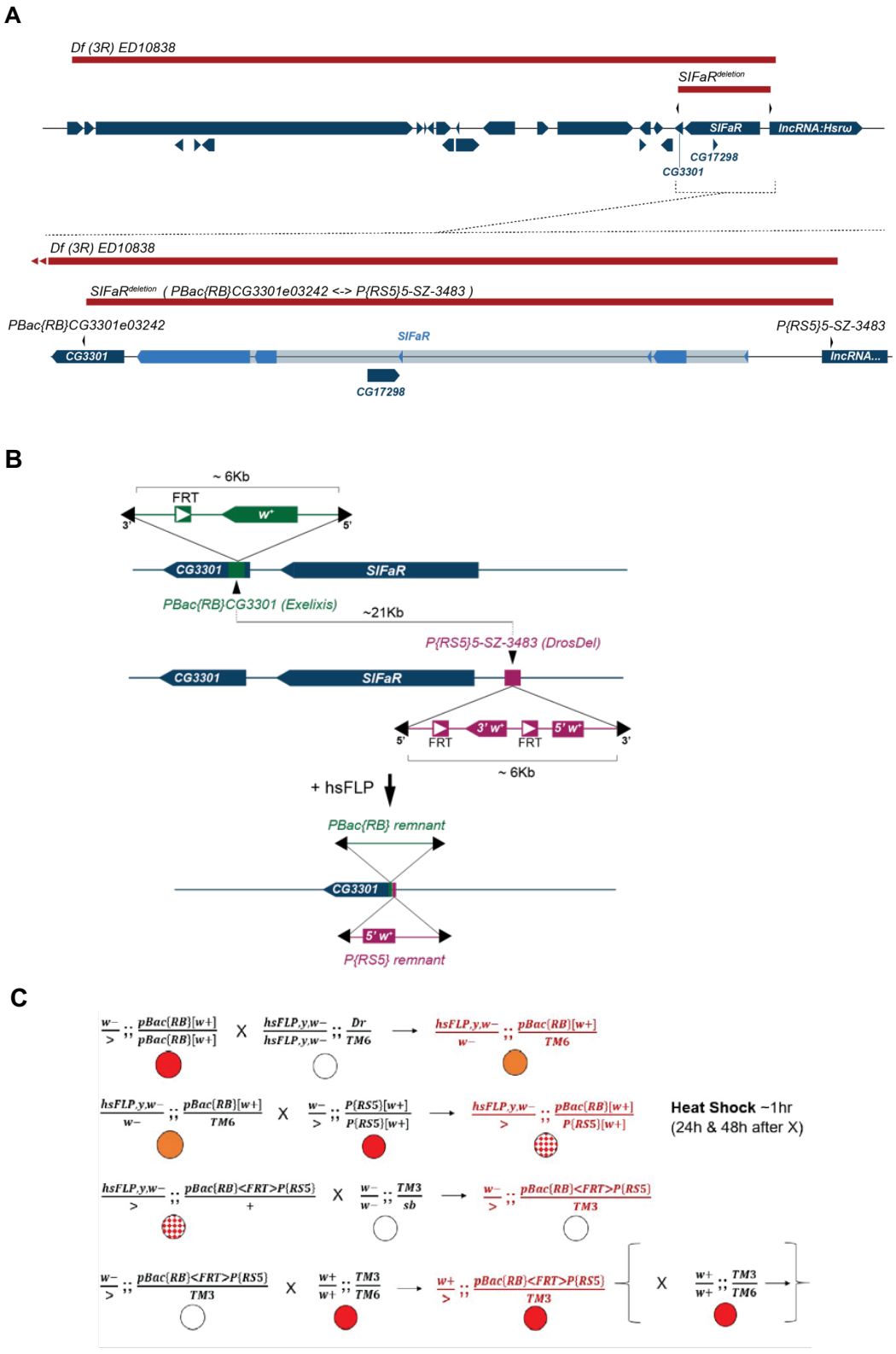


Figure 1: *SIFaR* Deletion generation. (A) (Top): Schematic of *SIFaR* region, showing extent of Df(3R)ED10838 deletion; location of transposable elements (TE's) selected for novel deletion generation (black arrows) and extent of this deletion. (Bottom): Close up of *SIFaR* locus, showing *P{RS5}5-SZ-3483* and *PBac{RB}CG3301e03242* TE insertion sites. **(B)** (Top): SIFaR locus, with expanded view of both TE's used for recombination. (Bottom): locus after Flp mediated recombination between FRT sites within TE's. **(C).** Overview of the crossing schedule used for deletion generation.

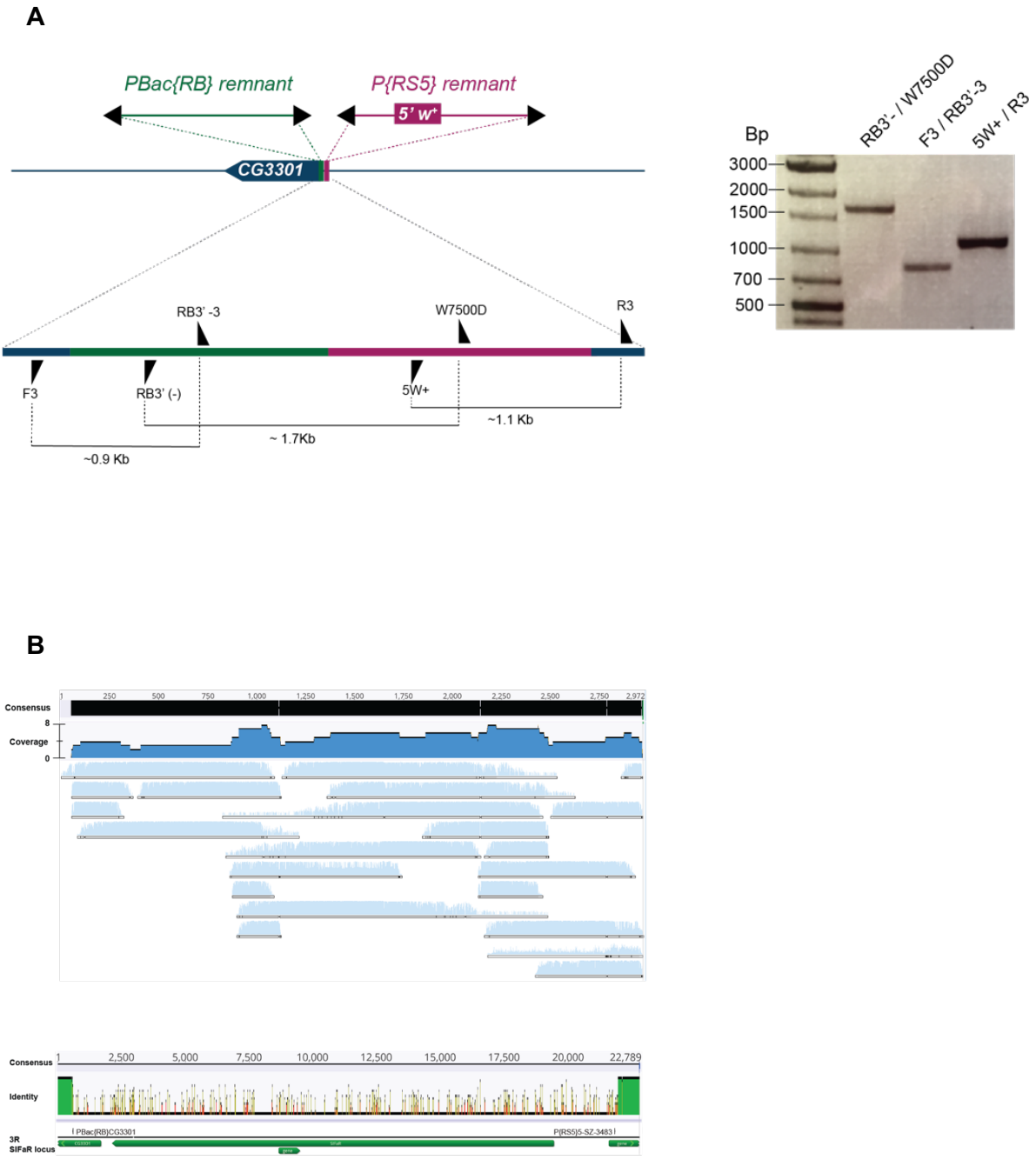


Figure 2 SIFaR *Deletion* Confirmation. (A) Schematic: locus after Flp mediated recombination between FRT sites within TE's. Expanded view shows location of primer binding sites for confirmation of recombination by PCR. Gel image: image of DNA fragments amplified by PCR, run on poly-acrylamide gel, confirming recombination of the locus. (**B**). Overview of sanger sequencing to confirm deletion generation. Top: Full coverage of the W7500D-RB3'(-) fragment used for sequencing achieved using a range of primers (coverage for each primer shown in light blue, overall coverage in dark blue above). Bottom: Analysis of sequence identity of W7500D-RB3'(-) fragment using sequencing primers, showing a drop off in coverage in the predicted deletion region.

5.2.2 *SIFaR-p65AD[MiMIC]* as a *SIFaR* mutant allele

To investigate the expression pattern of *SIFaR,* the Trojan-MiMIC system was used to generate a *splt-GAL4 line*, via insertion of the T2A-p65AD-polyA 'Trojan' exon at [*MI05375];* a MiMIC in the 3rd intron of *SIFaR.* Evidence from the first use of the Trojan-MiMIC system showed that T2A Trojan exons function as effective gene traps when inserted in coding introns, strongly disrupting the expression of the genes into which they are inserted (Venken et al., 2011b). The mutagenic capabilities of MiMIC-T2A Trojan exon insertions made a compelling case for using *SIFaRp65AD[MiMIC]* line to investigate SIFaR loss-of-function in the current study. Accordingly, the reproductive behaviour of *SIFaRp65AD[MiMIC]* homozygotes*;* hereafter *'AD/AD',* was analysed, to ascertain whether loss of *SIFaR* phenocopies loss of *SIFamide.*

5.2.2.1 Reproductive behaviour phenotypes in *SIFaR-p65AD[MiMIC]* homozygous males

It was previously shown that knockdown of SIFamide does not the affect the courtship behaviour of males paired with receptive females (Terhzaz et al., 2007). To test whether this is also true when *SIFaR* is perturbed, *AD/AD* males were paired with wild-type (CS) virgin females and courtship behaviours were analysed. These males displayed the same degree of courtship as wild-type and heterozygous (*AD/+)* controls, suggesting that courtship towards receptive females is unaffected in *AD/AD* males (fig.3A). Next, courtship behaviour was analysed in *AD/AD* males paired with wild-type males. In this context, these males showed significantly more courtship than wild-type and heterozygous controls (fig.3B). Thus, loss of *SIFaR* phenocopies SIFamide perturbation. To rule out possible effects of off-target mutations in this line, the *SIFaRp65AD[MiMIC]* allele was placed in trans to the *SIFaR^{del}* allele generated by recombination between TEs flanking *SIFaR* (see section 5.2.1) '*AD/Del'* males also showed increased levels of male directed courtship (fig.3C), suggesting the phenotype arises specifically due to perturbation of *SIFaR.*

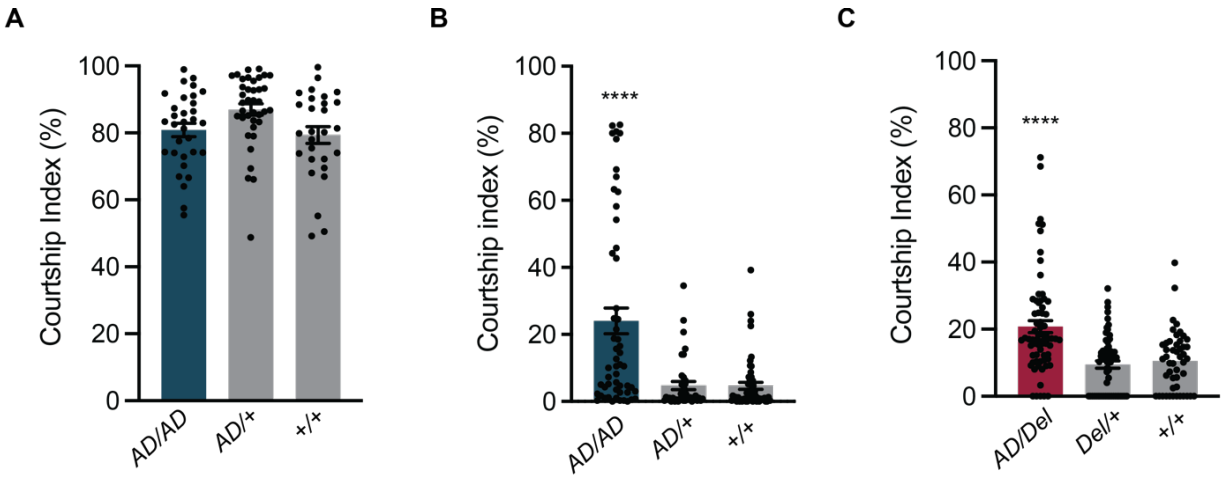


Figure 3: Courtship behaviour in SIFaR -/- males. (A) Courtship indices of *SIFaRp65AD*$^{[MiMIC]}$ homozygous *(AD/AD*), and control (*AD/+, +/+)* males paired with wild-type virgin females. **(B)** Courtship indices of males paired with wild-type males. **(C)** Courtship indices of *AD/Del* and control (Del/+, +/+) males paired with wild-type males. Bars and error bars show means ± SEM. One-way ANOVA $p < 0.0001$ (****), with Brown-Forsythe and Welch ANOVA tests (equal SDs not assumed) and Tukey's multiple comparisons.

The indiscriminate courtship displayed by *AD/AD* males could reflect a state of heightened, or disinhibited, sexual arousal. It was of interest to determine whether this state might manifest in other phenotypes, not previously examined in the context of SIFamide loss-of-function. A 'hypersexual' phenotype is unlikely to be discerned in a standard assay of courtship towards receptive females, since wild-type males display unabetted courtship in this context. Displays of courtship towards females do, however, decrease with repeated mating: a phenomenon known as reproductive satiety. This was illustrated in a 2016 study by Zhang et al., which showed that, when placed with multiple receptive females for a ~4-hour period, males copulate ~3 times and show decreasing courtship behaviours over time (Zhang et al., 2016). A state of increased sexual arousal in males could, therefore, be reflected by an increase in the threshold for reproductive satiety. To determine whether reproductive satiety is affected by *SIFaR* loss-of-function, individual *AD/AD* males were placed in vials containing 15 wild-type virgin females and courtship and copulatory behaviours were scored over a 4-hour period. Both control (CS) and *AD/AD* males copulated an average of 6-7 times within this time frame (fig.4A). Monitoring of these males for a further 2 hours revealed that they did not copulate thereafter. Interestingly,

this is double the number of copulation events recorded in wild-type males over the same time frame in the original study by Zhang et al. However, the result suggests that *SIFaR* loss-of-function males do not copulate more times with receptive females. Further analysis revealed that the period between copulation events increased at the same rate for both genotypes and the duration of each copulation event did not change (fig.4B,C). Moreover, the reproductive satiety index; the difference between time spent engaging in courtship or copulation in the last hour vs the first hour of the assay, was unaffected in *AD/AD* males (fig.4D). Taken together, these findings suggest that male reproductive satiety is unaffected by *SIFaR* loss-of-function. As such, if *AD/AD* males do exist in a state of heightened sexual arousal, the mechanisms that drive reproductive satiety (for example, dopaminergic input to courtship circuits (Zhang et al., 2016)) presumably override this state.

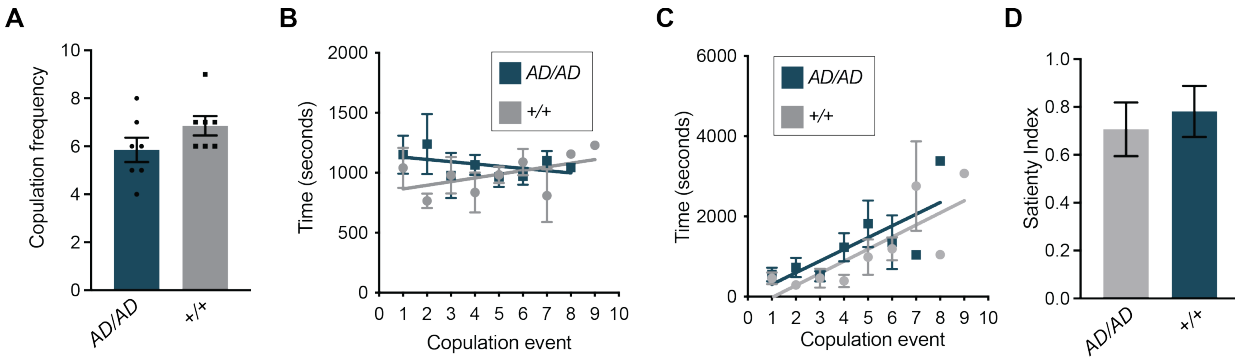


Figure 4: Reproductive satiety in SIFaR -/- males. (A) Number of separate copulation events for *SIFaRp65AD[MiMIC]* homozygous *(AD/AD*) and CS males (+/+) placed with ~15 CS virgin females for a ~4.5-hour period (unpaired t-test, with Welch's correction p = ns). **(B)** Duration of separate copulation events (t-test, Welch's correction & Linear regression analysis: no difference between genotypes (p= ns) and no change among copulation events). **(C)** Time between copulation events: increasing time between copulation events (linear regression pooled slope = 297.6), no difference among genotypes (t-test, Welch's correction, p = ns). (**D**) Reproductive satiety index (Satiety index); the proportion of time spent in-courtship or copulation during the last hour, compared to the first hour of the assay (unpaired t-test with Welch's correction, p = ns)

When analysing the data from the male-male courtship assay, it became apparent that the penetrance of this phenotype was variable: a few *AD/AD* males courted wild-type males at similar levels as they did females; others, like controls, showed little or no male directed courtship, and some displayed courtship levels between these extremes (fig.3B). In the original experiments assaying SIFamide loss-of-function (Terhzaz et al.,2007), sample sizes were significantly smaller (~10, *versus* ~ 60 in the current study) and individual data points were not plotted. However, the mean courtship indices and reported standard errors from these experiments suggest that the penetrance of the phenotype was both stronger (*SIFamide-less* males were reported to court males 'as intensely' as they did females (Terhzaz et al, 2007)), and less variable. Interestingly, in a follow-up study from the same research group, knockdown of SIFaR by RNAi in males also produced male-directed courtship, but to a lesser extent than in the context of SIFamide perturbation, and with a similar degree of variability to that observed in *AD/AD* males the current study (Sellami & Veenstra, 2015).

The phenotypic variability observed in the context of SIFaR loss-of-function could occur for several reasons. One possibility is a 'biological' explanation. For instance, the broad expression pattern of *SIFaR*, and the importance of *SIFamide* in the regulation of sleep and feeding, means that global *SIFaR* perturbation likely produces a syndrome of effects. Subtle differences in the conditions to which *SIFaR* disrupted individuals are subjected, particularly those affecting sleep or feeding, could, therefore, be amplified or exacerbated, leading to phenotypic differences in specific behavioural contexts. To rule out the possible influence of these environmental factors, experiments in *AD/AD* males were repeated in the current study, controlling more stringently for subtle variance in rearing or experimental conditions: experimental progeny were generated from crosses comprising a fixed number of flies, were collected within 1 hour of eclosure and were tested at a fixed time of day. However, none of these measures altered the experimental outcome. A further argument against this hypothesis is that phenotypic variability of this degree is not observed in the context of SIFamide perturbation, which should, likewise, globally disrupt SIFa/SIFaR signalling. Another potential biological explanation is that SIFamide has SIFaR independent actions, presumably signalling

through an additional receptor. However, this is unlikely, as this ligand/receptor couple shows particularly strong evolutionary conservation, and there are no paralogues of SIFaR (Jorgensen et al., 2005). The possibility of other SIFamide receptors has not been specifically investigated, however.

Alternatively, the phenotypic variability associated with SIFaR perturbation could reflect incomplete SIFaR LOF in these SIFaR perturbed genotypes. This could well apply to the findings of Sellami & Veenstra (2015), as RNAi is known to confer incomplete knockdown of target genes. In the current study, this explanation would imply variability in the extent to which the *SIFaRp65AD*$^{[MiMIC]}$ allele functions as an effective gene trap. Assuming successful splicing of the *T2A-p65AD*-polyA Trojan exon into the endogenous transcript, the position of the insertion; upstream of 3 *SIFaR* exons encoding >50 % of the protein, makes it inconceivable that the truncated form of SIFaR produced from the locus could retain its function. Therefore, incomplete loss-of-function could only result from failure of the Trojan exon to splice into the endogenous transcript, due to 'exon skipping' (Rueter et al.,1999). In the first study to test the mutagenic capacity of T2A polyA Trojan exons, complementation tests were performed on 100 lines with intronic MiMIC T2A-GAL4-poly-A exons in essential genes. When combined with a deficiency for the same gene, an allele in which the gene is entirely absent, 90% of these insertions produced lethality (Lee et al., 2018). This finding indicates that exon skipping of T2A polyA Trojan exons is not sufficiently widespread for the lethality of an essential gene to be rescued. It does not, however, rule out its occurrence entirely. For instance, if, due to exon skipping, the expression of a target gene persisted in 20% of cells, but expression of that gene were required in > 50% of cells for survival, lethality would still be observed. Moreover, in the current study, the very fact that the *SIFaRp65AD*$^{[MiMIC]}$ allele is homozygous viable, when other lines of evidence (acquired subsequently) suggest that *SIFaR* is essential in development, implies that it produces incomplete *SIFaR* loss-of function. If differences in the extent of exon skipping do account for the phenotypic variability of *AD/AD* males*,* these differences could be detected in assays that measure overall *SIFaR* expression level. However, it's also possible that the phenotypic variability is explained not by differences in the *extent* of exon skipping

among individuals, but in the *identities* of the cells in which exon skipping has occurred. For example, in two individuals, exon skipping in a population of cells similar in size, but which is non-overlapping and functionally distinct, could have different phenotypic consequences. Unfortunately, the lack of specificity of the SIFaR antibody produced in the current study and the failure to establish reliable reference genes across genotypes in a qPCR assay, precluded further investigation into the strength for SIFaR knockdown across genotypes.

5.2.2.2 Reproductive behaviour phenotypes in SIFaR-p65AD[MiMIC] homozygous females

Previous reports suggest that females lacking SIFamide are more sexually receptive, copulating more quickly with wild type males (Terhzaz et al., 2007). To determine whether SIFaR loss-of-function produces the same effect, virgin *AD/AD* females were paired with wild-type males and latency to copulation was scored. In this experiment, both *AD/AD* and *AD/+* females copulated more quickly than wild-type (CS) females (fig.5A). This was unexpected, as haploinsufficiency is rarely observed, and the male behavioural data had suggested that AD/+ individuals are, behaviourally, wild-type. Repeat experiments are therefore needed to ascertain whether *SIFaR* loss-of-function phenocopies loss of SIFamide in virgin females. Previous studies of SIFamide's role in female reproductive behaviour did not consider female post-mating behaviours. These behaviours were therefore analysed in *AD/AD* females in the current study. The number of eggs laid within 24 hours of mating did not differ among AD/AD and control females (fig.5C), suggesting *SIFaR* loss-of-function does not affect egg laying. However, 4 out of 30 *AD/AD* females copulated 24 hours after mating, whereas none of the AD/+ or wild-type females re-mated (fig.5B). An ~11% remating percentage was not statistically significant with this sample size, however, since wild-type females invariably fail to remate, it is noteworthy. Further experiments with larger sample sizes should be conducted to determine whether there is a remating phenotype associated with *SIFaR* LOF.

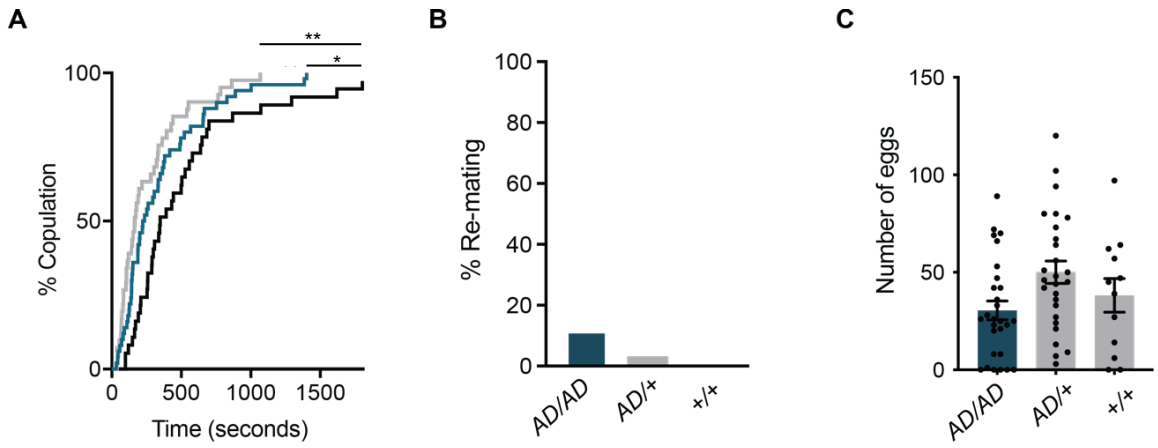


Figure 5: Reproductive behaviour in SIFaR -/- females (A) Percentage copulation over time for *SIFaRp65AD*$^{[MiMIC]}$ homozygous (*AD/AD)* (n=50), and control (*AD/+)* (n=41*), (+/+)*(n=37*)* females paired with wild-type males. **(B)** Percentage remating 24hrs after initial mating **(C)** Number of eggs laid within 24hrs of initial mating. Bars and error bars show means ± SEM. One-way ANOVA p = 0.0034, *p = 0.047, **p = 0.002 (Tukey's multiple comparisons).

5.3 Investigating the role of SIFaR signalling in *dsx* neurons

That loss of *SIFaR*, at least in males, results in 'ectopic' courtship behaviour, suggests a role for *SIFaR* signalling in the modulation of this behaviour. However, loss of SIFamide has also been shown to affect feeding and sleep (Park et al., 2014; Martelli et al.,2017). Thus, it is conceivable that abnormal courtship in *SIFaR* mutants arises secondarily to disrupted feeding and sleep patterns, rather than reflecting *direct* modulation of courtship circuits by the peptide. The use of transgenic approaches and analysis of scRNAseq data, however, suggests that *SIFaR* is expressed by neurons within these circuits, affording evidence of a direct modulatory mechanism. To ascertain whether SIFaR signalling in these neurons is required for wild-type male courtship behaviour, the receptor was selectively perturbed in *dsx* neurons via the expression of RNAi, driven by the *SIFaR-p65AD* and *dsx-DBD* hemidrivers. Though the effects of global SIFamide loss-of-function on female reproductive behaviour were less clear, reproductive behaviours were also assayed in females expressing the RNAi construct in this genetically intersected neuronal population.

5.3.1 Phenotypic screen for selection of a suitable RNAi construct

At the time of conducting these experiments, four RNAi constructs designed to target *SIFaR* were available from *Drosophila* stock centres: two from the Transgenic RNAi Project (TRiP) collection - *TRiP 25831* and *TRiP 34947*, which express long dsRNA hairpins; one from the Vienna GD stock library - *VDRC 1783*, which also expresses a long hairpin, and one from the Vienna shRNA stocks library - *VDRC 330383*, expressing a short hairpin microRNA (fig.6A). To test the efficacy of these lines, each *UAS-RNAi* line was expressed pan-neuronally under the control of *elav-GAL4* driver. Consistent with the findings of Agrawal et al. (2013) progeny expressing the *VDRC 1783* RNAi construct in all neurons reached the late pupal stage but did not eclose as adults. Interestingly, when using *Actin-GAL4* to drive expression, lethality in this line was observed earlier; in the larval stage, and pupal lethality was also observed in the *TRiP 25831* line (fig B). The finding that two separate *SIFaR*-directed RNAi constructs (from different RNAi collections, and with different target sequences) produce lethality when expressed in all cell types lends further support to the conclusion that *SIFaR* may serve an essential function in development. In addition, the delayed lethality of the *VDRC 1783* RNAi line when expressed under *elav-GAL4* control (versus *Actin-GAL4*), suggests that SIFaR signalling in non-neuronal tissues may contribute to this function. Three out of four *UAS-RNAi* lines were, however, viable when driven by *elav-GAL4*, permitting phenotypic analysis at the adult stage. Given that *SIFaR* mutant males display increased levels of male directed courtship, successful knockdown of SIFaR by RNAi should produce a similar phenotype. To test this, males expressing each viable RNAi construct under the control of *elav-GAL4* were paired with wild-type (CS) males and courtship behaviours were analysed. *SIFaRp65AD*$^{[MiMIC]}$ homozygous (AD/AD) males, were included in this experiment for comparison. Levels of courtship greater than those typically observed in wild-type males were displayed by all genotypes (fig.6B). Of the 3 RNAi lines tested, the highest levels of courtship, which closely matched those observed in *AD/AD* males, were seen in the *TRiP 25831* line. On this basis, the *TRiP 25831 SIFaR* RNAi line was selected for subsequent RNAi experiments.

A

		Lethality	
RNAi line	**Construct ID**	***Actin-GAL4 >***	***elav-GAL4 >***
TRiP 34947	TRiP.HMS00299	N	N
TRiP 25831	TRiP.JF01849	Late pupae	N
VDRC 330283	VSH330283	N	N
VDRC 1783	GD698	Pre-pupae	Late pupae

B

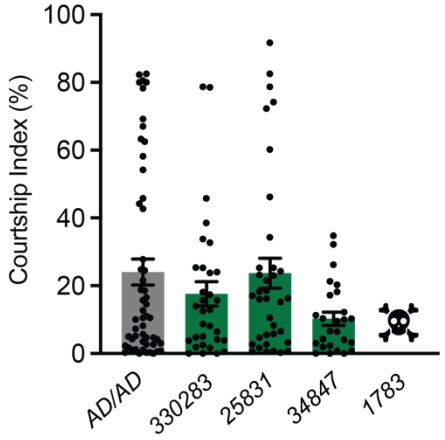


Figure 6: Screen of candidate RNAi lines for SIFaR knockdown experiments. (A) Table showing available RNAi lines targeting *SIFaR.* 'Lethality' column shows construct lethality when expressed in all cells (*Actin5C-GAL4)* or all neurons (*elav-GAL4*); 'N' = not lethal. **(B)** Courtship indices of *SIFaRp65AD[MiMIC]* homozygous males, and males expressing RNAi constructs under the control of *elav-GAL4,* paired with wild-type males. Bars and error bars show means ± SEM.

5.3.2 *SIFaR* knockdown in *dsx* neurons by RNAi affects reproductive behaviours

To target the *TRiP 25831 SIFaR RNAi* construct in *SIFaR∩dsx* cells, the *SIFaR-p65AD* and *dsx-DBD* hemidrivers were used to drive *UAS-RNAi* expression. When paired with wild-type males, males from this line showed significantly higher levels of courtship than control genotypes (fig.7A), suggesting that expression of *SIFaR* in *dsx* cells is required for normal male courtship behaviour. A reduction in time to copulation, a hallmark of increased sexual receptivity, was previously reported in SIFamide-less females (Terhzaz et al., 2007). In the current study, virgin females expressing RNAi against *SIFaR* in *SIFaR∩dsx* cells copulated with males more quickly than controls (fig.7B), suggesting that *SIFaR* expression in *dsx* cells is important in establishing wild-type levels of sexual receptivity in virgin females. Female post-mating behaviours were not analysed in this genotype, leaving open the question of whether loss of *SIFaR* in *dsx* neurons affects reproductive behaviour in mated females. Considered alongside the evidence that *SIFaR* Is expressed in dsx neurons, these behavioural findings provide strong evidence that SIFaR signalling in these circuits serves a function in modulating their functional outputs.

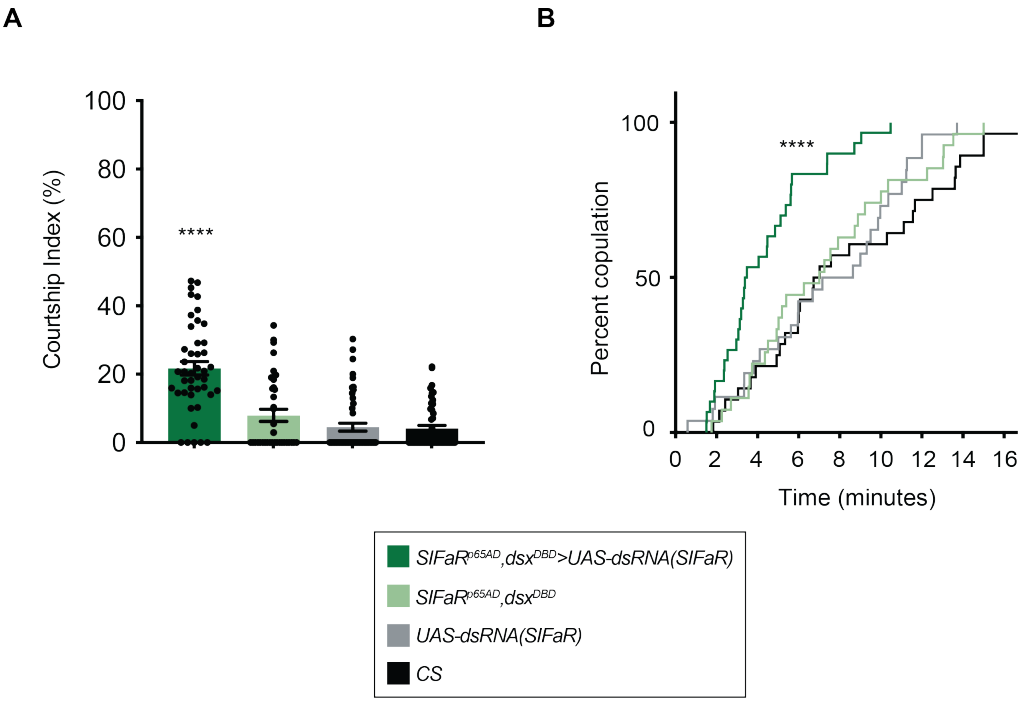


Figure 7: Reproductive behaviour phenotypes associated with RNAi- knockdown of SIFaR in *dsx* cells (A) Courtship indices of males expressing *UAS-dsRNA (Trip 25831)* against *SIFaR*, under the control of *SIFaR-p65AD,dsx-DBD* (dark green) and control males (light green, grey, black), paired with wild-type males. **(B)** Percentage copulation over time for virgin females expressing *UAS-dsRNA (Trip 25831)* against *SIFaR*, under the control of *SIFaR-p65AD,dsx-DBD* (dark green) and controls (light green, grey, black), paired with wild-type males. Bars and error bars show means ± SEM. One-way ANOVA $p < 0.0001$, ****$P < 0.0001$ (Tukey's multiple comparisons against all control genotypes)

5.4 Generation of a *SIFaR*-FlpSTOP allele

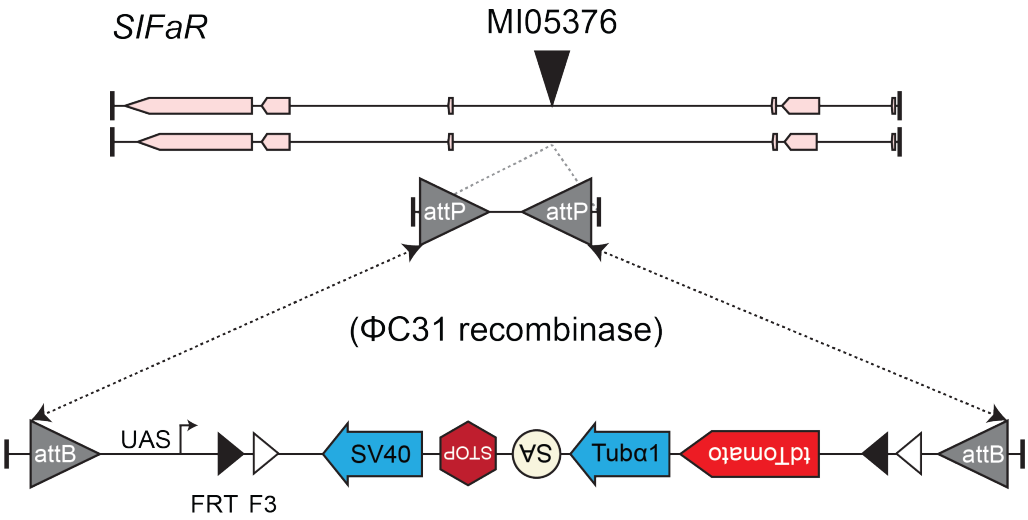
In recognising the limitations of RNAi as a method of gene silencing; namely incomplete gene knockdown and the possibility of off-target effects, it was of interest to devise an independent means of silencing *SIFaR* in genetically defined sub-sets of cells. For this purpose, a *SIFaR*-FlpSTOP allele (Fisher et al., 2017) was generated. *attB* sites at each end of the FlpSTOP cassette allow its insertion into an intronic MiMIC in the target gene. Once inserted, 'disrupting' elements in the cassette silence the target gene via two mechanisms. First, SV40 and Tuba1 transcriptional terminator sequences halt transcription. Second, the presence of an MHC intron 18 splice acceptor, followed by STOP codons in all reading frames, ensures that any remaining transcripts are not translated. MiMIC insertions can occur in both orientations. In the case of

the FlpSTOP cassette, the disrupting elements silence target gene expression in one orientation (the 'disrupting' (D) orientation) but not the other (the 'non-disrupting' (ND) orientation). FRT sites flanking the disrupting elements of the cassette allow the intervening region to be inverted by Flp-recombinase (Flp) protein, switching the cassette from the ND to D orientation, or *vice versa*. Thus, the target gene can be silenced in specific cells by placing the FlpSTOP(ND) allele in trans to a null allele of the target gene and expressing Flp specifically in those cells using, for example, the GAL4/UAS system. Likewise, the target gene can be rescued in a mutant background via the same approach but, in this case, utilising the FlpSTOP(D) allele. Where Flp expression is driven by GAL4, the addition of a *tdTomato* sequence, encoding a variant of RFP, within the cassette, and a UAS sequence *outside* the FRT sites, allows inversion of the cassette in target cells to be visualised (fig.8) (Fisher et al., 2017).

Insertion of the FlpSTOP cassette at the MI05376 intronic MiMIC in SIFaR was achieved by injection of the FlpSTOP plasmid into embryos from this MiMIC strain (injections performed by BestGene). Progeny with the insert in both orientations were isolated. PCR reactions, using primers for endogenous sequences flanking MI05376 and sequences within the FlpSTOP construct, identified the [+] orientation as the non-disrupting, ND, allele and the [-] allele as the disrupting, D, allele. Producing a line in which FlpSTOP could be used to silence *SIFaR* in *dsx* cells involved a complex crossing strategy to combine the necessary transgenes, namely: the FlpSTOP(ND) allele; UAS-Flp (to invert the FlpSTOP cassette); *dsx-GAL4* (use of the *SIFaR-65AD* allele was not appropriate due to its mutagenic affects), and $SIFaR^{del}$ (serving as the null allele of SIFaR, as per the recommended FlpSTOP protocol (Fisher et al., 2017)). Analysis of RFP fluorescence in this line (*UAS-Flp;dsx-GAL4,SIFaRdel/SIFaR(ND*)) revealed TdTomato expression in *dsx* cells (fig.9A), indicating successful inversion of the cassette in these cells. Animals of this genotype should be *SIFaR* mutants in *dsx* cells only. Thus, behavioural assays using this line could be used to corroborate RNAi experiments, which show that loss of *SIFaR* in *dsx* cells affects reproductive behaviour. However, preliminary behavioural experiments using the newly generated FlpSTOP allele produced questionable results: though levels of

male-directed courtship were as high as those observed in *AD/AD* males, or in the context of SIFaR knockdown by RNAi, similar levels were observed in the control genotypes (fig.9C). Moreover, the courtship indices of wild-type ‘target’ males were higher than those previously observed in this study. The same phenomenon was observed when testing males with both copies of the *SIFa*- FlpSTOP(D) allele, which should confer ubiquitous *SIFaR* LOF (fig.9B). At present, the basis of this effect remains unclear. One possibility is that, by the time these experiments were conducted, the wild-type (CS) strain to which the transgenic lines were crossed to generate control genotypes had become genetically contaminated.

A



B

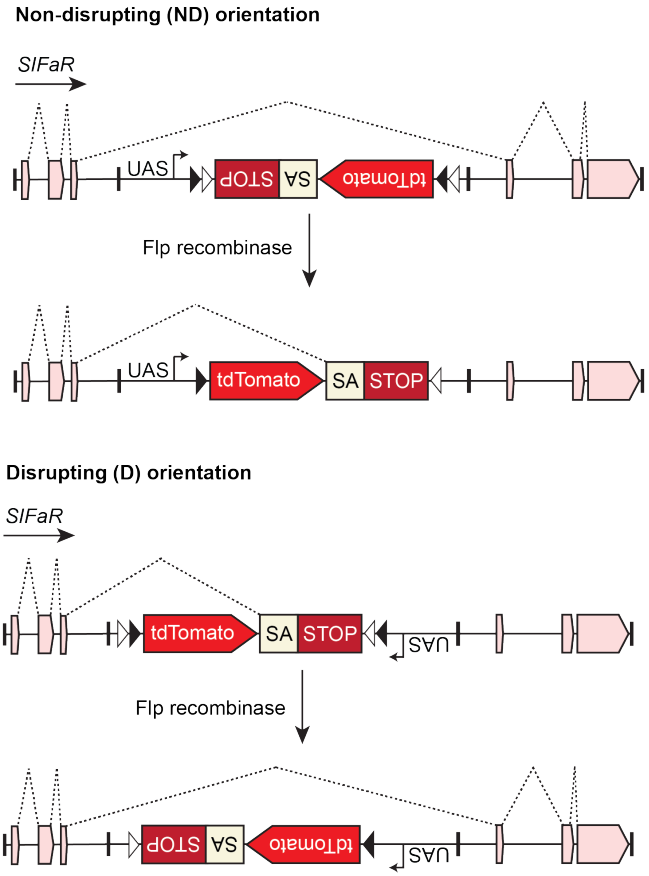


Figure 8: Generation of a *SIFaR*-FlpSTOP allele. (A) *SIFaR* locus, showing MiMIC site (MI05376) into which the FlpSTOP cassette was inserted, and overview of FlpSTOP cassette sequence. **(B)** Top: Inversion of the cassette from the (ND) orientation to the (D) orientation induces disruption of SIFaR expression. Bottom: Inversion from the (D) orientation to the (ND) orientation to 'rescue' SIFaR expression.

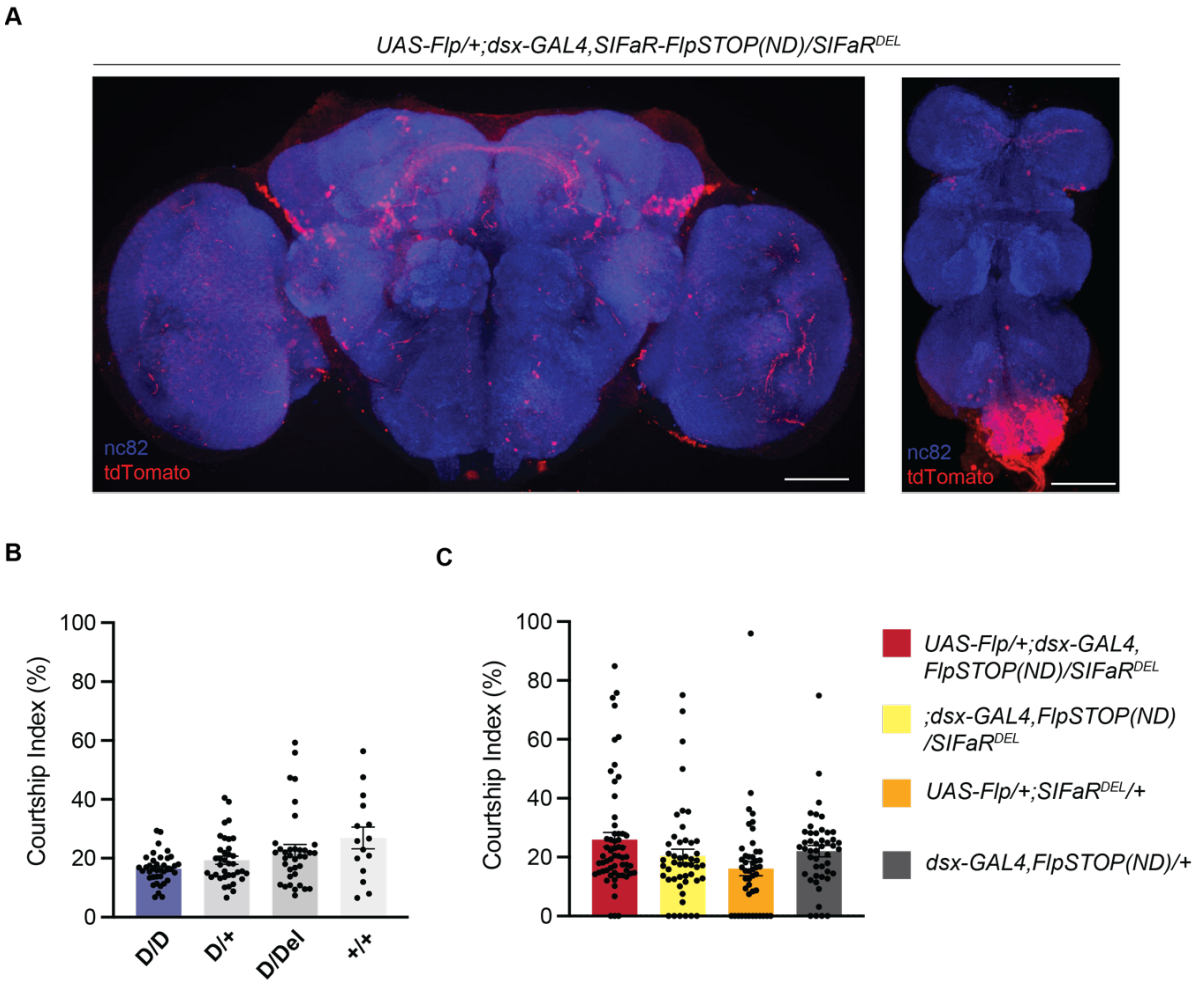


Figure 9: Preliminary validation of the *SIFaR*-FlpSTOP allele. (A) Adult Brain (left) and VNC (right) stained with nc82 (blue), and mCherry (red) to reveal TdTomato expression. (TdTomato-expressing cells are those in which the SIFaR-FlpSTOP cassette has been inverted by Flp mediated recombination). **(B)** Preliminary behavioural experiments using the *SIFaR*-FlpSTOP(D) allele. Courtship indices of males paired with wild-type males *(SIFaR*-FlpSTOP(D) homozygous = D/D, *SIFaR*-FlpSTOP(D) heterozygous = D/+, *SIFaR*-FlpSTOP(D) / $SIFaR^{Deletion}$ = D/Del, CS = +/+). **(C)** Preliminary behavioural experiments using the *SIFaR*-FlpSTOP(ND) allele to disrupt *SIFaR* in *dsx* cells: courtship indices of males paired with wild-type males. Bars and error bars show means ± SEM.

5.5 Quantifying SIFaR loss-of-function alleles

In the current study, various tools to disrupt *SIFaR* expression were generated and applied experimentally to investigate the role of SIFaR signalling in reproductive behaviour. The conclusions drawn from these experiments would clearly be substantiated by evidence showing an absence, or reduction, of SIFaR expression in these lines. The ability to quantify

SIFaR loss of function from these alleles could also elucidate some apparent inconsistencies in the data as it stands. For example, the fact that *SIFaR* loss-of-function from the *SIFaRp65AD*$^{[MiMIC]}$ allele only partially phenocopies loss of SIFamide, or that the *SIFaRp65AD-MiMIC* and *SIFaR*-FlpSTOP(D) lines are homozygous viable, despite evidence to suggest that *SIFaR* is required in development. A key motivator for producing SIFaR antiserum in the current study was for its application in this context. However, this possibility was ruled out by its lack of specificity, as revealed by IHC and WB assays. As an alternative, qPCR experiments were set up to quantify *SIFaR* mRNA levels in the various lines used in SIFaR LOF experiments (see chapter 2, Materials and Methods for protocol). However, the occurrence of multiple/non-specific qPCR products observed using a range of primer sets and conditions and a lack of stability in the expression levels of reference genes across genotypes precluded the use of qPCR data in quantifying *SIFaR* expression levels from these alleles.

5.6 Visualising SIFaR signalling in *dsx* circuits

5.6.1 Introduction

The expression of *SIFaR* in *dsx* neurons facilities direct modulation of these neurons by SIFamide. However, the mechanism by which SIFamide modulates target neurons, and the consequences of this modulation on neuronal activity, remain unclear. In the current study, two approaches were taken to visualise the effects of SIFaR signalling in target neurons. First, a FRET based method was used to visualise the signalling pathways that are directly activated when SIFamide binds to SIFaR in target cells (section 5.5.2). Second, *in-vivo* and *ex-vivo* calcium imaging was used to determine the ultimate consequences of SIFamide signalling on stimulus-evoked neuronal responses (section 5.5.3)

5.6.2 Pilot experiments to measure cAMP modulation via EPAC-FRET

By what mechanism does SIFamide affect neuronal physiology? Answering this question demanded closer examination of the pathways activated by SIFaR signalling. GPCR's, like SIFaR, can trigger a range of signalling pathways depending on the type of G proteins they activate (Hanlon & Andrew, 2015). Unpublished data provided by Erik Johnson, a collaborator from Wakeforest university CA, revealed that, in HEK cells expressing *SIFaR*, cAMP increases in response to Forskolin (Fsk) (a potent activator of adenylate cyclase), measured using a luciferase reporter, were attenuated by SIFamide (fig.10A). This suggests that SIFaR activation reduces cAMP levels and is, therefore, likely to be coupled to the G(i) pathway (fig.10B). To ascertain whether SIFamide has the same effect on *dsx* target neurons *in vivo,* a genetically encoded cAMP sensor *Epac-Camp1* (Nikolaev et al.,2004; Shafer et al.,2008*)*, was used. *Epac* encodes a protein complex in which a cAMP binding protein, exchange protein directly activated by cAMP (EPAC), is tethered to CFP and YFP fluorophores. Excitation of CFP triggers Förster resonance energy transfer (FRET), whereby energy is transferred from CFP to YFP, leading to emission at both wavelengths. The ratio of CFP/YFP emission is referred to as the FRET signal. Binding of cAMP to EPAC changes the confirmation of the protein, moving the fluorophores further apart and reducing FRET (fig10.C). Thus, increases in intracellular cAMP are detected as a reduction in the FRET signal, and vice versa. FRET imaging with the EPAC sensor has been used previously to show both increases and decreases in cAMP in neurons in the *Drosophila* brain in response to the addition of neurotransmitters and neuropeptides (for example: Schafer et al.,2008; Vecsey et al., 2014). Based on the hypothesis that SIFamide decreases cAMP levels, addition of SIFamide should increase the FRET signal in cells expressing the Epac construct.

In the current study, it was first necessary to determine whether EPAC expression could be induced *in vivo.* In animals expressing *Epac* under the control of the *SIFaRp65AD* and *dsx-DBD* hemidrivers, YFP fluorescence was observed in *SIFaR*∩*dsx* neurons in the brain, indicative of successful *Epac* expression in these cells (fig.10D). The majority of the studies

reporting FRET changes in neurons in the *Drosophila* brain utilised confocal fluorescence microscopy to measure these changes. In the current study, initial pilot experiments were, therefore, conducted using this set-up. However, FRET responses could not be reliably measured following the addition of compounds known to elevate cAMP levels. As a result, subsequent pilot experiments (data from which is shown in fig.10E-I), were conducted on a wide-field fluorescence microscope in the lab of Manuela Zaccolo; whose group routinely use this set-up for FRET based measurement of cAMP modulation. The wide range of factors that can obfuscate FRET measurements demands thorough validation of the experimental set-up, to ensure that 'real' FRET responses can be reliably measured in response to compounds known to modulate cAMP levels. Thus, pilot experiments involved measurement of FRET changes in response to manual addition of 3-isobutyl-1-methylxanthine (IBMX) and Fsk to the imaging solution. IBMX inhibits phosphodiesterases, which break down cAMP, while Fsk activates the enzyme producing cAMP. Thus, this cocktail strongly increases intracellular cAMP levels. Overall, data from these pilot experiments suggested that cells did respond to IBMX/Fsk with a moderate increase in cAMP (fig.10E). However, close analysis of the data highlighted a number of confounding factors that would need to be addressed going forward. The first issue related to 'background' fluorescence. Analysis of emission in CFP and YFP channels was indicative of a high degree of baseline signal from designated background regions. This could be explained either by tissue autofluorescence (a degree of which is inevitable in relatively thick tissue samples), or by the presence of out-of-focus neurites associated with expressing cells. Evidence of the latter is afforded by the fact that these background regions appeared to respond (show changes in emission) to the addition of IBMX/Fsk (fig.10G). In many cases, the background subtracted responses of expressing cells were masked by this effect. Moreover, the extent of these responses varied across different background regions (fig.10G). As such, FRET measurements in target cells varied significantly according to which background region(s) were (arbitrarily) selected. A second issue was the heterogeneity of responses seen in target cells: only some cells responded to the addition of IBMX and, after the subsequent addition of Fsk, FRET changes in both directions were

observed among cells within the target cluster (fig.10H) A third issue was the occurrence of movement artifacts (seen in fig.10F,H), resulting from the osmotically induced shrinking of cells within the tissue. This was most likely due to the presence of Dimethly sulfoxide (DMSO) in the applied solutions; a compound necessary to prevent Fsk and IBMX from precipitating in solution, but which significantly contributes to osmotic pressure. The relatively high concentrations of applied compounds needed to elicit measurable responses in pilot experiments could have, conceivably, contained DMSO levels that cells were unable to fully tolerate. A final issue highlighted by the analysis of pilot experiments related to the high pH sensitivity of the YFP fluorophore present in the EPAC construct. Under HEPES buffered conditions, intracellular pH tends to rise, due to dysfunction of bicarbonate transporters in these conditions. This unquenches the YFP fluorophore, resulting in an artificial rise in YFP emission (fig.10I). This effect can easily confound measurement of 'real' FRET changes.

Taken together, these confounding factors suggested a significant degree of optimisation would be needed to establish a set-up in which FRET signals could be reliably measured. The issue of background signal would likely be improved by switching to confocal imaging, though this could compromise the FRET signal strength, which was relatively quite weak in pilot wide-field experiments. Changes to the composition of the imaging solution, to counteract the osmotic effects of DMSO, could reduce cell movement and leave more cells in a viable state for imaging. In addition, the use of a perfusion system, rather than manual application of compounds via a pipette, could help to prevent large fluctuations in DMSO concentration and, also, limit movement artefacts. Finally, the pH sensitivity of the EPAC construct suggested that alternative buffers (for example, a bicarbonate-based buffer) that have smaller effects on pH, should be considered, or alternatively, a different FRET sensor used.

These modifications may have allowed robust FRET changes to be observed in response to the addition of IBMX/Fsk. However, the ultimate purpose of these experiments was to determine whether SIFamide modulates cAMP levels. The FRET changes associated with this modulatory effect are likely to be far more subtle. In addition, though inverse FRET changes,

corresponding to decreases in cAMP can be measured, they occur in a smaller dynamic range (Shafer et al., 2008) and are, thus, harder to detect. Considering these points, investigation into the physiological effects of SIFamide signalling was pursued via more established calcium imaging protocols.

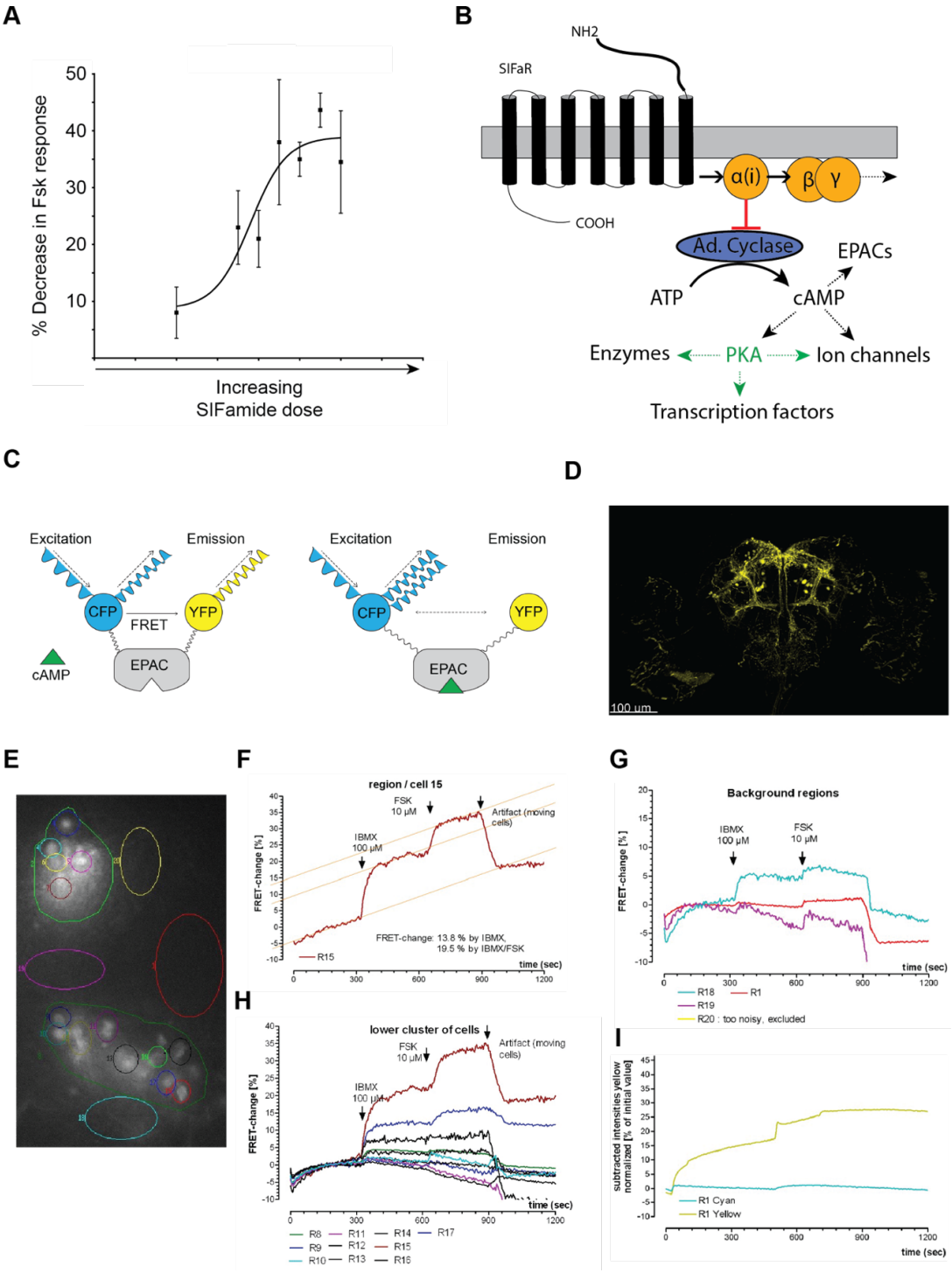


Figure 10: Pilot EPAC/FRET experiments. (A) From Erik Johnson (Wakeforest, NC): Dose-response curve for HEK cells expressing *SIFaR* stimulated with Forskolin (Fsk). (**B**) Schematic showing overview of predicted SIFaR signalling pathway. (**C**) Schematic showing how binding of cAMP to the EPAC alters FRET signal. (**D**) YFP expression (yellow) in animals expressing UAS-Epac1-camps under the control of *SIFaR-p65AD* and *dsx-DBD*. (*E-I*) Preliminary wide-field FRET imaging experiments: **(E)** Widefield image showing regions selected for analysis. **(F)** FRET change in one labelled cell in response to IBMX/Fsk addition. (**G**) Analysis of FRET change in 3 separate background regions (shown in (E)). (**H**) FRET responses of all cells selected for analysis. (**I**) CFP and YFP emission of a single background region, normalized to the initial emission value, showing 'YFP' drift; predicted to be due to unquenching of the fluorophore. Analysis in (F-I) was conducted by Andreas Koschinski (Zaccolo Group, Oxford)).

5.6.3 Calcium imaging to measure the effect of SIFamide on *SIFaR dsx* neuronal activity

5.6.3.1 Pilot *in-vivo* 2-photon calcium imaging experiments

The hypersexual phenotypes associated with perturbation of SIFaR in *dsx* neurons are indicative of altered activity in these circuits. In an effort to visualise and measure this, calcium imaging experiments were designed and validated. Genetically encoded calcium indicators (GECIs) are widely used as a proxy for neuronal activity measurements. GECI's utilise molecules engineered to emit fluorescence on binding to intracellular calcium. Since the concentration of intracellular calcium increases when neurons depolarise, via the opening of voltage gated- calcium channels, a change in this fluorescence provides a readout of neuronal activity (Venken et al., 2011a). In the current study, expression of the GECI GCaMP was induced in *SIFaR∩dsx* neurons using S*IFaR-p65AD and dsx-DBD* to drive expression of *UAS-GCaMP6(s).* GFP-labelled somata and neurites in the brain and VNC indicated successful expression of GCamp in the target cells (fig.11A). Initially, pilot experiments were conducted using *in-vivo* 2 photon calcium imaging. With this approach, neuronal responses can be recorded in awake animals during the performance of specific behaviours, or in response to various stimuli. For example, this method has been used to visualise activity in male P1 interneurons; a subset of the sexually dimorphic *dsx* pC1 cluster, activation of which drives most aspects of male courtship behaviour; including wing extension (Kimura et al., 2008, Kohatsu et al., 2011). *In-vivo* calcium imaging has revealed that P1 neurons are activated when the male touches the abdomen of the female; specifically, by contact pheromones present on her abdomen, which activate gustatory neurons in the male's foreleg (Kohatsu et al., 2011; Inagaki et al., 2014; Kohatsu & Yamomoto, 2015; Clowney et al., 2015). Thus, P1 is considered to be a key locus for courtship initiation in response to the integration of relevant sensory stimuli. In the current study, when placed with wild-type males, males expressing RNAi against *SIFaR* displayed courtship behaviours, including wing extension. Thus, in-vivo calcium imaging experiments were designed to test whether this correlates with greater activation of

P1 neurons in this context. Initially, however, it was important to establish an experimental set-up in which GCaMP activity in P1 neurons could be reliably measured in awake, head-fixed males. Thus, in pilot experiments, *SIFaR-p65AD,dsx-DBD>UAS-GCaMP6(s)* males were presented first with an immobilised virgin female: a stimulus that should evoke a significant response. Fluorescence measurements were recorded from the dendritic arbors of GFP-labelled cells thought to constitute a sub-set of the *dsx* pC1 cluster (fig.11A). Responses were observed in some males. However, there was often a weak stimulus-response correlation: only some female touches evoked a GCaMP signal and responses were sometimes observed in the absence of an obvious touch (example traces shown in fig.11E). Rather than reflecting a 'real' stimulus-response mismatch, this could simply reflect the difficultly of annotating stimulus frames in the data. For example, following the presentation of the female, males often continued to extend a foreleg long after the female had been retracted. In other cases, it was unclear whether contact with the female was due to active extension of the foreleg or 'passive' contact with the female. In general, the activity levels of males, and their propensity to produce 'active' touches of the female, varied significantly according to how well they recovered from the process of anaesthesia, dissection, and head fixation prior to imaging. Perhaps because of this, though only males capable of stimulus-evoked movement were tested experimentally, a paucity of a neuronal responses was observed in a significant proportion of males. Improvements to the experimental protocol and set-up could, conceivably, address some of these issues. However, given that SIFaR signalling may only modulate slightly the activity of dsx circuits, devising an alternative experimental paradigm, in which more robust calcium responses could be reliably measured, was deemed preferable.

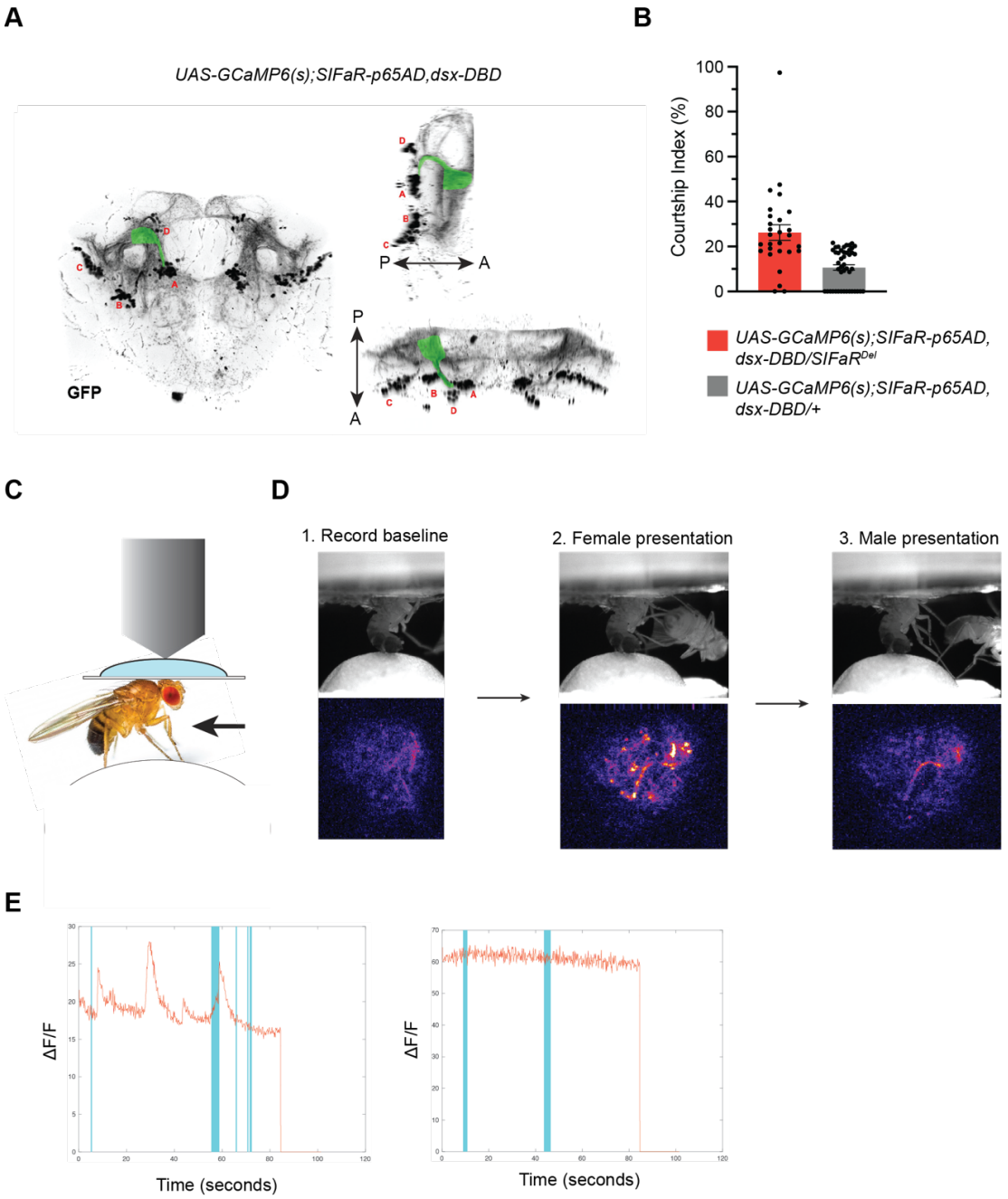


Figure 11: Pilot *in-vivo* 2-Photon calcium imaging experiments. (**A**) Maximal Z-projection of brain expressing *UAS-GaMP6(s)* (GFP fluorescence in black) under the control of *SIFaR-p65AD* and *dsx-DBD*, in coronal (left), sagittal (top right) and transverse (bottom right) planes. Fluorescence measurements were recorded from the neurites (highlighted in green) of cluster 'A' neurons; thought to be a sub-set of *dsx* pC1. (**B**). *AD/Del* (*SIFaR* mutant) males expressing *UAS- GCaMP6(s)* under the control of *SIFaR-p65AD* and *dsx-DBD* court wild-type males at increased levels. **(C)** Schematic to show experimental set up allowing presentation of stimuli (arrow) to head fixed male fly during imaging (**D**) Overview of experimental procedure showing stills from videos recorded to capture stimulus presentation (black and white) and stills from fluorescence image capture at the same time points, coloured according to GFP fluorescence (dark purple; low to white; high). (**E**) Example fluorescence measurements from individual experiments. Red trace shows normalised fluorescence over time, blue bars correspond to 'active' touches of the immobilised female by the head-fixed male. Left: some small responses observed, but weak stimulus-response correlation. Right: example of a 'non-responding' male.

5.6.3.2 Pilot *ex-vivo* wide-field calcium imaging experiments

Switching from an *in-vivo* to an *ex-vivo* set-up for calcium imaging alters the nature and scope of hypotheses that can be tested. Clearly, it is not possible to measure neuronal responses to 'naturalistic' stimuli in an *ex-vivo* preparation. However, the central question of whether and how a molecule modulates the excitability of a circuit can still be addressed. In addition, an *ex-vivo* preparation gives better access to the target neurons and permits greater control over the initial activation of the circuit. For example, using wide-field calcium imaging in explant brains, Barnsdedt et al. showed that the neuropeptide sNPF potentiates the response of M4/6 mushroom body output neurons (MBONs) to Acetyl choline (Ach) (Barnstedt et al.,2016). In the current study, several experiments were designed to assay the modulatory effects of SIFamide on *SIFaR∩dsx* neurons, based on a similar set-up. In the first, neuronal responses to application of Ach; thought to excite a significant proportion of *dsx* neurons, are analysed in samples co-expressing RNAi against *SIFaR* and GCaMP in *dsx* cells, to test whether their excitability is affected by SIFaR loss-of function. In the second, responses to Ach are compared with and without the addition of synthetic SIFamide, to investigate how the peptide modulates the excitability of its target neurons in the adult brain over different time scales (fig.12C). An example of the latter experiment was carried out by Barnstedt et al, in their study of sNPF function. In this case, sNPF was co-applied with Ach. However, while sNPF is thought to be co-released with FANs to modulate their effects at the synapse, SIFamide is presumed to exert a slower but, potentially, longer-lasting effect on neuronal physiology. Thus, it may be preferable to add SIFamide directly to the imaging solution, allowing sufficient time for the peptide to elicit its modulatory effect. The disadvantage of this approach is that the action of the peptide is not limited to *SIFaR∩dsx* neurons from which Ach-evoked responses are subsequently recorded. Thus, changes in the response properties of *dsx* neurons could reflect altered input from other neurons modulated by globally applied SIFamide. However, if these changes were absent in samples expressing dsRNA against SIFaR in *SIFaR∩dsx* neurons, this would offer evidence of a direct modulatory effect of the peptide on these cells. As with *in-*

vivo calcium imaging experiments, it was first important to establish a paradigm in which *dsx* neuronal activity could be reliably measured. To allow local delivery of Ach to GCaMP labelled neurites, a microinjection system was set up to facilitate the ejection of small quantities of solution to the target region. The addition of a red fluorescent dye to the applied solution provided visual evidence of successful ejection and allowed the diffusion of applied compounds to be monitored (fig.12B). An LED was also set up and programmed to deliver a pulse concurrent with pressure delivery to the injection system, allowing this event to be detected in the GFP imaging channel. Preliminary experiments, conducted in explant brains expressing GCamp6(s) in SIFaR dsx neurons, have yielded promising results, with evidence of calcium responses to applied Ach. However, at this stage, achieving an ejection pressure that does not move or damage the neurons, but which delivers a volume sufficient to elicit a response, still presents a challenge and, as such, there remains significant variability in the quality of the acquired imaging data. With further optimisation, however, this paradigm offers an opportunity to visualise the modulatory effects of SIFamide on neuronal activity.

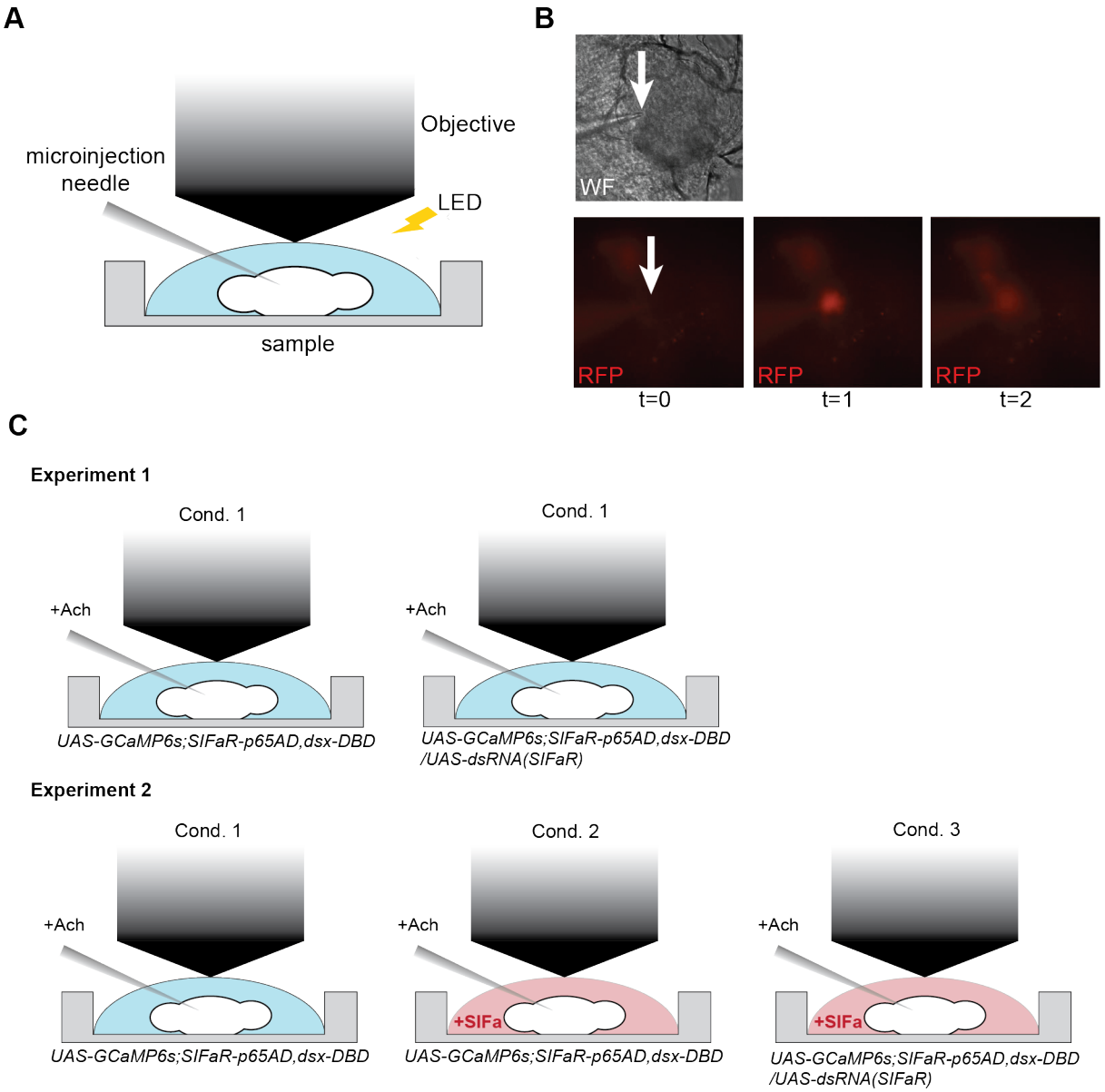


Figure 12: Pilot *ex-vivo* wide-field calcium imaging experiments. (A) Schematic showing overview of experimental set up, (**B**) Top: Bright-field image showing top of microelectrode (white arrow) positioned adjacent to GCaMP-expressing *SIFaR∩dsx* neurons in the Abdominal ganglion. Bottom panel: imaging in the red fluorescence channel showing successful ejection of solution containing TX-red fluorescent dye. (**C**) Theoretical experiments to assay the effects of SIFamide on activity in *SIFaR∩dsx* neurons. Experiment 1 assays whether SIFaR LOF affects the responses of *SIFaR∩dsx* neurons to Acetyl choline (Ach). Experiment 2 assays whether these responses are modulated in real-time by the presence of SIFamide (Cond.2 vs Cond.1), and whether this is mediated by signalling through SIFaR (Cond.3).

5.7 Summary and discussion

A true null allele for a gene of interest (GOI) is an invaluable tool: it can be used to assess the importance of the GOI in specific functions, provide a mutant background for other loss-of-function (LOFs) alleles, and serve as a negative control in assays designed to test the efficacy of these alleles. A *SIFaR* null allele was generated in the current study by Flp mediated recombination between two TE's flanking the *SIFaR* gene, producing a deletion of the entire *SIFaR* gene (*SIFaR*^{deletion}). Assays of viability in this line revealed this deletion to be embryonic lethal. Since this deletion encompassed two other genes, it is conceivable that the lethality does not reflect an essential function for SIFaR *per se*. However, as discussed, several independent lines of evidence suggest the lethality may indeed stem from an essential function for *SIFaR* in development. To confirm this however, a more 'targeted' *SIFaR* null allele would need to be generated, for example via CRISPR-Cas9

Adult reproductive behaviours were assayed in animals harbouring two copies of a mutagenic MiMIC insertion in *SIFaR* (*SIFaR-p65AD*$^{[MiMIC]}$*).* The phenotypes observed in these animals support the conclusion that wild-type reproductive behaviours require *SIFaR* expression and signalling. Since these animals are global *SIFaR* mutants, however, it is not clear from this finding *where* SIFaR signalling occurs to modulate reproductive behaviour. To ascertain whether SIFaR signalling specifically in *dsx* neurons is important for normal reproductive behaviour, these behaviours were analysed in animals expressing *SIFaR-RNAi* in these cells. The existence of reproductive behaviour phenotypes in this context was a key finding in this study: affording evidence that SIFamide signals directly through the neural circuits governing reproductive behaviours to modulate their outputs. Interestingly, evidence from the expression pattern of the *SIFaR-p65AD* line, used in conjunction with *dsx-DBD,* was indicative of *SIFaR* expression in many neurons throughout all levels of *dsx* circuit architecture: from neurons in the periphery (notably the foreleg and female reproductive tract), to ascending and descending neurons in the VNC, and higher-order brain neurons (for example, those in the PC1 cluster), which integrate information and initiate appropriate behavioural responses. Thus, while these

RNAi experiments suggest a role for SIFaR signalling in cells, exactly where in the circuit this signalling is important remains unclear. To shed light upon this, alternative DBD hemi-driver lines, which express in smaller sub-sets of *dsx* cells could be used in conjunction with the *SIFaR-p65AD* line to knockdown *SIFaR* in more restricted *dsx* neuronal populations.

In addition, as RNAi constructs do not confer temporally restricted knockdown of expression, it remains uncertain as to *when* the expression of *SIFaR* in *dsx* neurons is important for engendering wild-type reproductive behaviours. For instance, SIFaR signalling could occur during development, as these neural circuits are established, to set their 'basal' response properties, with no modulation of the circuit by SIFamide occurring thereafter. Alternatively, SIFamide signalling in these neurons could serve a context-specific function in the adult nervous system. The *SIFaR*-FlpSTOP line was generated in the current study to investigate these hypotheses further. However, further characterisation of this line is needed (specifically to ensure that it ensure that it engenders *SIFaR* loss-of-function) before its application to behavioural assays.

In-vitro assays can provide insights into how signalling through GPCRs, the targets of neuropeptides, occurs. For example, unpublished data for SIFaR ectopically expressed in HEK cells, suggests that this receptor may be coupled to Gα(i) pathways that decrease cAMP. However, whether this is also the case *in vivo*, where the complement of proteins and second messengers present in neurons, and their downstream interactions, may be far more complex, remains unclear. Methods which permit direct imaging of second messengers, for example FRET-based methods to measure cAMP fluctuations, can be used to investigate this. Further optimisation of the FRET protocol outlined in this study may permit the direct modulatory effects of SIFamide to be assayed via this methodology. The consequences of NPR signalling are, however, ultimately realised in changes in the physiological properties of target neurons, reflected in their responses to normal inputs. Thus, imaging the activity of *dsx* neurons either in the context of SIFaR LOF, or following the addition of the peptide, may provide insights into the normal modulatory action of SIFamide on these neurons. The *ex-vivo* calcium imagining set up piloted in the current study could be used for such a purpose.

Chapter 6: Elucidating SIFamide Signalling Mechanisms and Functions in the Adult Nervous System

6.1 Introduction

A wide range of phenotypes have been reported as consequences of disrupted *SIFa/SIFaR* signalling (summarised in table 1): altered reproductive behaviour (Terhzaz et al., 2007, Sellami & Veenstra, 2015, Wong et al., 2018); circadian rhythms (Cavanaugh et al., 2014); sleep duration (Park et al., 2014, Bai et al., 2018, Huang et al., 2018), and feeding (Martelli et al., 2017).

A key question is how SIFamide; a peptide expressed in just 4 neurons, performs these integrative functions. Due to the interaction between the neural circuits underlying these behaviours, disruption of a molecule implicated in just one of them is likely to produce a 'compound' phenotype, in which the remaining behaviours are affected (Nässel, Pauls & Hutteroth, 2019). The broad expression of *SIFaR* evidenced in the current study, however, suggests that SIFamide may be involved in the direct modulation of each of these behaviours. In support of this hypothesis: *SIFaR* is expressed in the *fru* and *dsx* neurons underlying sexual behaviour; there is evidence of SIFamide signalling in antennal lobe (AL) neurons implicated in feeding behaviour (Martelli et al., 2017); while Park et al. found that sleep is disrupted by *SIFaR* knockdown in a sub-set of cells in the pars intercerebralis (PI) - a region known to contain sleep-regulating neurons (Crocker and Sehgal, 2010).

Evidence of SIFa/SIFaR signalling across these distinct functional areas itself raises intriguing questions: does signalling occur simultaneously in these disparate locations or does SIFamide act on discrete targets at different points in time? If the former, what contexts trigger this 'global' release of the peptide? If the latter, how is signalling directed to specific target cells? This chapter considers several approaches that may help in addressing these questions. Section 6.1 considers what clues SIFamide distribution, and the morphology and connectivity of *SIFamide* neurons, can provide as to the functions of the peptide in the adult CNS. Section 6.3 presents preliminary data from experiments designed to test a specific hypothesis: that SIFamide simultaneously potentiates feeding and suppresses courtship in the context of starvation.

Genotype	(Intended) Effect(s)	Temporal dynamics	Phenotype(s)	Study
SIFaGAL4> UAS-hid/rpr	*SIFa* neuron ablation	constitutive	♂ - Indiscriminate courtship; ♀ - Increased sexual receptivity	Terhzaz et al., 2007
			Decreased total sleep (♀ and ♂). (Normal wake locomotor activity)	Park et al., 2014
			Circadian rhythm disruption (♂ only tested)	Cavanaugh et al., 2014
SIFa-CRISPR knockout	*SIFa* gene silencing	"	Decreased total sleep (♂ only tested) (Circadian rhythmicity normal).	Bai et al., 2018
SIFaGAL4> UAS SIFa-RNAi	*SIFa* knockdown in *SIFa* cells	"	♂ - Indiscriminate courtship; ♀ - Increased sexualreceptivity	Terhzaz et al., 2007
elavGAL4> UAS SIFa-RNAi	pan-neuronal *SIFa* knockdown	"	Decreased total sleep (♀ only tested). (Increased wake locomotor activity)	Park et al., 2014
			Disruption of SMD & LMD (♂)	Wong et al., 2018
elavGAL4> UAS SIFaR-RNAi	pan-neuronal *SIFaR* knockdown	"	Decreased total sleep (♀ only tested). (Increased wake locomotor activity)	Park et al., 2014
SIFaRGal4> UAS SIFaR-RNAi	knockdown of *SIFaR* in *SIFaR -GAL4* cells	"	♂ - Indiscriminate courtship (phenotypic variability noted) (♀ not tested)	Sellami & Veenstra,2015
SIFaR-p65AD [MiMIC]	*SIFaR* gene silencing	"	♂ - Indiscriminate courtship (Phenotypic variability noted)	Current
SIFaR-p65AD, dsx-DBD> UAS SIFaR RNAi	knockdown of *SIFaR* in *SIFaR* + *dsx* cells	"	♂ - Indiscriminate courtship; ♀ - Increased sexual receptivity (phenotypic variability noted)	Current
121Y-GAL4 and c767-GAL4 > UAS-SIFaR RNAi	knockdown of *SIFaR* in PI cells	"	Decreased total sleep (♀ only tested).	Park et al., 2014
SIFaGAL4> UASdTrpA1	*SIFa* neuron activation	For assay duration (~1 hour)	Increased attraction to EA (attractant food odour); increased sucrose PER; increased food consumption(♀ only tested).	Martelli et al., 2017
SIFaGAL4> UASChrimson	*SIFa* neuron activation	1 hour	Transiently increased sleep (♀ - rapid and stronger effect); ♂ - weaker, delayed effect)	Huang et al., 2021

Table 1: Summary of phenotypic consequences of SIFa/SIFaR manipulation observed in different studies

6.2 Investigating SIFamide neuron morphology and connectivity.

A biological system's structure often reveals much about its function. This concept is central to the study of neural circuits, but also to that of single neurons. An example of this is afforded by a recent study investigating the role of specific *dsx* neurons in sexual behaviour. The authors identified a sexual dimorphism in the dendritic arborisations of aDN *dsx* neurons. In males only, these dendrites innervate the AOTu; an area to which neurons encoding visual information project, while in females they innervate regions associated with other sensory modalities, in particular the lateral horn (LH), which is involved in higher olfactory processing. And these morphological characteristics were found to have predictable functional consequences: in males aDN's relay visual information that guides courtship, while in females they function in the selection of egg-laying sites by conveying relevant olfactory cues (Nojima et al., 2021). The caveat of applying this logic to peptidergic neurons is that signalling can occur at a distance from release sites (Nässel, 2009). As such, a neuropeptide's function is not constrained by the morphology or connectivity of the neuron(s) that release is. However, information about morphology and connectivity is not entirely uninformative. For example, the signals to which peptidergic neurons respond will often include fast acting neurotransmitters (FANs), released by upstream synaptic partners. Thus, the location and morphology of dendrites may reveal the type of information that is encoded by the activity of a peptidergic neurons. In addition, the arborisation of peptidergic axons presumably serves some functional purpose. This purpose could, of course, relate to a co-released FAN or monoamine. It could, however, reflect of the fact that while volume transmission is *possible* for neuropeptides, a significant degree of signalling occurs locally to release sites (Van den Pol, 2012). Considering these points, an analysis of SIFamide expression and SIFamide neuron morphology and connectivity was conducted in the current study, drawing on evidence from immunostaining, genetic reagents and connectomic data.

6.2.1 SIFamide expression in the CNS

SIFa-GAL4 driven GFP and SIFamide immunoreactivity have been used separately to examine SIFamide expression in the CNS (Terhzaz et al., 2007, Martelli et al., 2017). These methods were utilised in parallel in the current study, by staining the CNS of *SIFa-GAL4>UAS-mCD8::GFP* flies with anti-SIFa antisera. Consistent with previous reports, *SIFa-GAL4* drove GFP expression in 4 neurons in the pars intercerebralis (PI) region of the brain (fig. 1A), the somata of which showed strong anti-SIFa immunoreactivity (fig. 1B). The GFP-labelled processes of these neurons showed complex and extensive arborisations throughout the CNS (fig. 1A). Anti-SIFa immunoreactivity was evident in almost all regions of the CNS innervated by GFP-labelled neurites. This suggests that, at a given time point, the peptide is distributed throughout the neurites of *SIFamide* neurons in the adult brain, facilitating signalling through target cells in disparate CNS compartments, simultaneously. In line with previous reports (see Martelli et al., 2017), anti-SIFa immunoreactivity was particularly strong in regions of the superior-medial and superior-lateral protocerebrum (SMP & SLP, respectively), antennal lobes (AL) and around the dorsal fan shaped body (DFSB). Whether the signal in these regions is derived from *SIFamide* neuronal processes or peptide that has been released is unclear, though either interpretation points to a relatively high degree of release in these regions. Dense innervation of the SMP was indicated by *mCD8::GFP* expression, yet in more medial regions of the SMP anti-SIFa immunoreactivity was relatively weak, suggesting these processes are predominantly dendritic (*SIFamide* neuron input regions are considered in more detail in the next section). As per Terhzaz et al (2007), no gross difference in the morphology of SIFamide neurons was evident between the sexes, though subtle differences (with potential functional implications) cannot be ruled out.

In the current study, *SIFamide-GAL4* appeared to drive much weaker GFP expression in additional cells in the PI region, in which anti-SIFa immunoreactivity was not observed (fig.1C,D). This has not been reported previously. The extent of this ‘ectopic’ expression varied according to which *UAS*-reporter line was used. For example, the additional cells were more

strongly labelled by a *[20X]UAS-mCD8::GFP* reporter (fig.1E,F) and, interestingly, also when using the *UAS-CaLexA* reporter in later experiments to measure *SIFamide* neuron activation (see fig.3, section 6.3.1). However, when *SIFamide-GAL4* was used to drive a reporter localising to the nucleus, only 4 labelled cells were evident. Thus, the ectopic expression could well reflect some 'leaky' expression in the reporter lines. Alternatively, the *SIFamide-GAL4* line, which was generated by fusion of the *GAL4* coding sequence to a putative *SIFamide* promotor (Terhzaz et al., 2007), may capture some off-target transcription. Finally, this expression could reflect *SIFamide* transcription in these cells occurring at a minimal level and/or for a brief time frame, which does not persist in the adult brain.

The question of whether SIFamide is produced only in brain neurons is an intriguing one. Originally, the 4 PI neurons were identified as the only cells producing SIFamide (Terhzaz et al., 2007). However, subsequently, anti-SIFa immunoreactivity in putative somata in the VNC was observed (Martelli et al., 2017). In the current study, anti-SIFa immunoreactivity was generally weaker in the VNC than in the brain and there was a higher degree of background signal (fig. 1A). Nonetheless, more intense 'rings' of anti-SIFa immunoreactivity, approximately the size of neuronal somata, were evident in the Abg region of the VNC (fig. 1C). However, corresponding structures were not GFP-labelled by *SIFamide-GAL4>UASmCD8::GFP* (fig. 1C). Moreover, cell nuclei were not visible in the Abg when *SIFamide-GAL4* was used to drive a nucleus-localising reporter. Thus, it is not clear whether this immunoreactivity reflects SIFamide-producing cells or, else, an accumulation of the peptide at axon terminals in the Abg. Interestingly, *SIFamide* transcripts are detected in the scRNAseq VNC data set. However, very few transcripts are present in expressing cells, which is inconsistent with the evidence that neuropeptides are, generally, very highly expressed genes (Allen et al., 2019). These transcripts could, therefore, reflect capture of ambient *SIFamide* transcripts in non-expressing cells.

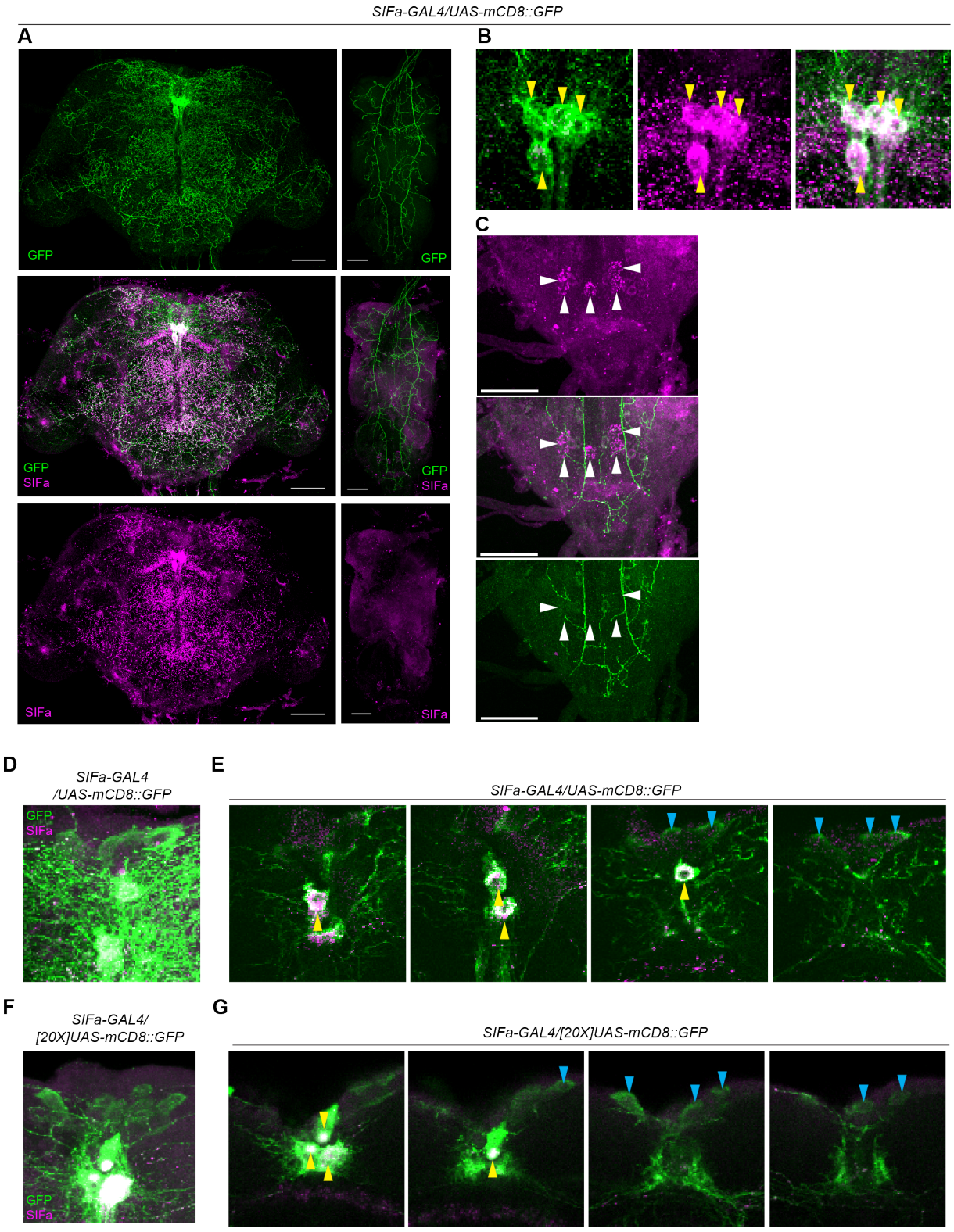


Figure 1: SIFamide expression in the CNS. (**A**) Brain (left) and VNC (right) of *SIFa-GAL4>UASmCD8::GFP* female, stained with anti-GFP (green) and anti-SIFamide (magenta). (**B**) Close up of region containing somata of *SIFamide* neurons (yellow arrows) (**C**) Close up of VNC Abg region. White triangles = *putative SIFamide* somata. (**D**) Region as shown in (B) (**E**) Individual optical sections of(D), showing GFP and anti-SIFamide co-labelled somata; yellow triangles, and somata labelled weakly by GFP only; blue triangles. (**F & G**) Same as (D & E), from *SIFa-GAL4>[20x]UASmCD8::GFP* female. Allimages (except E & G) are maximal z projections. Scale bars = 50µm.

6.2.2 Exploring heterogeneity in SIFamide neuron morphology and connectivity

The close proximity of the 4 SIFamide neurons, and the complexity of their neurites, prevents the morphology of each from being distinguished. The existence of heterogenous morphology could have important functional implications; allowing each neuron to modulate separate functions by integrating different information and/or acting on different target cells. The ability to analyse individual SIFamide neuron morphology could also reveal sex-differences that were previously indistinguishable. The possibility of sex differences is made particularly intriguing by the fact that SIFamide's role in modulating feeding is established only in females (Martelli et al., 2017); many of the sleep phenotypes associated with SIFamide perturbation have also been shown only in females (Park et al.,2014); and a potentially sexually dimorphic function in sleep regulation has been identified (Huang et al., 2021).

In the current study, initial attempts to map the morphology of individual SIFamide neurons utilised the Flybow system, which permits stochastic multi-colour labelling of cells within a genetically related population (Shimosako et al., 2014). However, despite numerous modifications to the protocol, progeny in which individual SIFamide neurons were heterogeneously labelled could not be produced via this approach.

The subsequent release of the *Drosophila* hemibrain connectome offered an alternative strategy for investigating the morphology and connectivity of individual SIFamide neurons. The connectome is a reconstruction of electron micrograph (EM) images of one hemisphere of the female central brain and contains information about the morphology and connectivity of ~22,000 neurons (Scheffer et al., 2020)(fig. 2A). All 4 SIFamide neurons have been annotated in this reconstruction. Visualisation of 3D reconstruction, or 'skeletons', of each neuron reveals that the positions of their cell bodies; at the anterior-dorsal midline, between the SMP neuropils of each hemisphere, and the gross morphology of their neurites are consistent with the distribution of both anti-SIFamide antisera and *SIFa-GAL4* driven GFP (fig. 2 B).

Analysis of each skeleton in parallel suggests that they constitute 2 pairs of neurons, with the neurites of each pair predominating on one side of the brain (fig. 2C-F). Intriguingly,

comparison of the two neurons innervating the mapped side of the hemibrain (ID:1418618235 – ‘SIFa-1’ and ID:699031185 - ‘SIFa-2’) highlights distinct morphologies (fig. 2E,F). This refers both to heterogeneous innervation of the same brain regions; seen in the AL, where both neurons show dense innervation, yet their neurites are located in largely segregated sub-compartments, and innervation of brain compartments by only one of the pair. For example, the Calyx (CA) and lateral horn (LH) is extensively innervated by SIFa-2, but SIFa-1 neurites are almost entirely absent in these regions. Conversely, SIFa-1 innervates all 3 lobes of the mushroom body (MB: aL,bL,gL), whereas SIFa-2 is confined almost exclusively to the aL of the MB (fig. 2E,F).

As well as allowing visualisation of neuronal morphology, the hemibrain data provides information about where input and output synapses are located. Some input and output regions are common to both mapped SIFamide neurons. For instance, for SIFa-1 and SIFa-2, the AL and SMP are the regions of greatest input and, for both neurons, a significant proportion of output synapses are located in the AL. This is consistent with the distribution of anti-SIFamide, and GAL4-driven expression of dendritic markers (see Martelli et al., 2017). However, there are some clear differences. For example, many SIFa-2 input and output synapses are located in the LH: a region largely devoid if innervation by SIFa-1. Thus, these neurons are capable of integrating different information and modulating different targets.

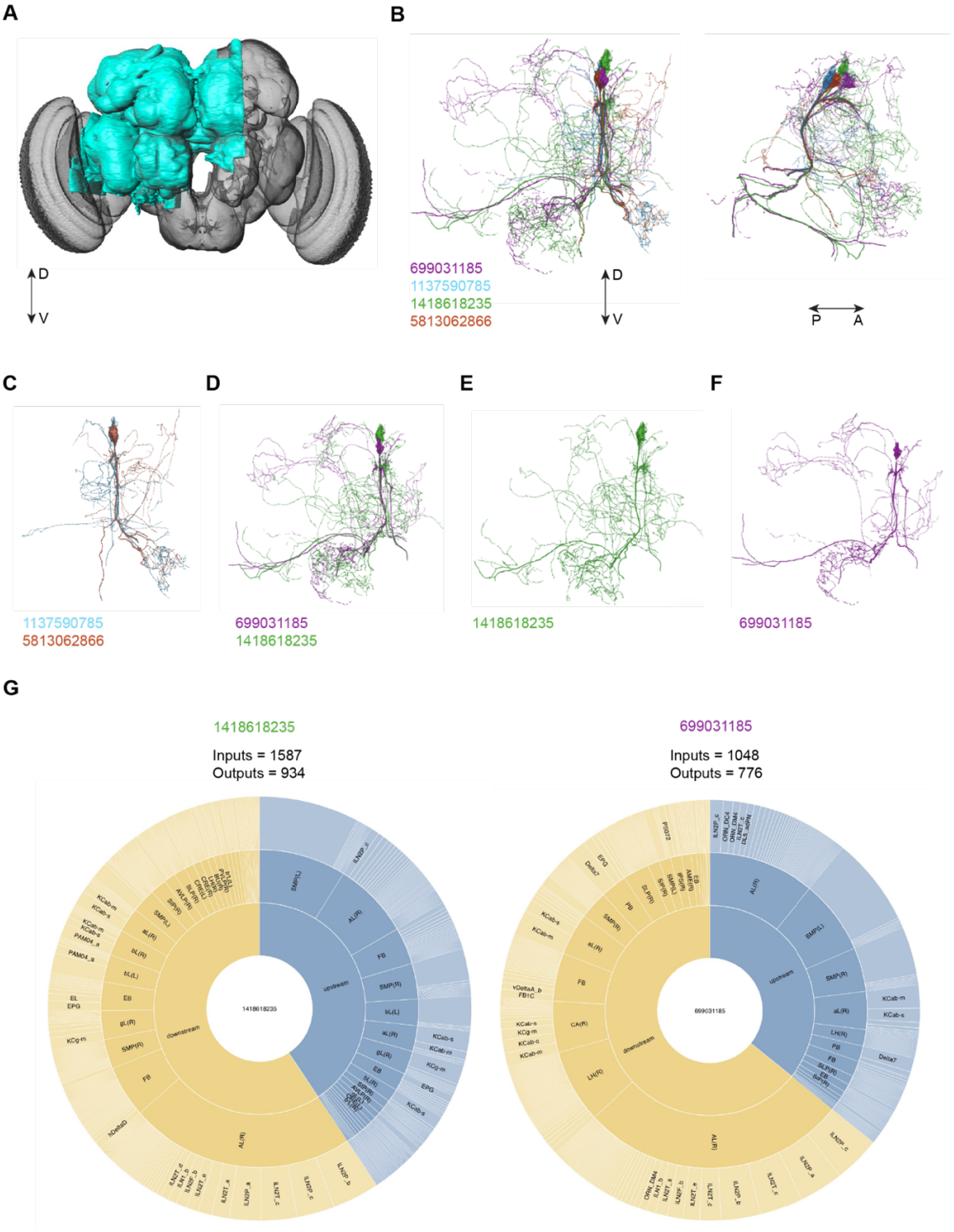


Figure 2: SIFamide neuron connectome data (A) (Adapted from Scheffer et al., 2020) Representation of the female brain (grey), showing region imaged and reconstructed in the creation of the connectome data set (turquoise). (**B**) 'Skeletons' of all 4 SIFamide neurons, from coronal (left) and sagittal (right) views. (**C**) *SIFamide* neurons predicted to innervate the un-mapped hemibrain (**D**) *SIFamide* neurons innervating the mapped hemibrain. (**E & F**) Neurons in (D), shown individually. (**G**) Synapse connectivity diagram for neurons shown in (E); left and (F); right. Blue = upstream connections, yellow = downstream connections. Inner concentric circle: segregated by upstream/downstream. Middle circle: neuropil regions with highest connectivity. Inner circle: neurons with highest connectivity. Skeletons and synapse connectivity diagrams were generated using the neuPrint web-tool to analyse the hemibrain:v1.2.1 data

6.3 Investigating the functions of SIFamide signalling in starvation

A role for SIFamide in the modulation of appetitive feeding behaviour was first evidenced in a 2017 study by Martelli et al., based on several observations. First, starvation increases the activity of SIFamide producing neurons. Second, thermogenetic activation of SIFamide neurons mimics the effects of starvation: increasing attraction to food-related odorants and driving feeding behaviours and food consumption. Third, SIFamide neurons receive excitatory input from neurons expressing hugin-PK; a neuropeptide with proposed anorexigenic functions (at least in the larval nervous system (Melcher et al., 2017)), and inhibitory input from neurons expressing MIP, which are also thought to regulate satiety (Min et al., 2016) (Martelli et al., 2017). Finally, SIFamide neurons densely innervate regions of the brain that process chemosensory signals, including those associated with food-related odorants. Collectively, these findings implicate SIFamide signalling in appetitive feeding behaviour.

Two points should be noted, however. First, food intake in starved flies was not reduced by silencing SIFamide neurons (Martelli et al., 2017), suggesting that other pathways may also drive this behaviour. Indeed, short neuropeptide F release in the AL is already known to stimulate food searching in starvation (Root et al., 2011), and a role for AL Dopamine signalling has also been established (Inagaki et al., 2012). Second, all experiments were carried out in female flies. Therefore, while the broadly similar distribution of SIFamide in males supports a function common to both sexes, functional evidence for this is lacking. In the current study, expression of *SIFaR* was identified in *dsx* circuits underlying sexual behaviours. RNAi-mediated loss of the receptor in these neurons had a stimulatory effect on courtship-associated behaviours in both sexes: increasing sexual receptivity in females and producing indiscriminate courtship in males, suggesting SIFamide suppresses activity in the neurons driving these behaviours. An intriguing question is whether SIFamide performs these functions simultaneously, stimulating feeding circuits and inhibiting courtship-relevant circuits in the context of starvation. If so, SIFaR signalling would mirror the function of signalling through Gonadotropin Inhibitory Hormone receptor GnIHR: the vertebrate homolog of SIFaR, which

occurs in mammals in physiological contexts that demand the prioritisation of feeding over reproduction (Smith & Clarke, 2010). Experiments were designed in the current study to test this hypothesis. However, first, two key premises of this hypothesis had to be tested. First, that SIFamide promotes appetitive feeding behaviour in both sexes. Second, that courtship behaviours are affected by starvation.

6.3.1 Does SIFamide promote appetitive feeding behaviour in males?

To address this question, key experiments performed by Martelli et al were repeated in both sexes. First, those showing that SIFamide neurons are activated by starvation and second, those showing that artificially activating these neurons increases food consumption.

To measure SIFamide neuron activation during starvation, the CalexA construct (Musayama et al., 2016) was used. In *SIFa-Gal4>UAS-CaLexA* flies, GFP expression was observed in the PI in both sexes; the region in which SIFamide neuronal somata are located. However, notably, the GFP expressing population comprised more than 4 cells, suggesting that this driver/reporter combination induces some degree of ectopic expression (fig.3A). This was not reported previously. Nonetheless, the effects of starvation on GFP signal strength were analysed. In contrast to previous reports, 24 hours of starvation did not produce a measurable change in the GFP signal intensity in females (fig.3B). GFP signal intensity was similarly unaffected by the same period of starvation in male flies. It is not clear why changes in GFP signal strength were not observed in the current study. Possibly, the sensitivity of the microscope detectors was insufficient: the Lecia SP5 was used in the current study, while Martelli et al. conducted their experiments on a Leica SP8, the sensitivity of which is reported to be significantly improved. In some cases, the activation of peptide producing neurons has been shown to produce a reduction in neuropeptide antibody staining intensity at the soma (for example, see Zhao et al., 2009). As an alternative to the use of CaLexA, the effects of starvation on the distribution of SIFamide immunoreactivity were investigated in the current study. However, SIFa immunoreactivity in SIFamide neuronal somata was not measurably

affected by starvation.

To ascertain whether activation of SIFa neurons promotes feeding in both sexes, short-term calorimetric assays were performed (as per Martelli et al., 2017). Briefly, flies were placed on food containing a red dye and SIFamide neurons were thermogenetically activated using dTrpA1, for a period of 1 hour. The quantity of red dye in homogenised flies was then measured to indicate food consumption. Consistent with the findings of Martelli et al., *SIFa-GAL4>UAS-TrpA1* satiated females consumed more food at the restrictive temperature (fig.3C). Satiated males of this genotype also increased food consumption at the higher temperature but so too did the control genotypes (fig.3D), suggesting an effect of temperature, *per se,* on food intake in satiated males. Therefore, whether i) *SIFamide* neurons are stimulated by starvation in males, and ii) their artificial activation increases food consumption, could not be established in the current study.

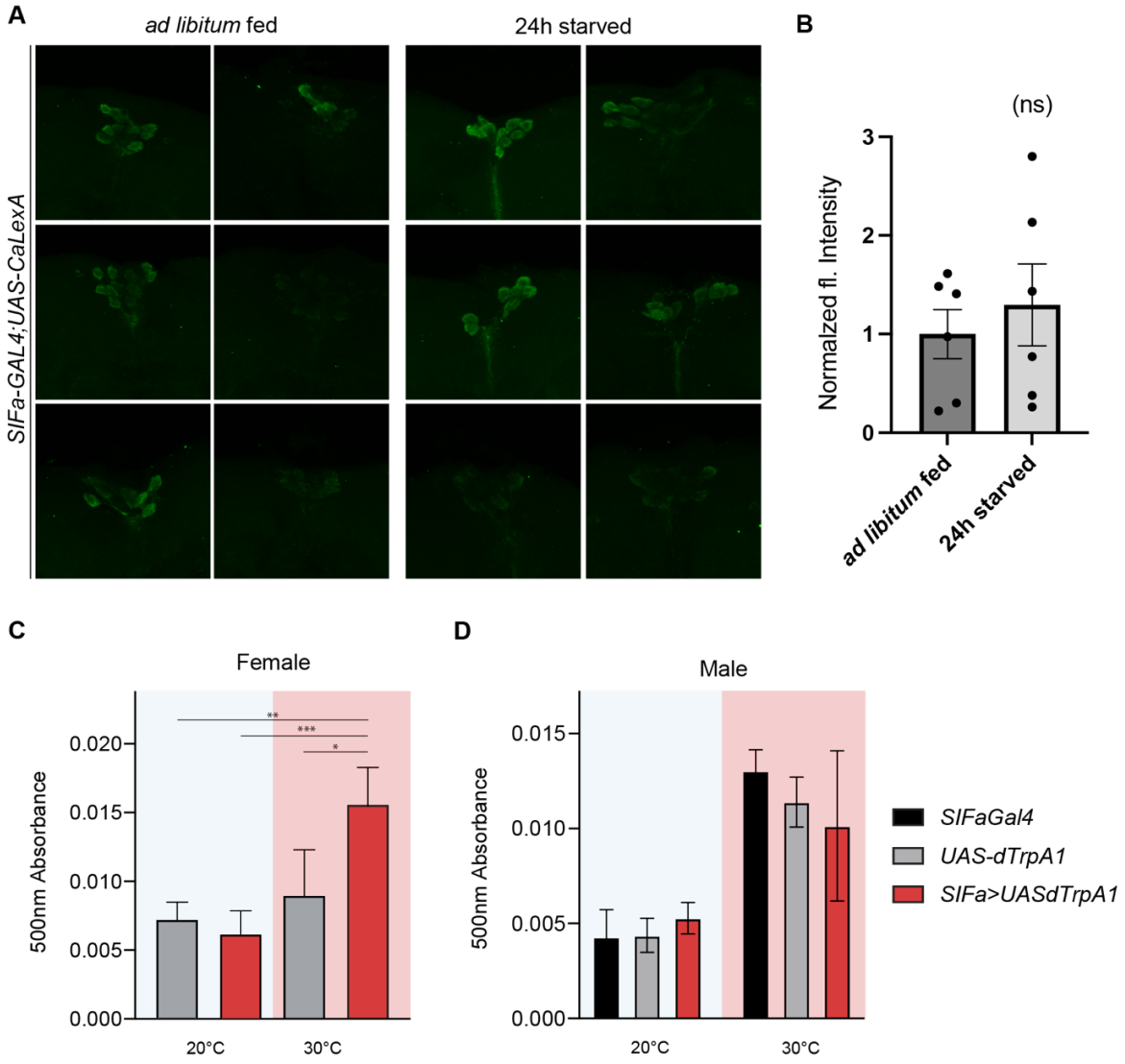


Figure 3. (**A**) Pars intercerebralis region from SIFa-GAL4>UAS-CalexA females, stained with anti-GFP (green). Left panel shows images taken from 6 females fed *ad libitum* prior to dissection. Right panel shows images taken from 6 females staved for 24 hours prior to dissection (**B**) Comparison of normalised GFP fluorescence intensity at the soma for samples shown in (A). (**C**) Effect of TrpA1 activation of SIFamide neurons on food consumption in satiated females, quantified by amount of ingested red dye. (**D**) As in (C), for satiated males. *p = 0.0217, p** = 0.0015, p*** = 0.0006. (Brown-Forsythe and Welch ANOVA test). n > 12 flies/genotype.

6.3.2 Does starvation affect courtship behaviour?

The effect of food deprivation on courtship behaviours in *Drosophila* remains largely unexplored. Interestingly, a relatively recent study showed that male courtship intensity towards receptive females is unaffected by starvation (Churchill et al., 2018), while another found that males from 'starvation-selected' lines do show decreased courtship levels (Billings et al., 2019). To investigate the effects of starvation on male courtship behaviour, a choice assay was devised in which males, starved or sated, were placed in a courtship arena containing a receptive female and a semi-viscous solution containing sucrose (fig.4A). Whereas satiated males showed a strong preference towards courtship, rather than feeding, starved males showed a significant preference for feeding (fig.4B). Over the course of the ~10 -minute assay, in which time, starved males were able to consume the sucrose-containing solution, the preference index shifted from feeding to courtship (fig.4C). Therefore, while starvation does not abolish courtship, it clearly modulates the male's preference for food over a potential mate. The effects of starvation on virgin female receptivity were not analysed, however a similar preference assay could be used for this purpose.

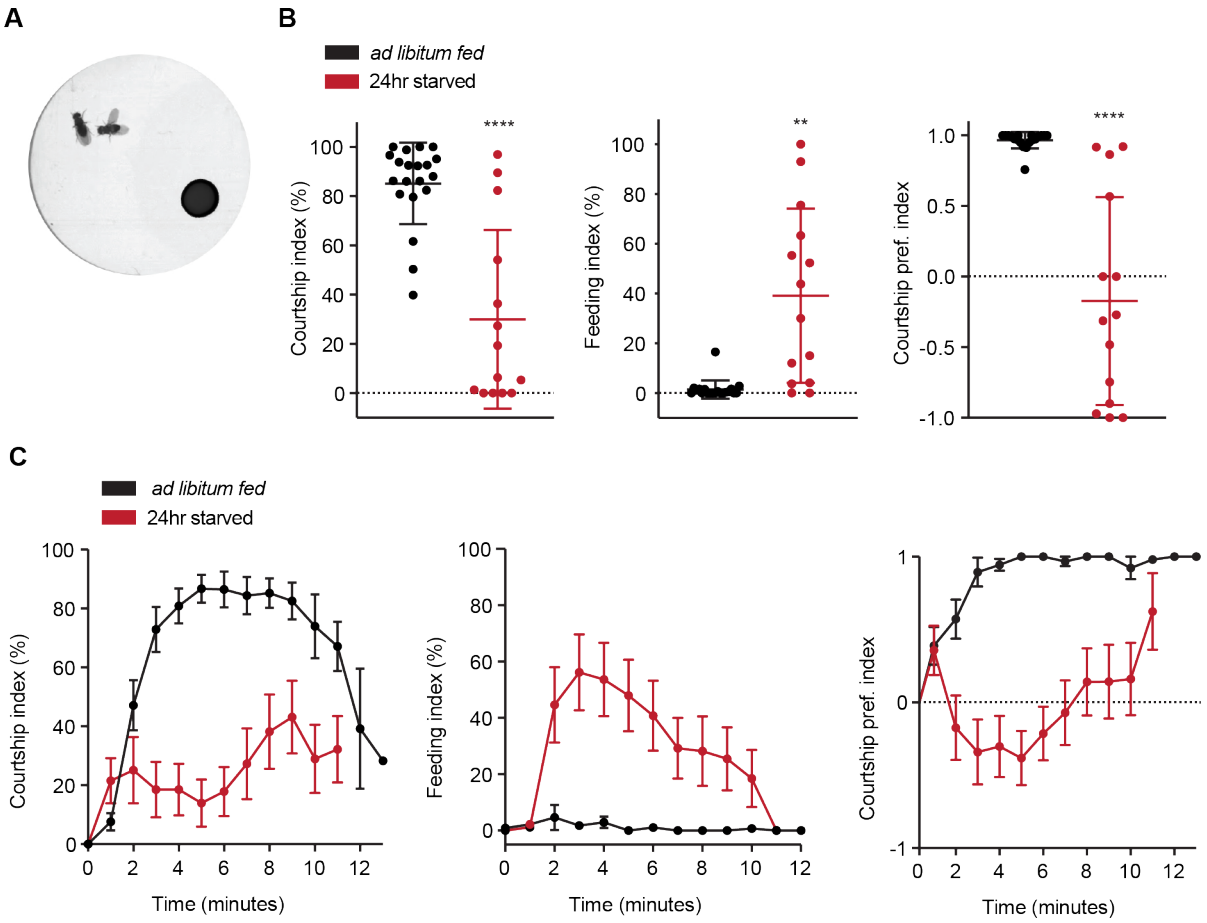


Figure 4: Effects of starvation on male courtship preference. (A) 2 choice assay: Male placed with virgin female and droplet of sucrose containing solution (black). **(B)** Courtship indices (left), feeding indices (middle) and courtship preference indices (right) of ad libitum fed and 24 hour starved wild-type males in 2-choice assay. **p = 0.0014, ***p > 0.0001 (Unpaired t-test with Welch's correction). (C) Courtship indices (left), feeding indices (middle) and courtship preference indices (right) of *ad libitum* fed and 24 hour starved wild-type males in 2-choice assay, plotted over time.

6.3.3 Follow up experiments to investigate modulation of sexual behaviour by SIFamide during starvation

Before further investigating the potential for simultaneous modulation of feeding and sex circuits by SIFamide in the context of starvation, some further data is clearly needed. Firstly, to show that SIFamide neurons are activated by starvation in males. This could not be demonstrated in the current study. However, this could reflect issues with the tools used to measure SIFamide neuron activation. In the case of the CaLexA experiments, an effect was absent in both sexes, despite previous data showing a measurable response in females. Though measurement of anti-peptide immunoreactivity at the soma has been used previously

as a proxy for peptidergic neuron activation, this may not be a reliable indicator of peptide release in all cases. Secondly, follow up behavioural assays are needed to assay the effects of starvation on female receptivity. Once this data is obtained, the next question to address is whether loss of SIFaR exclusively in *dsx* circuits affects the modulation of courtship behaviours in the context of starvation. To do so, flies expressing RNAi against SIFaR in *SIFaR/dsx* neurons can be tested in the feeding/courtship preference assay described above.

6.4 Summary and discussion

Certain aspects of how SIFamide functions in the adult CNS have been uncovered, but the overall picture remains incomplete. Evidence that SIFamide can reach most regions of the CNS, and that the receptor is expressed throughout discrete circuits that govern each of the behaviours it modulates, favours the view that SIFamide functions as a 'global' modulator of multiple behaviours. That SIFamide input sites predominate in the SMP region suggests that SIFamide release is more likely to be triggered by systemic physiological changes rather than, for example, the activity of just a few neurons with specific functions. Consistent with this, starvation has been shown to trigger SIFamide neuron activation, at least in females. However, many questions about SIFamide function remain unanswered. Some of these questions are considered below:

In what context(s) does SIFamide release occur?

The CaLexA approach was employed previously to show *SIFamide* neuron activation in the context of starvation (Martelli et al., 2017). Just as states like starvation are represented by multiple neuropeptides, a single neuropeptide may participate in the representation of multiple states (Bargmann, 2012). Thus, a key question is whether SIFamide release occurs in the adult nervous system in other contexts. This question is particularly intriguing given the evidence that SIFamide release promotes sleep, yet starvation inhibits it (Keene et al., 2010). Might this apparent inconsistency reflect differing levels of SIFamide release in contexts that promote sleep, versus those that necessitate feeding? Notably, sleep induction was observed following brief optogenetic activation of SIFamide neurons (Huang et al., 2021), while appetitive feeding was induced via thermogenetic activation (Martelli et al., 2017). Thus, perhaps these two phenotypes reflect differences in the extent to which the target neurons were activated. The extent of neuronal activation could dictate not only the quantity of released SIFamide, but also whether other molecules, potentially co-expressed by SIFamide neurons, are released simultaneously.

With reference to the current study, it would also be of interest to determine whether other contexts that modulate reproductive behaviours affect SIFamide release. For instance, the male-directed courtship displayed by *SIFaR*-perturbed males could, in theory, reflect blockade of SIFamide signalling during exposure to conspecific males in early life: a process of habituation that forms part of 'courtship learning' (Griffith & Ejima, 2010). Unfortunately, failure to establish an assay to reliably measure SIFamide release precluded further investigation of these possibilities in the current study.

Only a cursory analysis of SIFamide neuron connectome data was performed in the current study. Further scrutiny of where input synapses are located, and the identities of the upstream synaptic partners, could provide clues as to what signals are represented by the activation of SIFamide neurons and release of the peptide. However, the fact that *SIFamide* neurons may, themselves, be subject to the influence of neuromodulators acting extra-synaptically should be born in mind when interpreting this data. Analysis of the transcriptomic profile of SIFamide neurons in scRNAseq data sets could be used to compliment these anatomical approaches, for example by revealing the receptor expression profile of these cells.

Does signalling in discrete circuits occur simultaneously?

As important as identifying what triggers SIFamide release, is determining where and when signalling occurs. One hypothesis is that it occurs in discrete circuits, performing diverse functions, simultaneously. For example, assuming that SIFamide is released during starvation, in this scenario, signalling would occur in feeding, sex and sleep circuits to orchestrate a set of behavioural responses to this physiological state. However, direct evidence of SIFamide signalling in downstream neurons f*ollowing* starvation is lacking. This is true even of signalling in feeding circuits: though evidence suggests anatomical and functional connectivity between *SIFamide* neurons and neurons in the antennal lobe (AL) (Martelli et al., 2017), that signalling in these neurons occurs *following* starvation was not conclusively demonstrated. As such, though artificial activation of SIFamide neurons may trigger appetitive feeding, this is not *necessarily* due to modulation of AL neurons. In the current study, the use of the split-GAL4

driven RNAi restricted *SIFaR* loss-of-function to a specific set of neurons, but did so ubiquitously in time. Conversely, the most recent investigation of SIFamide's role in the regulation of sleep employed optogenetics to activate SIFamide neurons in a precise time window, which transiently increased sleep (Huang et al., 2021), but it is not clear where signalling was potentiated to elicit this effect. Thus, without control over the time and place in which SIFa/SIFaR signalling is manipulated, which circuits SIFamide acts on in specific contexts remains unclear.

Methods that achieve spatial and temporal disruption of SIFamide signalling simultaneously, could. The *SIFaR*-FlpSTOP line generated in the current study could be used for this purpose, as temporally restricted gene disruption can be achieved by temperature control of Flp expression (Fisher et al., 2017). The mechanism by which FlpSTOP achieves gene disruption may not permit SIFaR loss-of-function to be induced in a specific context in the adult (for example during starvation) but would allow disruption or rescue to be confined to developmental or adult stages. This could, for example, reveal whether the effects of SIFamide on behaviour are mediated by signalling through the relevant circuits in development, or in specific contexts in the adult. Another way to determine when SIFamide signalling occurs in specific sub-sets of neurons would be to express an engineered form of the receptor coupled to elements that can report its activation. This approach was used previously by Inagaki et al., who used an engineered form of the receptor for Dopamine; DopR-Tango, to visualise modulation of neurons by this monoamine *in-vivo* (Inagaki et al., 2012)

Do individual SIFamide neurons perform different functions?

The finding that SIFamide neurons traced in the fly hemibrain data set have distinct morphologies should be considered in subsequent investigations of SIFamide function. In particular, this finding should be considered when interpreting the effects of artificial activation or inhibition of SIFamide neurons, as these effects may not reflect the true function of individual SIFamide neurons *in vivo.* Gaining genetic access to individual SIFamide neurons to manipulate their activity *in vivo* would be an ideal solution. However, this would only be

possible with conventional methods if there are expressed genes not common to both cells. A more simple, but nonetheless important, question that needs addressing is whether SIFamide is expressed in neurons in the VNC. Current data on this (presented in the current study and by others (see Martelli et al., 2017) is inconclusive. If SIFamide is expressed in VNC neurons, it would be interesting to determine their function.

Chapter 7: Author’s Summary

A key objective of this study was to shed light upon the distribution of neuropeptide receptors (NPRs), particularly those with ligands implicated in the modulation of reproductive behaviour. For each of these 'candidate' NPRs, 'driver' and 'hemidriver' lines were generated to label expressing cells via the use of fluorescent reporters. The methods used to generate these lines; the Trojan MiMIC system and CRISPR-Cas9, were selected to maximise the fidelity of expression. The data produced was indicative of broad expression of *NPRs* in the CNS*,* with many hundreds, if not thousands, of cells labelled across almost all neuromeres. This finding supports a scenario in which the cognate neuropeptides elicit their effects by signalling through discrete, and functionally distinct, neural circuits. It should be noted, however, that the precise temporal dynamics of *NPR* expression cannot be deduced from the expression of reporters: cells expressing an *NPR* in development, or only transiently, may still be labelled in the adult nervous system, even when transcription of the *NPR* has ceased. It is, therefore, possible that expression of candidate *NPRs* could be widespread in development, but persist in a more restricted, functionally homogenous, neuronal population. However, data from single-cell RNA sequencing of 5-day-adult CNS tissue suggests this is not the case, as transcripts were detected in large numbers of cells across transcriptionally distinct clusters. Thus, the broad expression produced by *NPR* driver lines in the adult CNS likely reflects active expression of the *NPRs* in a significant proportion of these cells.

For *SIFaR* and *SPR*, which were examined in more detail, reporter expression was also evident in cells outside the CNS. With the exception of SP-SPR signalling in sensory neurons in the female reproductive tract (FRT), the functional significance of this expression remains largely unclear. In the case of *SPR*, expression was observed in non-neuronal cells in the FRT. Morphological changes in the FRT after mating have been documented, which require re-structuring of non-neuronal tissues, and are thought to depend on male accessory gland proteins (Acps) (Adams & Wolfner, 2007). It is possible that activation of SPR in these cells by SP: a key Acp, could underly some of these changes. To test this hypothesis, morphological changes in the FRT of mated females, expressing *SPR-RNAi* in these cells, could be analysed.

Though MIP, the ancestral ligand of SPR, is not an Acp, neurons in the female Abg are GFP-labelled by *MIP-GAL4* (Jang et al., 2017), neurites of which are visible in the abdominal nerve trunk. It is not clear whether these neurons innervate the FRT. If so, MIP-SPR signalling in non-neuronal FRT cells could, conceivably, also drive the morphological changes observed after mating. Thus, the relative contribution of these two ligands to any phenotypes observed in *SPR-RNAi* experiments should be investigated, for example by removing SP from the male seminal fluid, or blocking the release of MIP. Unlike *SPR*, peripheral *SIFaR* expression was restricted to neurons with afferent, or efferent, projections to the CNS. As such, though this expression could reflect SIFaR signalling in the periphery (which would necessitate secretion of the peptide into the circulation, given the absence of peripheral SIFamide projections), it is more likely to reflect SIFamide-SIFaR signalling on neurites located within the CNS.

Analysis of scRNAseq data was able to provide insights into the expression level of *NPRs:* information not revealed by the *NPR*-driver lines. This analysis indicated that *NPR*-expressing cells contain very few *NPR* transcripts. This raises several interesting questions. First, given that mRNA and protein abundance are not always directly correlated (Vogel & Marcotte, 2012), how do *NPR* transcript levels relate to the abundance of NPR proteins in different cells? It could be that a similarly low number of receptor proteins are expressed. The capacity for signal amplification by GPCRs makes this a realistic possibility. Alternatively, as proteins often have much longer half-lives than RNA molecules (Vogel & Marcotte, 2012), the number of copies of an NPR could be several times greater than the number of corresponding transcripts. Methods of detecting protein abundance would be necessary to investigate these possibilities further. Another interesting question is whether transcription invariably occurs at these low-levels or, else, is subject to fluctuations. It has been shown that, after mating, expression of *SPR* increases ~10 fold in the antennal lobe (AL) of the female brain (a region of chemosensory processing), driving an increase in SP-SPR signalling that, in turn, modulates odour and food preference (Hussain et al., 2017). The extent to which local regulation of *NPR* expression level is a ‘common’ mechanism for enhancing neuropeptide signalling is not clear, however.

Interestingly, in the current study, *SPR* transcript numbers were far more variable than those of other *NPRs* in the scRNAseq data. It could be that regulation of *SPR* expression level does, indeed, represent a 'special case'. Alternatively, regulation of the other *NPRs* analysed may occur under a more restricted set of conditions, for example at a particular stage of development, or in specific environmental or physiological contexts. There are efforts, currently underway, to generate scRNAseq data from *Drosophilae* subjected to a variety of conditions, and from different developmental stages. It will be interesting to see if transcriptional regulation of other *NPRs* is evident in this data.

By generating split-GAL4 lines, using Trojan-MIMIC/CRISPR-Cas9, *NPR* expression in more restricted sub-sets of cells could be examined. When used in combination with the *dsx-DBD* hemidriver, neurons were intersected and labelled in all these lines, suggesting *dsx* circuits could be subject to modulation by several neuropeptides. In the case of *NPFR-p65AD∩dsx-DBD* and *CrzR-p65AD∩dsx-DBD* there was some variability in expression pattern, however, potentially reflecting a degree of 'leaky' expression of one or more of the transgenes. *SPR-p65AD∩dsx-DBD* expression was consistent, but revealed prevalent expression in glial cells. Both immunohistochemistry, using glial markers, and analysis of scRNAseq data supported this observation. The *SIFaR-p65AD∩dsx-DBD* line produced a consistent pattern of expression, exclusively in a sub-set of *dsx* neurons. As a result, SIFaR was chosen as the subject of functional investigations. A key finding of these investigations was that RNAi-mediated knockdown of SIFaR in *SIFaR-p65AD∩dsx-DBD* neurons induces the same reproductive behaviour phenotypes as those associated with knockdown of the peptide (first described in Terhzaz et al., 2007) and global perturbation of *SIFaR* (evidenced by behavioural analyses of *SIFaR-p65AD* homozygotes in the current study). These were: indiscriminate courtship in males, and reduced time to copulation in females. This finding suggests that SIFamide modulates, directly, the neural circuits that govern these reproductive behaviours. A pertinent, and outstanding, question is where exactly in the circuit this modulation occurs. *SIFaR* expression was evidenced in *dsx* sensory neurons, interneurons and motor neurons,

facilitating modulation at any, or all, of these levels of circuit organisation. Available hemidriver lines expressing in sub-sets of *dsx* neurons could be used in conjunction with *SIFaR-p65AD* to further restrict RNAi expression. Analysis of reproductive behaviours in these genotypes could shed light upon the key loci of SIFaR signalling in the modulation of these behaviours.

A variety of approaches can be used to measure, or visualise, the effects of neuropeptides on target neurons. If the goal is to monitor the *immediate* consequences of NPR signalling, these approaches must target the relevant second messenger(s); in the case of SIFaR: cAMP (based on data from *in-vitro* assays). Though the FRET-based sensor EPAC has been used previously in *Drosophila* to measure cAMP fluctuations, several factors precluded reliable measurements from being made in the current study. In light of this, calcium imaging experiments were designed and conducted to address a slightly different question: how does SIFaR signalling ultimately affect the stimulus-evoked activity of *dsx* target neurons? Though the *ex-vivo* set-up outlined in chapter 5.6.3.2 does not allow responses to 'naturalistic' stimuli to measured, as they can in an *in-vivo* experiment, it provides more control over the delivery of stimuli, affords better access to the neurons from which responses are recorded, and allows a greater range of conditions to be tested (for example, the presence of ectopically applied SIFamide). For these reasons, this set-up represents a viable option for future investigations into the modulatory effects of SIFamide on downstream neurons.

As intriguing as the question of *where* in the *dsx* circuit SIFamide acts, is *when* this modulation occurs. For instance, SIFamide could regulate the development of *dsx* circuitry, establishing its blueprint, or intrinsic physiological properties, but play no part in its modulation thereafter. Alternatively, certain contexts could trigger modulation of *dsx* neurons by SIFamide in the adult nervous system. The novel *SIFaR*-FlpSTOP allele generated in the current study, which facilitates conditional disruption, or rescue, of *SIFaR* expression in a genetically defined cell population, could be a valuable tool in addressing this question. Another approach that may offer insights into the temporal dynamics of SIFamide signalling is measurement, or

visualisation, of the peptide's release. This can be achieved using a tool like CaLexA, which provides *post-hoc* evidence of activity in the peptide-producing neuron(s), and was used previously to show that SIFamide neurons are activated by starvation in females (Martelli et al., 2017). Alternatively, release of a neuropeptide can be reflected by a change in the distribution of relevant antisera. For example, in males, contact with other males; thought to enhance insulin-like peptide 2 (ILP2) release, decreases ILP2 immunoreactivity in the somata of ILP2-producing neurons (Zhang et al., 2022). To determine whether SIFamide modulates *dsx* circuits in certain contexts, approaches such as these could be used to test whether those contexts trigger SIFamide release. Two important caveats should be noted, however. First, such an experiment would not provide information as to where downstream signalling occurs. Thus, additional investigation would be needed to show that modulation of behaviour is mediated by signalling in *dsx* neurons *per se*. Second, potentiation of neuropeptide signalling may not always be driven by an increase in peptide release, a case in point being the potentiation in MIP-SPR signalling in the AL of females after mating, which occurs due to up-regulation of *SPR* expression in downstream neurons. Thus, neither a 'positive' CaLexA signal, or a change in anti-SIFamide distribution, are a prerequisite for the modulation of *dsx* circuits by SIFamide. This second point could explain the lack of evidence for SIFamide release afforded by CaLexa and SIFamide antibody experiments performed in the current study. However, at least in the case of CaLexA experiments, it is more likely that optimisation of the protocol is needed, as the findings of Martelli et al. could not be recapitulated.

Given the pleiotropic nature of most neuropeptides, the ability to observe and manipulate NPR expression in restricted neuronal sub-sets is crucial for elucidating mechanisms of action. That said, when adopting this more 'targeted' approach, it is important that the disparate functions of a neuropeptide are considered. In the case of SIFamide, *SIFaR* expression in *dsx* circuits, and the reproductive behaviour phenotypes observed when this expression is perturbed, provide strong evidence of direct modulation of reproductive behaviour. This should, however, be considered in the context of other findings, which suggest that SIFamide drives both

appetitive feeding (Martelli et al., 2017) and sleep (Park et al., 2014; Huang et al., 2021). In some cases, devising a model in which these functions are reconciled, and testing that model experimentally, may be the best approach in understanding the underlying mechanisms. For example, if SIFamide release is triggered by starvation, and modulates circuits in the AL to drive appetitive feeding (as is evidenced by Martelli et al., 2017), might signalling through *dsx* circuits in the same context modulate reproductive behaviours? In the current study, preliminary experiments were performed to test two key premises of this hypothesis: (i) that reproductive behaviours are modulated by starvation and (ii) that SIFamide drives appetitive feeding in both sexes. Using the 2-choice assay described in chapter 6.3.2, it was possible to test premise (i), and reveal that male courtship preference is modulated by starvation. Regarding premise (ii), starvation did not increase the CaLexA signal in SIFamide neurons in males, but neither did it in females (in contradiction to previous reports). Thus, repeat/optimisation of this experiment is necessary. Interestingly, starved males expressing the heat-activated cation channel, TrpA1, in SIFamide neurons did not consume more food than control genotypes when placed at the restrictive temperature, potentially suggesting that SIFamide promotes appetitive feeding only in females. However, as all male genotypes increased food consumption significantly under these conditions, the effects of temperature on food consumption could have masked those of SIFamide activation. If repeat experiments do support the two premises outlined above, follow-up experiments should ascertain whether modulation of reproductive behaviour in the 2-choice assay is affected by SIFaR-knockdown in *dsx* circuitry.

As illustrated above, where a neuropeptide is thought to modulate multiple behaviours in a specific context, assays highlighting interactions between these behaviours may prove useful. However, there may be some cases in which the modulatory actions of a neuropeptide on discrete behaviours do not occur simultaneously. For instance, it is conceivable that SIFamide potentiates appetitive feeding and suppresses courtship during starvation, but how can this be reconciled with evidence showing that SIFamide signalling promotes sleep? The ability of

neuropeptides, like SIFamide, to modulate behaviours that are mutually exclusive, or to perform different functions under certain circumstances, suggests more 'targeted' signalling must be possible. How this is achieved is an intriguing question for those working in the field. The question is especially pertinent for neuropeptides like SIFamide, which is expressed in just 4 neurons that innervate almost all of the CNS. In the current study, SIFamide immunoreactivity was observed throughout many of the neurites GFP-labelled by *SIFamide-GAL4*. There were, however, regions with a relatively strong signal. Thus, regulation of SIFamide trafficking could be a mechanism for targeted signalling. However, though differential trafficking of neuropeptides in different cell-types has been reported, the mechanisms are unknown (Ramamoorthy et al., 2012), thus it is not clear whether this process could be subject to regulation. Analysis of the *Drosophila* hemibrain connectome data suggested that SIFamide neurons do not have identical morphologies and, consequently, though many regions of innervation are shared, there are clear differences. Thus, different SIFamide functions could be mediated by differences in the extent to which each neuron is activated (according to the identity of its inputs) and, consequently, where release of the peptide occurs. This possibility should be considered in future investigations of SIFamide function, particularly where approaches are used to manipulate SIFamide neurons.

This study presents novel data relating to the distribution and expression of several *NPRs* in *Drosophila*, and provides evidence that one specific neuropeptide, SIFamide, directly modulates the performance of reproductive behaviours. Though, SIFamide, and its receptor, were the subject of functional investigations, the approaches and tools utilised could equally be applied to any of the other ~50 neuropeptides encoded in the *Drosophila* genome, the functions of which remain elusive in many cases. Considering the crucial role played by neuropeptides in all animals, understanding how these signalling molecules function will shed light, more generally, upon how environmental and physiological contexts are represented in the nervous system to affect changes in behaviour.

References

Adams, M.D. et al (2000) 'The genome sequence of *Drosophila* melanogaster'. Science, 287(5461):2185–2195.

Adams, E. M. and Wolfner, M. F. (2007). 'Seminal proteins but not sperm induce morphological changes in the Drosophila melanogaster female reproductive tract during sperm storage'. *Journal of insect physiology*, 53(4), 319–331.

Agrawal, T., Sadaf, S. and Hasan, G. (2013) 'A Genetic RNAi Screen for IP 3 /Ca $^{2+}$ Coupled GPCRs in *Drosophila* Identifies the PdfR as a Regulator of Insect Flight'. *Genetics,* 9(10): e1003849.

Allen, A. M. et al. (2020) 'A single-cell transcriptomic atlas of the adult *Drosophila* ventral nerve cord'. *eLife,* 9, e54074.

Amrein, H. and Thorne, N. (2005) 'Gustatory perception and behavior in *Drosophila melanogaster*'. *Current Biology,* 15(17), R673-684.

Asahina, K. (2018) 'Sex differences in *Drosophila* behavior: qualitative and quantitative dimorphism'. *Current Opinion in Physiology*, 6, 35–45.

Asahina, K. et al. (2014) 'Tachykinin-expressing neurons control male-specific aggressive arousal in drosophila'. *Cell*, 156(1–2), 221–235.

Auer, T. O. and Benton, R. (2016) 'Sexual circuitry in *Drosophila*'. *Current Opinion in Neurobiology*, 38, 18–26.

Bai, L. et al. (2018) 'A Conserved Circadian Function for the Neurofibromatosis 1 Gene'. *Cell Reports*, 22(13), 3416–3426.

Bargmann, C. I. (2012) 'Beyond the connectome: How neuromodulators shape neural circuits'. *BioEssays*, 34(6), 458–465.

Bargmann, C. I. and Marder, E. (2013) 'From the connectome to brain function'. *Nature Methods*, 10(6), 483–490.

Barnstedt, O. et al. (2016) 'Memory-Relevant Mushroom Body Output Synapses Are Cholinergic'. *Neuron*, 89(6), p. 1237.

Bellen, H. J. et al. (2015) 'A library of MiMICs allows tagging of genes and reversible, spatial and temporal knockdown of proteins in *Drosophila*'. *eLife*, 4, 1–28.

Bennet-Clark, H.C. and Ewing, A.W. (1967). 'Stimuli provided by Courtship of Male *Drosophila* melanogaster'. *Nature,* 215, 669-671.

Bennett, M. R. (1999) 'The early history of the synapse: From Plato to Sherrington', *Brain*

Research Bulletin, 50(2), 95–118

Bloch Qazi, M. C., Heifetz, Y. and Wolfner, M. F. (2003) 'The developments between gametogenesis and fertilization: ovulation and female sperm storage in drosophila melanogaster'. *Developmental Biology*, 256(2), 195–211.

Böhme, I. and Beck-Sickinger, A. G. (2009) 'Illuminating the life of GPCRs'. *Cell communication and signaling,* 7, 16.

Bowser, P. R. F. and Tobe, S. S. (2005) 'Immunocytochemical analysis of putative allatostatin receptor (DAR-2) distribution in the CNS of larval *Drosophila* melanogaster'. *Peptide.* 26(1), 81–87.

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

Burbach, J. P. H. (2011) 'What are neuropeptides?' *Methods in Molecular Biology*, 789, 1– 36.

Carlsson, M. A. et al. (2010) 'Multiple neuropeptides in the drosophila antennal lobe suggest complex modulatory circuits'. *Journal of Comparative Neurology*, 518(16), 3359–3380.

Carvalho, G.B., et al. (2006). 'Allocrine modulation of feeding behavior by the Sex Peptide of *Drosophila*'. *Current Biology*, 16(7), 692-696.

Cavanaugh, D. J. et al. (2014) 'Identification of a Circadian Output Circuit for Rest:Activity Rhythms in *Drosophila*'. *Cell*, 157(3), 689–701.

Chalfie, M. et al. (1985). 'The neural circuit for touch sensitivity in Caenorhabditis elegans'. *The Journal of neuroscience,* 5(4), 956–64.

Chapman, T. et al. (2003) 'The sex peptide of *Drosophila* melanogaster: Female post-mating responses analyzed by using RNA interference'*. Proceedings of the National Academy of Sciences of the United States of America*, 100(17), 9923–9928.

Chen, G., Ning, B. and Shi, T. (2019) 'Single-cell RNA-seq technologies and related computational data analysis'. *Frontiers in Genetics*,10, p.317.

Chen, P. S. et al. (1988) 'A male accessory gland peptide that regulates reproductive behavior of female D. melanogaster'. *Cell*, 54(3), 291–298.

Chung, B. Y. et al. (2017) '*Drosophila* Neuropeptide F Signaling Independently Regulates Feeding and Sleep-Wake Behaviour'. *Cell Reports*, 19(12), 2441–2450.

Clowney, E. J. et al. (2015) 'Multimodal Chemosensory Circuits Controlling Male Courtship in *Drosophila*'. *Neuron*, 87(5), 1036–1049.

Connolly, K. and Cook, R.M. (1973). 'Rejection responses by female *Drosophila* melanogaster: Their ontogeny, causality and effects upon the behaviour of the courting male'. *Behaviour,*

44(1/2), 142-166.

Cook, R.M. (1979). ‘The courtship tracking of *Drosophila* melanogaster’. *Biological Cybernetics*, 34, 91- 106.

Court, R. et al. (2020). ‘A Systematic Nomenclature for the *Drosophila* Ventral Nerve Cord’. *Neuron*, 107(6):1071-1079.

Crocker, A. et al. (2010) ‘Identification of a Neural Circuit that Underlies the Effects of Octopamine on Sleep:Wake Behavior’. *Neuron*. 65(5), 670–681.

Crocker, A. and Sehgal, A. (2010). ‘Genetic analysis of sleep’. *Genes & development*, 24(12), 1220-1235.

Croset, V., Treiber, C. D. and Waddell, S. (2018) ‘Cellular diversity in the *Drosophila* midbrain revealed by single-cell transcriptomics’. *eLife*, 7,e34550

Toker, D. (2013) ‘Ancient Greek Philosophy and the Birth of Neuroscience*’* – Daniel Toker. Available at: https://thebrainscientist.com/2013/11/04/greek-philosophy-and- neuroscience.

Demir, E. and Dickson, B.J. (2005) ‘fruitless splicing specifies male courtship behaviour in *Drosophila*’. *Cell,* 12(5), 785–794

Diao, F. et al. (2015) ‘Plug-and-play genetic access to drosophila cell types using exchangeable exon cassettes’. *Cell Reports*, 10(8), 1410–1421.

Dickson, B. J. (2008) ‘Wired for sex: the neurobiology of *Drosophila* mating decisions’. *Science,* 322(5903), 904–909.

Dierick, H. A. and Greenspan, R. J. (2007) ‘Serotonin and neuropeptide F have opposite modulatory effects on fly aggression’. *Nature Genetics*, 39(5), 678–682.

Ejima, A., Nakayama, S. and Aigaki, T. (2001) ‘Phenotypic association of spontaneous ovulation and sexual receptivity in virgin females of *Drosophila* melanogaster mutants’. *Behavior genetics, 31(5),* 437–444..

Ellendersen, B. E. and von Philipsborn, A. C. (2017) ‘Neuronal modulation of D. melanogaster sexual behaviour’. *Current Opinion in Insect Science*, 24, 21–28.

Fernandez, M.P. and Kravitz, E.A. (2013) ‘Aggression and courtship in *Drosophila*: pheromonal communication and sex recognition’. J*ournal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology,* 199(11), 1065–1076.

Ferveur, J. F. (2010) ‘*Drosophila* female courtship and mating behaviors: Sensory signals, genes, neural structures and evolution’. *Current Opinion in Neurobiology*, 20(6), 764–769.

Finger, S. (2005) ‘Santiago Ramón y Cajal: From Nerve Nets to Neuron Doctrine’, in: *Minds Behind the Brain*. Oxford University Press, 197–216.

Fisher, Y. E. et al. (2017) ‘FlpStop, a tool for conditional gene control in *Drosophila*’. *eLife*, 6, p. e22279.

Fredriksson, R. and Schiöth, H. B. (2005) ‘The repertoire of G-protein-coupled receptors in fully sequenced genomes’. *Molecular Pharmacology*, 67(5), 1414–1425.

Gailey, D.A. and Hall, J.C. (1989) ‘Behavior and cytogenetics of fruitless in *Drosophila* melanogaster: different courtship defects caused by separate, closely linked lesions’. *Genetics*, 12(4), 773–785.

Gayon, J. (2016) ‘De Mendel à l’épigénétique: histoire de la génétique’. *Comptes Rendus – Biologies*, 339(7–8), 225–230

Getting, P. A. (1989) ‘Emerging principles governing the operation of neural networks’. *Annual review of neuroscience*, 12, 185–204.

Golic, KG and Lindquist, S. (1989) ‘The FLP recombinase of yeast catalyzes site-specific recombination in the *Drosophila* genome’. *Cell,* 59(3), 499-509.

Golic, K. G. and Golic, M.M. (1996) ‘Engineering the *Drosophila* genome: chromosomal rearrangements by design’. *Genetics*, 144(4), 1693–1711.

Goodwin, S. F. and Hobert, O. (2021) ‘Molecular Mechanisms of Sexually Dimorphic Nervous System Patterning in Flies and Worms’. *Annual Review of Cell and Developmental Biology*, 37, 519–547.

Gratz, S. J., et al. (2015) ‘CRISPR-Cas9 Genome Editing in *Drosophila*’. *Current protocols in molecular biology,* 111, 31.2.1–31.2.20.

Greenspan, R.J. (2012) ‘Seymour Benzer. 15 October 1921 — 30 November 2007’. *Biographical Memoirs of Fellows of the Royal Society*, 58, 23–32.

Greenspan, R.J. and Ferveur, J.F. (2000) ‘Courtship in *Drosophila*’. *Annual review of genetics,* 34, 205–232

Griffith, L. C. and Ejima, A. (2009) ‘Courtship learning in *Drosophila* melanogaster: diverse plasticity of a reproductive behavior’. *Learning & memory*, 16(12), 743–750.

Guiilemin, R., 1978. ‘Biochemical and physiological correlates of hypothalamic peptides: The new endocrinology of the neuron’. *Research publications - Association for Research in Nervous and Mental Disease*, 56, 155–194..

Hall, J.C. (1994) ‘The mating of a fly’. *Science*, 264(5166), 1702–1714. 7.

Hanlon, C. D. and Andrew, D. J. (2015) ‘Outside-in signalling - A brief review of GPCR signalling with a focus on the *Drosophila* GPCR family’. *Journal of Cell Science*, 128(19), 3533–3542.

Hartenstein, V. (2011) 'Morphological diversity and development of glia in *Drosophila*'. *Glia*, 59(9), 1237–1252.

Häsemeyer, M. et al. (2009) 'Sensory Neurons in the *Drosophila* Genital Tract Regulate Female Reproductive Behavior'. *Neuron*, 61(4), 511–518.

He, C. et al. (2013) 'Regulation of circadian locomotor rhythm by neuropeptide Y-like system in *Drosophila* melanogaster'. *Insect Molecular Biology*, 22(4), 376–388.

Hewitt, S. M. et al. (2014) 'Controls for Immunohistochemistry: The Histochemical Society's Standards of Practice for Validation of Immunohistochemical Assays', J*ournal of Histochemistry and Cytochemistry,* 62(10), p. 693.

Hilger, D., Masureel, M. and Kobilka, B. K. (2018) 'Structure and dynamics of GPCR signaling complexes'. *Nature Structural and Molecular Biology*, 25(1), 4–12.

Hodgkin, A. L. and 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.

Huang, H., Possidente, D. R. and Vecsey, C. G. (2021) 'Optogenetic activation of SIFamide (SIFa) neurons induces a complex sleep-promoting effect in the fruit fly *Drosophila* melanogaster'. *Physiology and Behavior*, 239, p. 113507.

Hussain, A. et al. (2016) 'Neuropeptides Modulate Female Chemosensory Processing upon Mating in *Drosophila*'. *PLoS Biology*, 14(5), 1–28.

Inagaki, H.K, et al. (2012) 'Visualizing neuro- modulation in vivo: TANGO-mapping of dop-amine signaling reveals appetite control of sugar sensing'. *Cell*, 148(3), 583–95.

Inagaki, H.K. et al. (2014) 'Optogenetic control of *Drosophila* using a red-shifted channelrhodopsin reveals experience-dependent influences on courtship'. *Nature Methods,* 11(3), 325–332.

Isaac, R.E., Bland, N.D. and Shirras, A.D. (2009) 'Neuropeptidases and the metabolic inactivation of insect neuropeptides'. *General and comparative endocrinology,* 162(1), 8–17.

Isaac, R. E. et al. (2010) '*Drosophila* male sex peptide inhibits siesta sleep and promotes locomotor activity in the post-mated female.' *Proceedings. Biological sciences / The Royal Society*, 277(1678), 65–70.

Jan, L.Y. and Jan, Y.N. (1982) 'Peptidergic transmission in sympathetic ganglia of the frog'. *The Journal of physiology,* 327, 219–246.

Jang, Y. H., Chae, H. S. and Kim, Y. J. (2017) 'Female-specific myoinhibitory peptide neurons regulate mating receptivity in *Drosophila* melanogaster'. *Nat Commun,* 8(1), 1630.

Janssen, I. et al. (1996) ‘Isolation of NEB-LFamide, a novel myotropic neuropeptide from the grey fleshfly’. *Molecular and Cellular Endocrinology*, 117(2), 157–165.

Jékely, G. et al. (2018) ‘The long and the short of it -A perspective on peptidergic regulation of circuits and behaviour’. *Journal of Experimental Biology*, 8, 221(Pt 3)

Johard, H. A. D. et al. (2001) ‘A putative tachykinin receptor in the cockroach brain: Molecular cloning and analysis of expression by means of antisera to portions of the receptor protein’. *Brain Research*, 919(1), 94–105.

Johnson, E. C. et al. (2003) ‘Identification and characterization of a G protein-coupled receptor for the neuropeptide proctolin in *Drosophila* melanogaster’*. Proceedings of the National Academy of Sciences*, 100(10), 6198–6203

Jørgensen, L. M. et al. (2006) ‘Molecular identification of the first SIFamide receptor’. *Biochemical and Biophysical Research Communications*, 340(2), 696–701.

Jumper, J. et al. (2021) ‘Highly accurate protein structure prediction with AlphaFold’. *Nature,* 596(7873):583-589

Kanca, O., Bellen, H. J. and Schnorrer, F. (2017) ‘FLYBOOK Gene Tagging Strategies To Assess Protein Expression, Localization, and Function in *Drosophila*’. *Drosophila. Genetics*, 207(2), 389–412

Kapan, N. et al. (2012) ‘Identified peptidergic neurons in the *Drosophila* brain regulate insulin-producing cells, stress responses and metabolism by coexpressed short neuropeptide F and corazonin’. *Cellular and Molecular Life Sciences*, 69(23), 4051–4066.

Keene, A.C., et al. (2010) ‘Clock and cycle limit starvation- induced sleep loss in *Drosophila*’. *Current Biology*, 20(13), 1209-1215.

Kim, Y.J. et al. (2004) ‘Corazonin receptor signaling in ecdysis initiation’. Proceedings of the National Academy of Sciences of the United States of America, 101(17), 6704–6709.

Kim, W. J., Jan, L. Y. and Jan, Y. N. (2013) ‘A PDF/NPF neuropeptide signalling circuitry of male *Drosophila* melanogaster controls rival-induced prolonged mating’. *Neuron*, 80(5), 1190–1205.

Kim, Y. J. et al. (2010) ‘MIPs are ancestral ligands for the sex peptide receptor’. *Proceedings of the National Academy of Sciences of the United States of America*, 107(14), 6520–6525.

Kimura, K. et al. (2008) ‘Fruitless and Doublesex Coordinate to Generate Male-Specific Neurons that Can Initiate Courtship’. *Neuron*, 59(5), 759–769.

Knapek, S. et al. (2013) ‘Short Neuropeptide F Acts as a Functional Neuromodulator for Olfactory Memory in Kenyon Cells of *Drosophila* Mushroom Bodies’. *The Journal of Neuroscience*. *Society for Neuroscience*, 33(12), p. 5340.

Kohatsu, S. and Yamamoto, D. (2015) 'Visually induced initiation of *Drosophila* innate courtship-like following pursuit is mediated by central excitatory state'. *Nature Communications*, 6, 6457

Kohatsu, S., Koganezawa, M. and Yamamoto, D. (2011) 'Female Contact Activates Male-Specific Interneurons that Trigger Stereotypic Courtship Behavior in *Drosophila*'. *Neuron*, 69(3), 498–508.

Kubli, E. (2003) 'Sex-peptides: seminal peptides of the *Drosophila* male'. *Cellular and molecular life sciences,* 60(8), 1689–1704.

Kubli, E. (2008) 'Sexual behaviour: a receptor for sex control in *Drosophila* females'. *Current Biology*, 18(5), R210-2.

Lee, G., Hall, J.C. and Park, J.H. (2002) 'Doublesex gene expression in the central nervous system of *Drosophila* melanogaster'. *Journal of Neurogenetics,* 16(4), 229–48.

Lee, G., Bahn, J. H. and Park, J. H. (2006) 'Sex- and clock-controlled expression of the neuropeptide F gene in *Drosophila*'. *Proceedings of the National Academy of Sciences*, 103(33), 12580–12585.

Lee, P. T. et al. (2018) 'A gene-specific T2A-GAL4 library for drosophila'. *eLife*, 22,7:e35574

Lelito, K. R. and Shafer, O. T. (2012) 'Reciprocal cholinergic and GABAergic modulation of the small ventrolateral pacemaker neurons of *Drosophila* 's circadian clock neuron network'. *Journal of Neurophysiology*, 107(8), 2096–2108.

Levitis, D. A., Lidicker, W. Z. and Freund, G. (2009) 'Behavioural biologists do not agree on what constitutes behaviour'. *Animal Behaviour*, 78(1), 103–110.

Liu, H. and Kubli, E. (2003) 'Sex-peptide is the molecular basis of the sperm effect in *Drosophila* melanogaster'. *Proceedings of the National Academy of Sciences*, 100(17), 9929–9933.

Liu, H. and Kubli, E. (2003) 'Sex-peptide is the molecular basis of the sperm effect in *Drosophila* melanogaster'. *Proceedings of the National Academy of Sciences of the United States of America,* 100(17), 9929–9933.

Luan, H., et al. (2006) 'Refined spatial manipulation of neuronal function by combinatorial restriction of transgene expression'. *Neuron,* 52(3), 425–436.

Lung, O. and Wolfner, M. F. (1999) '*Drosophila* seminal fluid proteins enter the circulatory system of the mated female fly by crossing the posterior vaginal wall'. *Insect biochemistry and molecular biology,* 29(12), 1043–1052.

Macosko, E. Z. et al. (2015) 'Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets'. *Cell,* 161(5), 1202–1214

Manning, A. (1959) ‘The sexual isolation between *Drosophila* melanogaster and *Drosophila* simulans’. *Animal Behaviour*, 7, 60-65.

Marder, E., O’Leary, T. and Shruti, S. (2014) ‘Neuromodulation of Circuits with Variable Parameters: Single Neurons and Small Circuits Reveal Principles of State-Dependent and Robust Neuromodulation’. *Annual Review of Neuroscience*, 37(1), 329–346.

Martelli, C. et al. (2017) ‘SIFamide Translates Hunger Signals into Appetitive and Feeding Behavior in *Drosophila*’. *Cell Reports*, 20(2), 464–478.

Michel, M. C., Wieland, T. and Tsujimoto, G. (2009) ‘How reliable are G-protein-coupled receptor antibodies?’. *Naunyn-Schmiedeberg’s Archives of Pharmacology, Springer*, 379(4), 385–388.

Middleton, C.A. et al. (2006) ‘Neuromuscular organization and aminergic modulation of contractions in the *Drosophila* ovary’. *BMC Biology*, 4, 17.

Min, S. et al. (2016) ‘Identification of a Peptidergic Pathway Critical to Satiety Responses in *Drosophila*’. *Current Biology*, 26(6), 814–820.

Nagarkar-Jaiswal, S. et al. (2015) ‘A genetic toolkit for tagging intronic MiMIC containing genes’. *eLife,* 23, 4:e08469

Nässel, D. R. and Winther, Å. M. E. (2010) ‘*Drosophila* neuropeptides in regulation of physiology and behavior’. *Progress in Neurobiology*, 92(1), 42–104.

Nässel, D. R. and Zandawala, M. (2019) ‘Recent advances in neuropeptide signaling in *Drosophila*, from genes to physiology and behaviour’. *Progress in Neurobiology*, 179, p. 101607.

Nässel, D. R., Pauls, D. and Huetteroth, W. (2019) ‘Neuropeptides in modulation of *Drosophila* behavior: how to get a grip on their pleiotropic actions’. *Current Opinion in Insect Science*, 36, 1–8.

Nayak, R. and Hasija, Y. (2021) ‘A hitchhiker’s guide to single-cell transcriptomics and data analysis pipelines’. *Genomic*s, 113(2), 606–619.

Nikolaev, V.O. et al. (2004). ‘Novel single chain cAMP sensors for receptor-induced signal propagation’. T*he Journal of biological chemistry,* 279(36), 37215–37218.

Nilsen, S. P. et al. (2004) ‘Gender-selective patterns of aggressive behaviour in *Drosophila* melanogaster’. *Proceedings of the National Academy of Sciences of the United States of America,* 101(33), 12342–12347

Nojima, T. et al. (2021) ‘A sex-specific switch between visual and olfactory inputs underlies adaptive sex differences in behavior’. *Current Biology*, 31(6):1175-1191.e6.

Nusbaum, M. P. et al. (2001) ‘The roles of co-transmission in neural network modulation’. *Trends in Neurosciences*, 24(3):146-54

Ölveczky, B. P. (2014) ‘Ordered randomness in fly love songs’. *Nature* 507(7491), 178–178.

Park, S. et al. (2014) ‘SIFamide and SIFamide Receptor Defines a Novel Neuropeptide Signaling to Promote Sleep in *Drosophila*’. *Molecules and cells*, 37(4), 295–301.

Pavlou, H. J. and Goodwin, S. F. (2013) ‘Courtship behavior in *Drosophila* melanogaster: Towards a “courtship connectome”’. *Current Opinion in Neurobiology*, 23(1), 76-83

Pavlou, H.J. et al. (2016). ‘Neural circuitry coordinating male copulation’. *eLife*, 15, 5:e20713

Ramamoorthy, P., Wang, Q. and Whim, M. D. (2011). ‘Cell type-dependent trafficking of neuropeptide Y-containing dense core granules in CNS neurons’. *The Journal of neuroscience*, 31(41), 14783–14788.

Ran, F. et al. (2013). ‘Genome engineering using the CRISPR-Cas9 system’. *Nature Protocols,* 8(11), 2281–2308

Reinhard, J. (2019) ‘Proboscis Extension Reflex’ in: *Encyclopedia of Animal Behavior* (Second Edition).

Rezával, C. et al. (2012) ‘Neural circuitry underlying *Drosophila* female postmating behavioral responses’. *Current Biology*, 22(13), 1155–1165.

Rezával, C. et al. (2014) ‘Report Sexually Dimorphic Octopaminergic Neurons Modulate Female Postmating Behaviors in *Drosophila*’. *Current Biology*, 24(7), 725–730.

Rideout, E.J., Billeter, J.C. and Goodwin, S.F. (2007) ‘The sex-determination genes fruitless and doublesex specify a neural substrate required for courtship song’. *Current Biology*, 17(17), 1473–78.

Rideout, E.J. et al. (2010) ‘Control of sexual differentiation and behavior by the doublesex gene in *Drosophila* melanogaster’. *Nature Neuroscience*, 13(4), 458–66.

Rings, A. and Goodwin, S. F. (2019) ‘To court or not to court – a multimodal sensory decision in *Drosophila* males’. *Current Opinion in Insect Science*, 35, 48–53.

Ritchie, M.G., Halsey, E.J. and Gleason, J.M. (1999) ‘*Drosophila* song as a species-specific mating signal and the behavioural importance of Kyriacou & Hall cycles in D. melanogaster song’. *Animal Behaviour*, 58, 649-657.

Rodríguez-Valentín, R. et al. (2006) ‘Oviduct contraction in *Drosophila* is modulated by a neural network that is both, octopaminergic and glutamatergic’. *Journal of cellular physiology,* 209(1), 183–198.

Root, C.M. et al. (2011) ‘Presynaptic facilitation by neuropeptide signalling mediates odor-

driven food search'. *Cell*, 145(1), 133–44.

Rubinstein, C.D. and Wolfner, M.F. (2013) '*Drosophila* seminal protein ovulin mediates ovulation through female octopamine neuronal signaling'. *Proceedings of the National Academy of Sciences of the United States of America,* 110(43), 17420–17425.

Rueter, S.M., Dawson, T.R. and Emeson, R.B. (1999). 'Regulation of alternative splicing by RNA editing'. *Nature*, 399(6731), 75–80.

Ryder, E. and Russell, S. (2003) 'Transposable elements as tools for genomics and genetics in *Drosophila*'. *Briefings in functional genomics & proteomics,* 2(1), 57–71.

Ryder, E. et al. (2004) 'The DrosDel collection: A set of P-element insertions for generating custom chromosomal aberrations in *Drosophila* melanogaster'. *Genetics*, 167(2), 797–813.

Sadaqat, Z., Kaushik, S. and Kain, P. (2021) 'Gut Feeding the Brain: *Drosophila* Gut an Animal Model for Medicine to Understand Mechanisms Mediating Food Preferences'. *Animal Models in Medicine and Biology.* Published online April 6th 2021.

Saper, C. B. and Sawchenko, P. E. (2003) 'Magic peptides, magic antibodies: Guidelines for appropriate controls for immunohistochemistry'. *Journal of Comparative Neurology*, 465(2), 161–163.

Schally, A. V., Arimura, A. and Kastin, A. J., (1973) 'Hypothalamic regulatory hormones'. *Science,* 179(4071), 341-350.

Scheffer, L. K. et al. (2020) 'A connectome and analysis of the adult drosophila central brain'. *eLife*, 9, e57443.

Sellami, A. and Veenstra, J. A. (2015) 'SIFamide acts on fruitless neurons to modulate sexual behavior in *Drosophila* melanogaster'. *Peptides*, 74, 50–56.

Sewell, W (2006). Genetic Basis for Arthropod Limb Diversification. (Available online at: https://www.researchgate.net/publication/277744377_Genetic_Basis_for_Arthropod_Limb_Diversification)

Shafer, O. T. et al. (2008) 'Widespread Receptivity to Neuropeptide PDF throughout the Neuronal Circadian Clock Network of *Drosophila* Revealed by Real-Time Cyclic AMP Imaging'. *Neuron*, 58(2), 223–237.

Shankar, S. et al. (2015) 'The neuropeptide tachykinin is essential for pheromone detection in a gustatory neural circuit'. *eLife*, 4, e06914.

Shao, L. et al. (2017) 'Dissection of the *Drosophila* neuropeptide F circuit using a high-throughput two-choice assay'. *Proceedings of the National Academy of Sciences*, 114(38), E8091–E8099.

Shen, P. and Cai, H. N. (2001) '*Drosophila* neuropeptide F mediates integration of chemosensory stimulation and conditioning of the nervous system by food'. *Journal of Neurobiology*, 47(1), 16–25.

Shimosako, N., Hadjieconomou, D. and Salecker, I. (2014) 'Flybow to Dissect Circuit Assembly in the *Drosophila* Brain'. *Methods in molecular biology,* 1082, 57–69.

Shohat-Ophir, G. et al. (2012) 'Sexual experience affects ethanol intake in *Drosophila* through Neuropeptide F'. *Science*, 335(6074), 1351–1355.

Siwicki, K.K. and Kravitz, E.A., (2009) 'Fruitless, doublesex and the genetics of social behavior in *Drosophila* melanogaster'. *Current opinion in neurobiology*, 19(2), 200–206.

Sousa-Nunes, R., Yee, L. L. and Gould, A. P. (2011) 'Fat cells reactivate quiescent neuroblasts via TOR and glial insulin relays in *Drosophila*'. *Nature*, 471(7339), 508–513.

Spieth, H.T. (1974) 'Courtship behavior in *Drosophila*'. Annual review of entomology, 19, 385–405.

Spieth, H.T. and Ringo, J.M. (1983). 'Mating behavior and sexual isolation in *Drosophila*'. In: *The Genetics and Biology of Drosophila, Volume 3c*, M. Ashburner, H. Carson and T. J., eds. (New York: Academic Press), 223–284.

Starling, E.H. (1904) 'The chemical regulation of the secretory process' (Croonian Lecture to the Royal Society). *Proceedings of the Royal Society*, 73B, 310–322.

Stockinger, P. et al. (2005) 'Neural circuitry that governs *Drosophila* male courtship behavior'. *Cell*, 121(5), 795-807.

Strausfeld, N.J. and Meinertzhagen, I.A. (1998) 'The insect neuron: types, morphologies, fine structure, and relationship to the architectonics of the insect nervous system'. *Microscopic anatomy of invertebrates, Volume IIB*: Insecta, F.W. Harrison and M. Locke, eds. (New York: Wiley), 487-538.

Tan, K. J. et al. (2015) 'The neuropeptide tachykinin is essential for pheromone detection in a gustatory neural circuit'. *eLife*, 4, e06914.

Tang, W. et al. (2009) 'Faithful expression of multiple proteins via 2A-peptide self-processing: A versatile and reliable method for manipulating brain circuits'. *Journal of Neuroscience*, 29(27), 8621–8629.

Tawfik, A.I. et al. (1999) 'Identification of the gregarization-associated dark-pigmentotropin in locusts through an albino mutant'. *Proceedings of the National Academy of Sciences of the United States of America*, 96(12), 7083–7087.

Tayler, T. D. et al. (2012) 'A neuropeptide circuit that coordinates sperm transfer and copulation duration in *Drosophila*'. *Proceedings of the National Academy of Sciences of the United States*

of America, 109(50), 20697–20702.
Terhzaz, S. et al. (2007) 'The neuropeptide SIFamide modulates sexual behavior in *Drosophila*'. *Biochemical and Biophysical Research Communications*, 352(2), 305–310.

Thibault, S. T. et al. (2004) 'A complementary transposon tool kit for *Drosophila* melanogaster using P and piggyBac'. *Nature Genetics*, 36(3), 283–287.

Thistle, R., et al (2012) 'Contact chemoreceptors mediate male-male repulsion and male-female attraction during *Drosophila* courtship'*. Cell,* 149(5), 1140–1151.

Thorne, N., Bray, S. and Amrein, H. (2005) 'Function and expression of the *Drosophila* gr genes in the perception of sweet, bitter and pheromone compounds'*. Chemical Senses* 30 Suppl1, i270-272.

Valenstein, E. S. (2005) 'The War of the Soups and the Sparks'. Columbia University Press.

Van den Pol, A. N. (2012) 'Neuropeptide Transmission in Brain Circuits'. *Neuron*, 76(1), 98–115.

Vandesompele, J. et al. (2002) 'Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes' *Genome Biol*, 18,3(7).

Vecsey, C. G., Pirez, N. and Griffith, L. C. (2014) 'The *Drosophila* neuropeptides PDF and sNPF have opposing electrophysiological and molecular effects on central neurons'. *Journal of Neurophysiology*, 111(5), 1033–1045.

Veenstra, J.A. (1989). 'Isolation and structure of corazonin, a cardioactive peptide from the American cockroach'. *FEBS Letters,* 250(2), 231–34.

Venken, K. J. T. and Bellen, H. J. (2007) 'Transgenesis upgrades for *Drosophila* melanogaster'. *Development*, 3571–3584.

Venken, K. J. T., Simpson, J. H. and Bellen, H. J. (2011a) 'Genetic Manipulation of Genes and Cells in the Nervous System of the Fruit Fly'. *Neuron*, 72(2), 202–230.

Venken, K. J. T. et al*.* (2011b) 'MiMIC: A highly versatile transposon insertion resource for engineering *Drosophila* melanogaster genes'. *Nature Methods*, 8(9), 737–747.

Verleyen, P. et al*.* (2004) 'SIFamide is a highly conserved neuropeptide: a comparative study in different insect species'. *Biochemical and Biophysical Research Communications*, 320(2), 334–341.

Villella, A. et al. (1997) 'Extended reproductive roles of the fruitless gene in *Drosophila* melanogaster revealed by behavioral analysis of new fru mutants'. *Genetics* 147(3), 1107–1130.

Villella, A. and Hall, J.C. (2008) 'Neurogenetics of courtship and mating in *Drosophila*'. *Advances in Genetics*, 62, 67-184.

Vogel, C. and Marcotte, E. M. (2012). 'Insights into the regulation of protein abundance from proteomic and transcriptomic analyses'. *Nature reviews. Genetics*, 13(4), 227–232.

Von Schilcher, F. (1976) 'The function of pulse song and sine song in the courtship of *Drosophila* melanogaster'. *Animal Behaviour,* 24, 622-625.

Wen, T. et al. (2005) '*Drosophila* neuropeptide F and its receptor, NPFR1, define a signaling pathway that acutely modulates alcohol sensitivity'. *Proceedings of the National Academy of Sciences*, 102(6), 2141–2146.

White, J.G. et al. (1986) 'The structure of the nervous system of the nematode Caenorhabditis elegans'. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences,* 314(1165), 1–340..

Winther, A.M. et al. (2003) 'Neuronal expression of tachykinin-related peptides and gene transcript during postembryonic development of *Drosophila*'. T*he Journal of comparative neurology,* 464(2), 180–196.

Wise, S. et al. (2002) 'Neuroanatomical studies of period gene expression in the hawkmoth, Manduca sexta'. The Journal of comparative neurology, 447(4), 366–380.

Wong, K. et al. (2019) 'Neuropeptide relay between SIFa signaling controls the experience-dependent mating duration of male *Drosophila*', *bioRxiv* 819045.

Wu, M. N. et al. (2009) 'The Effects of Caffeine on Sleep in *Drosophila* Require PKA Activity, But Not the Adenosine Receptor', *Journal of Neuroscience*, 29(35), 11029–11037.

Yang, C.H. et al. (2009) 'Control of the postmating behavioral switch in *Drosophila* females by internal sensory neurons'. *Neuron,* 61(4), 519–526.

Yapici, N. et al. (2008) 'A receptor that mediates the post-mating switch in *Drosophila* reproductive behaviour.' *Nature*, 451(7174), 33–37.

Yew, J.Y. et al. (2009) 'Analysis of neuropeptide expression and localization in adult *Drosophila* melanogaster central nervous system by affinity cell-capture mass spectrometry'. J*ournal of proteome research,* 8(3), 1271–1284.

Zhang, S. X., Rogulja, D. and Crickmore, M. A. (2016) 'Dopaminergic Circuitry Underlying Mating Drive'. *Neuron*, 91(1), 168–181.

Zhang, Y. Y., Price, M. L. and Zhao, J. L. (2013) 'Regulation of Sleep by Neuropeptide Y- Like System in *Drosophila* melanogaster'. *PLoS ONE*, 8(9), p. 74237.

Zhang, L., Guo, X. and Zhang, W. (2022). 'Nutrients and pheromones promote insulin release to inhibit courtship drive'. *Science advances*, 8(10), eabl6121.

Zhao, Y. et al. (2010) 'Corazonin neurons function in sexually dimorphic circuitry that shape behavioral responses to stress in *Drosophila*'. *PLoS ONE*, 5(2), p. 9141.

Zupanc, G. K. H. (1996) ‘Peptidergic transmission: From morphological correlates to functional implications’. *Micron*. 27(1), 35–91.

Accompanying Materials

RESEARCH ARTICLE

A single-cell transcriptomic atlas of the adult *Drosophila* ventral nerve cord

Aaron M Allen†*, Megan C Neville†, Sebastian Birtles, Vincent Croset, Christoph Daniel Treiber, Scott Waddell, Stephen F Goodwin*

Centre for Neural Circuits and Behaviour, University of Oxford, Oxford, United Kingdom

Abstract The *Drosophila* ventral nerve cord (VNC) receives and processes descending signals from the brain to produce a variety of coordinated locomotor outputs. It also integrates sensory information from the periphery and sends ascending signals to the brain. We used single-cell transcriptomics to generate an unbiased classification of cellular diversity in the VNC of five-day old adult flies. We produced an atlas of 26,000 high-quality cells, representing more than 100 transcriptionally distinct cell types. The predominant gene signatures defining neuronal cell types reflect shared developmental histories based on the neuroblast from which cells were derived, as well as their birth order. The relative position of cells along the anterior-posterior axis could also be assigned using adult Hox gene expression. This single-cell transcriptional atlas of the adult fly VNC will be a valuable resource for future studies of neurodevelopment and behavior.

***For correspondence:**
aaron.allen@dpag.ox.ac.uk (AMA);
stephen.goodwin@dpag.ox.ac.uk (SFG)

†These authors contributed equally to this work

Competing interests: The authors declare that no competing interests exist.

Funding: See page 25

Received: 30 November 2019
Accepted: 03 April 2020
Published: 21 April 2020

Reviewing editor: Richard S Mann, Columbia University, United States

© Copyright Allen et al. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Introduction

The adult *Drosophila* central nervous system (CNS) consists of the brain in the head capsule and the ventral nerve cord (VNC; also known as ventral nervous system) in the thorax (***Court et al., 2017***; ***Ito et al., 2014***). The VNC receives and integrates sensory input from the periphery and sends this information to the brain in ascending neurons through the cervical connective (***Tsubouchi et al., 2017***). The brain, in turn, sends sensory-motor signals to the VNC via descending neurons (***Namiki et al., 2018***). The VNC transforms these signals into locomotor actions (***Harris et al., 2015***). It controls muscles in the thorax in unique ways, depending on whether it is steering the wings during flight or generating acoustic communication signals during both reproductive and agonistic behaviors (***Clyne and Miesenböck, 2008***; ***Jonsson et al., 2011***; ***Shirangi et al., 2013***; ***von Philipsborn et al., 2011***). It coordinates muscles in the legs to walk, jump, groom, reach, touch, and taste (***Bidaye et al., 2014***; ***Card and Dickinson, 2008***; ***Chen et al., 2018***; ***Gowda et al., 2018***; ***Harris et al., 2015***; ***Howard et al., 2019***; ***Kim et al., 2017***; ***Mamiya et al., 2018***; ***Mendes et al., 2013***; ***Mendes et al., 2014***; ***Seeds et al., 2014***; ***Tuthill and Wilson, 2016***; ***Wosnitza et al., 2013***). The VNC also controls musculature in the abdomen relevant to copulation and reproduction including abdominal bending, attachment, and ejaculation in the male (***Crickmore and Vosshall, 2013***; ***Jois et al., 2018***; ***Pan et al., 2011***; ***Pavlou et al., 2016***; ***Tayler et al., 2012***), and sperm storage and oviposition in the female (***Kimura et al., 2015***; ***Lee et al., 2015***).

These different functions are orchestrated by anatomically discrete segments of the VNC. Whereas the thoracic neuromeres control the legs and wings, the abdominal neuromere controls the abdominal muscles, gastric system, and reproductive organs. Much of the developmental mechanisms that generate and assemble the VNC have been characterized (***Venkatasubramanian and Mann, 2019***). Like other holometabolous insects, *Drosophila* undergo two stages of neurogenesis building two distinct nervous systems. Embryonic neurogenesis gives rise to the larval nervous system, and post-embryonic neurogenesis produces adult specific neurons. The structure of the adult nervous system is established in the embryo where pioneering neurons setup networks of tracts,

which later developing neurons follow (**Hartenstein, 2018**). The adult VNC is comprised of 4 primary neuromeres, three thoracic (one per each pair of legs: the prothoracic neuromere, ProNm; mesothoracic neuromere, MesoNm and metathoracic neuromere, MetaNm) and one fused abdominal neuromere (ANm) (**Court et al., 2017**; **Niven et al., 2008**). The neuromeres, composed of one or more fused neuropils, are segment-specific parts of the CNS which process sensory signals and control movements of their specific segments. The thoracic neuromeres are homologous structures and are thus morphologically similar to each other (**Smarandache-Wellmann, 2016**). The neurons in these segments are derived from a set of repeating type I neuroblasts. Cells derived from a given neuroblast produce a lineage, which can be split into $Notch^{+}$ and $Notch^{-}$ hemilineages (**Truman et al., 2010**). The formation of tagmatic boundaries, which group these segments into morphological units along the body axis, is established through differential expression of Hox family transcription factors (TFs) (**Angelini and Kaufman, 2005**). Cell types within each neuromere are genetically encoded by developmental programs. However, it is unclear whether the mature terminal identity of an individual neuron can be determined from its adult transcriptome.

Although a large body of work has investigated how the adult *Drosophila* CNS is established from that of the embryo, many outstanding questions remain. It is for example unknown whether the TFs involved in establishing the cellular diversity of the nervous system also play a role in the form and function of these same neurons in the adult. Although certain features of individual neurons can be plastic, the identity of a terminally differentiated neuron is likely to remain stable throughout the animal's life. It is therefore important to understand gene expression and regulation that permit mature neurons to preserve their subtype identity, morphology and connectivity, and maintain neural circuit function throughout adulthood. The VNC of the genetically tractable vinegar fly, *Drosophila melanogaster*, is ideally suited for these studies.

Here we used single-cell RNA-sequencing to characterize the transcriptomes of individual cells in a 5 day old adult *Drosophila* VNC. We generated an atlas of the VNC with 26,768 single-cell transcriptional profiles that define more than 100 cell types. This analysis reveals that the VNC has a roughly equal number of inhibitory GABAergic neurons and excitatory cholinergic neurons, and that genes encoding preproneuropeptides are amongst the most highly expressed. The segmentally repeating nature and developmental history of the VNC is born out in the cellular transcriptomes, as maintained expression of Hox and several other neuronal lineage marker genes persist. In addition, these profiles provide many novel markers for known and new cell types. This single-cell atlas of the adult *Drosophila* VNC provides a useful resource to help connect cellular identity to behaviorally relevant neural circuit function.

Results

Single-cell transcriptomic atlas of the adult VNC

We created a single-cell transcriptome atlas of the VNC using single-cell RNA-sequencing with 10x Chromium chemistry. We processed 80 VNCs, 20 per independent replicate (**Figure 1A**; see Materials and methods). The data were filtered according to the number of genes (nGene), the number of unique molecule identifiers (nUMI), and the proportion of mitochondrial genes (prop.mito) (**Figure 1—figure supplement 1A,B**; see Materials and methods). A similar number of expressed genes, and UMIs were retrieved in each replicate (**Figure 1—figure supplement 1C**). Pooling the data, we recovered 26,768 high quality cells with a median of 1170 genes and 2497 transcripts per cell. We compared the expression levels observed in this single-cell data set with those reported from bulk sequencing, to evaluate the extent to which the tissue dissociation and the 10x procedure effect gene expression. The filtered single-cell data showed high correlation ($r = 0.76$) to FlyAtlas' bulk VNC sequencing data (**Figure 1—figure supplement 1D**; **Leader et al., 2018**). Comparisons of the pseudo-bulk expression levels between replicates of the filtered cells yielded correlations coefficients from 0.93 to 0.95 (**Figure 1—figure supplement 2**).

We performed a Canonical-Correlation Analysis (CCA) on the transcriptomes and reduced the first 45 dimensions into two t-SNE dimensions (**van der Maaten, 2014**; **Figure 1B**). We evaluated the clustering resolution with a clustering tree, comparing the relationship between clusters at different resolutions (**Figure 1—figure supplement 3**). A cluster resolution of 12 was chosen as it best resolved established substructures within our data (as described below). This resulted in 120 distinct

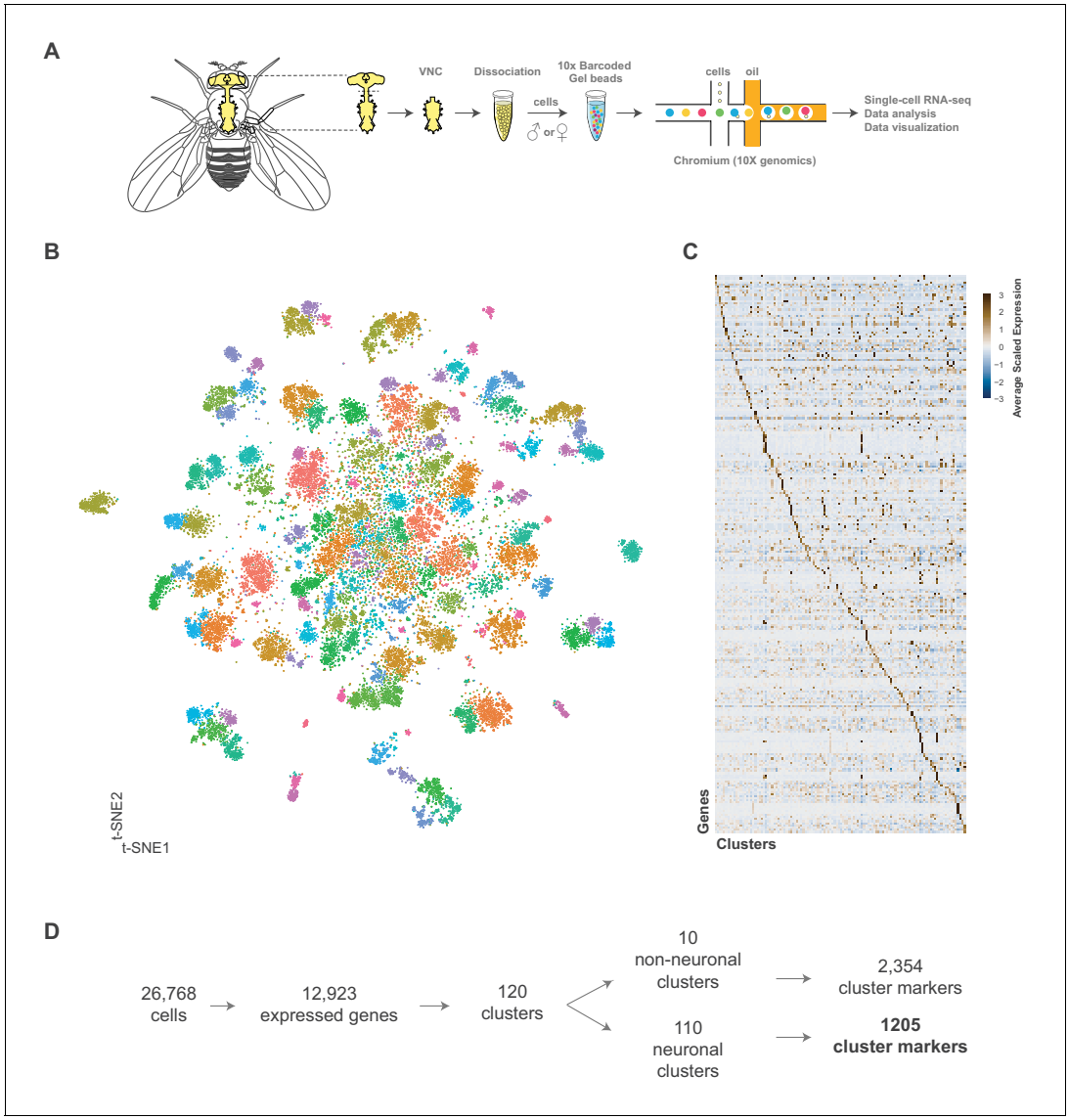


Figure 1. Single cell sequencing of 5 day old adult *Drosophila* VNC. (**A**) Schematic workflow showing single cell sequencing data generation. Male and female *Drosophila* VNCs were dissected and dissociated prior to droplet encapsulation of individual cells with barcoded beads, forming gel beads in emulsion (GEMs). Following barcode incorporation, molecular amplification and sequencing the transcriptional profiles of individual cells were determined. (**B**) Two-dimensional representation (t-SNE) of 26,768 *Drosophila* VNC cells grouped into 120 clusters. Clusters were assigned using the Louvain algorithm, using 45 CCA dimensions with a cluster resolution of 12. Each dot is a cell colored by cluster identity. (**C**) Heatmap showing scaled, log-normalized expression of top 5 cluster discriminative genes per cluster. (**D**) Flow diagram representing neuronal and non-neuronal cluster identification, including total number of genes identified as cluster markers.

The online version of this article includes the following source data and figure supplement(s) for figure 1:

Source data 1. List of marker genes for the 120 clusters shown in ***Figure 1B*** and ***Figure 1—figure supplement 4***.
Source data 2. Cluster summary table.
Figure supplement 1. Cut-offs and correlation to bulk sequencing.
Figure supplement 2. Pseudo-bulk comparisons between replicates of filtered VNC data.
Figure supplement 3. Cluster resolution analysis of adult VNC dataset.
Figure supplement 4. t-SNE plot of 5 day old adult VNC with cluster numbers labeled.
Figure supplement 5. Replicate contributions.
Figure supplement 6. t-SNE spatial distributions.
Figure supplement 7. Identification of neuronal and glia clusters.

clusters (***Figure 1—figure supplement 4***). Cells from independent experimental replicates were proportionally distributed across the clusters with an alignment metric of 94.6% (***Figure 1—figure supplement 5A,B***). All replicates contributed to all clusters (with one exception) and showed high correlation of cluster level expression (***Figure 1—figure supplement 1C***). Replicate one did not have any constituents in cluster 118, which had only 22 cells total. The number of genes and transcripts showed a uniform distribution across the t-SNE (***Figure 1—figure supplement 6A***).

These 26,768 cells represent an approximate 1.5x coverage of the adult VNC. Lineage based counts of the adult specific post-embryonic neurons in the VNC range from 10,000 to 11,000 cells (***Birkholz et al., 2015***; ***Lacin et al., 2019***), with a total number of cells in the VNC forecasted to be ~16,000 (***Lacin et al., 2019***). Connectomic analyses of the adult VNC using electron microscopy estimated a larger total number of ~20,000 cells (***Bates et al., 2019***). Based on these prior numbers, our cell atlas has 1.3–1.7x coverage of the VNC.

Cluster-defining marker genes

The 120 t-SNE clusters are defined by unique combinations of significantly enriched genes that we refer to as cluster markers (***Figure 1C***, ***Figure 1—source data 1***). The number of significantly enriched genes and the maximum observed log fold-change varies by cluster (***Figure 1—figure supplement 6B***; ***Figure 1—source data 2***). These cluster markers are likely to be important for the development and/or maintenance of the cell types represented by the clusters. We compared the significantly enriched cluster markers from this data set (excluding the salivary and sperm clusters, see below) to the cluster markers of re-analyzed mid-brain Drop-seq (***Croset et al., 2018***) and whole brain 10x (***Davie et al., 2018***) data sets (***Figure 2—figure supplement 1***; ***Figure 2—figure supplement 1—source data 1***; ***Figure 2—figure supplement 1—source data 2***). 85% of the mid-brain cluster markers were also found in our VNC cluster markers, while 67% of the whole brain markers overlapped. This suggests that the majority of the genes that define neuronal and glial identity in the brain also define neuronal and glial identity in the VNC.

We defined 110 neuronal clusters using expression of known neuron-specific and neuron-enriched genes (***Figure 1—figure supplement 7***; see methods), such as *embryonic lethal abnormal vision* (*elav*), *neuronal Synaptobrevin* (*nSyb*) and the long non-coding (lnc) RNA *noe* (*noe*) (***Davie et al., 2018***; ***DiAntonio et al., 1993***; ***Kim et al., 1998***; ***Robinow et al., 1988***). These neuronal clusters can be distinguished from each other by differential expression of 1205 additional cluster markers (***Figure 1D***, ***Figure 1—source data 1***). Functional enrichment analysis of these cluster-defining marker genes identified those encoding immunoglobulin (Ig)-like domains as most significantly enriched (***Figure 2A***), including many cell adhesion molecules that are known to establish neuronal connectivity during development (***Özkan et al., 2013***). Specifically, many members of the Immunoglobulin superfamily (IgSF) continue to be expressed in the adult and contribute to neuronal cluster identity (***Figure 2B***; ***Figure 2—figure supplement 2A***). The defective proboscis extension response (Dpr) subfamily, and their binding partners, the Dpr-interacting protein (DIP) subfamily, are markers for over 50% of neuronal clusters identified in the adult VNC (***Figure 2—figure supplement 2B***). Dpr and DIP genes provide a complex interaction network regulating neural circuit assembly and have been proposed to act as neuronal 'identification tags' during development. Distinct combinations of Dpr and DIP proteins are expressed by different neuronal classes in the developing optic lobe, the antennal lobe, and the VNC (***Carrillo et al., 2015***; ***Özkan et al., 2013***). Specific combinations of Beat and Side IgSF proteins play an important role in the sculpting the nervous system (***Özkan et al., 2013***; ***Pipes et al., 2001***). Almost all Beat and Side family members were found as cluster markers in the VNC (***Figure 2B***), and their enrichment was largely mutually exclusive, with 80% of the identified clusters enriched for either Beats or Sides (***Figure 2—figure supplement 2C***). The enrichment of IgSFs in the adult VNC supports their importance in establishing cell type specificity and suggests an ongoing role in maintenance of neuronal connectivity in the mature nervous system (***Figure 2B***).

G-protein coupled receptors (GPCRs) were also significantly over-represented as cluster markers (***Figure 2A***, ***Figure 2—figure supplement 3A***). GPCRs play a critical role in intercellular communication by interacting with a diverse group of extracellular ligands such as neurotransmitters, modulatory neuropeptides and biogenic amines, peptide hormones, and gases (***Hanlon and Andrew, 2015***). We observed a median of 8 GPCRs expressed per cell and a median of 22 transmembrane receptors in total per cell. Due to the potential for drop-out events (a gene is expressed in a cell,

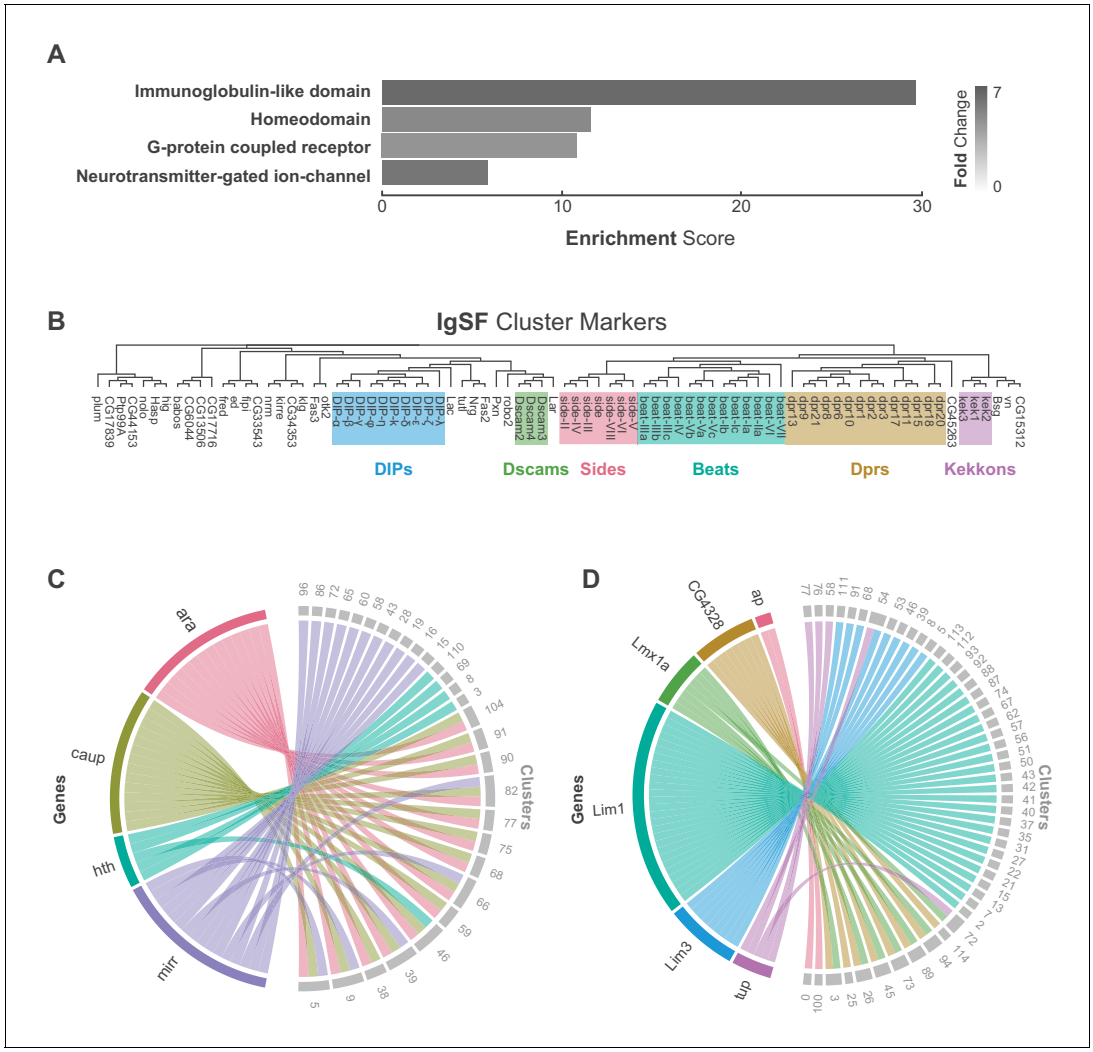


Figure 2. Characterization of neuronal cluster markers. (**A**) Functional analysis of neuronal cluster markers (DAVID) showing representative functional terms (>5 fold-change) and their corresponding enrichment scores (representing mean p-values). (**B**) Phylogenetic tree of IgSF neuronal cluster markers. IgSF subfamily members are highlighted in color. (**C**) Chord diagram comparing the relationship between Tale Homeobox TF cluster markers (left) and the clusters in which they are significantly enriched (right). (**D**) Chord diagram comparing the relationship between LIM Homeobox TF cluster markers (left) and the clusters in which they are significantly enriched (right). (**B–D**) Gene were classified based on Flybase Gene Group annotations (www.flybase.org). (**C,D**) Chord diagrams are visual aids based on data available in ***Figure 1—source data 1***.

The online version of this article includes the following source data and figure supplement(s) for figure 2:

Figure supplement 1. Comparison of central nervous system single-cell data sets.

Figure supplement 1—source data 1. List of marker genes for the 32 clusters shown in ***Figure 2—figure supplement 1A*** of re-analyzed data from ***Croset et al. (2018)***.

Figure supplement 1—source data 2. List of marker genes for the 88 clusters shown in ***Figure 2—figure supplement 1B*** of re-analyzed data from ***Davie et al. (2018)***.

Figure supplement 2. Relationships between IgSF subfamilies in defining cluster identity.

Figure supplement 3. GPCR neuronal cluster markers.

Figure supplement 4. Co-expression of gene family members in the VNC.

Figure supplement 5. Ion channel neuronal cluster markers.

Figure supplement 6. TF neuronal cluster markers.

but not detected) these values may be an underestimate. We saw a median of 12 GPCRs when we consider the expression at the cluster level (with a threshold of 20% positive cells per cluster; ***Figure 2—figure supplement 4B***). Although many GPCRs are significantly enriched in the individual clusters, they were, in general, expressed at low levels, in few cells per cluster, and across many

clusters (***Figure 2—figure supplement 3A***). The neuropeptide receptor family members had more inter-cluster variation in expression levels than the other GPCR subtypes, suggesting that they confer an added level of specialization. We also saw that 80% of biogenic amine receptor family members, a diverse group of neurotransmitter receptors sensitive to biogenic amine neurotransmitters including dopamine, octopamine, serotonin and tyramine, are found as cluster markers (***Figure 2—figure supplement 3B***). Odorant receptors, gustatory receptors, ionotropic receptors, and pickpocket sodium channels are predominantly found in first-order sensory neurons (***Joseph and Carlson, 2015***) and in general, were not expressed in the VNC, as expected. There were exceptions, however, as *Or63a*, *Gr28b*, *Ir76a*, *ppk31* are all expressed in the VNC, albeit at low levels and in few cells. Neurotransmitter-gated ion-channels were also highly enriched as neuronal cluster markers (***Figure 2A***, ***Figure 2—figure supplement 5***, ***Figure 1—source data 1***), including 8 of the 10 nicotinic acetylcholine receptor subunits (***Littleton and Ganetzky, 2000***). Expression of transmembrane receptors that provide fast- and slow-acting responses to neurotransmitters, therefore, contribute to cellular identity.

Finally, over 20% of all *Drosophila* TFs were cluster markers in the adult VNC, the most enriched class of which was the homeodomain (HD) family (***Figure 2A***, ***Figure 2—figure supplement 6***). Homeodomain TFs play central roles in establishing regional-, tissue- and cell-specific fates (***Bürglin and Affolter, 2016***). We observed a median of 6 HD TFs per cell, and a median of 52 TFs in total per cell. We saw a median of 10 HD TFs when we consider the expression at the cluster level (20% positive cells per cluster threshold; ***Figure 2—figure supplement 4D***). We detected high variance in HD TF expression levels across clusters (***Figure 2—figure supplement 6A***). Four members of the three amino acid loop extension (TALE) class of HD TFs mark specific clusters. This included all the three members of the evolutionarily conserved *Iroquois* gene complex (Iro-C) *araucan* (*ara*), *caupolican* (*caup*), and *mirror* (*mirr*) (***Cavodeassi et al., 2001***). The Iro-C arose through two duplication events, one ancient event in arthropods led to independent *ara*/*caup* and *mirr* genes, followed by a more recent event in dipterans which gave rise to the *caup* and *ara* genes (***Kerner et al., 2009***). Cluster expression of Iro-C genes appears to reflect this evolutionary history. Whereas *caup* and *ara* overlap as cluster markers, *mirr* expression is only partially overlapping and *mirr* is an independent marker for 11 additional clusters, suggesting specialization (***Figure 2C***). In addition, we found largely mutually exclusive enrichment of individual members of the LIM class of HD TFs defining cell clusters in the VNC (***Figure 2D***). LIM TFs are known to specify distinct neuronal identities in the embryo (***Thor et al., 1999***). In the VNC only *Lmx1a* and *CG4328* appear to be co-expressed which given their genomic linkage likely represents a relatively recent tandem duplication event. These findings suggest that aspects of the TF-code that establishes cellular identity during development actively maintains neuronal identity throughout the life of the animal (***Deneris and Hobert, 2014***).

Hox genes define neuromere identity

Hox genes encode homeodomain proteins that confer positional identities along the antero-posterior axis in all bilaterian animals (***McGinnis and Krumlauf, 1992***). In both vertebrates and invertebrates, the Hox family of transcription factors are known to govern key aspects of nervous system development, notably the formation of neuromuscular networks (***Philippidou and Dasen, 2013***). The neuromeres of the VNC show distinct segment-specific properties under Hox gene control (***Baek et al., 2013***; ***Suska et al., 2011***).

The Hox genes *Antennapedia* (*Antp*), *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*), and *Abdominal-B* (*Abd-B*), which specify the thoracic, and abdominal segments of the fly VNC, define specific clusters in the single-cell atlas (***Figure 3A***). Hox gene expression patterns were significantly anti-correlated at the individual cell level, except for *abd-A* and *Abd-B* (***Figure 3B***). Similar patterns were seen with the correlation of cluster-level expression, and the patterns were consistent between replicates. *Antp* and *Ubx* are both expressed throughout the t-SNE plot, with many clusters expressing both, albeit in mutually exclusive cells, while *abd-A* and *Abd-B* expression is more restricted and overlapping (***Figure 3C***). Immunostaining the adult VNC with antibodies against these Hox proteins showed belts of expression along the anterior-posterior axis (***Figure 3D***). These patterns of expression are different to those in the larva but are consistent with those observed in the mid-pupal stage (***Baek et al., 2013***). Antp protein is most highly expressed in the mesothoracic neuromere (MesoNm), whereas Ubx expression is highest in the metathoracic neuromere (MetaNm). Abd-A and Abd-B protein showed overlapping expression restricted to the abdominal neuromere (ANm), with

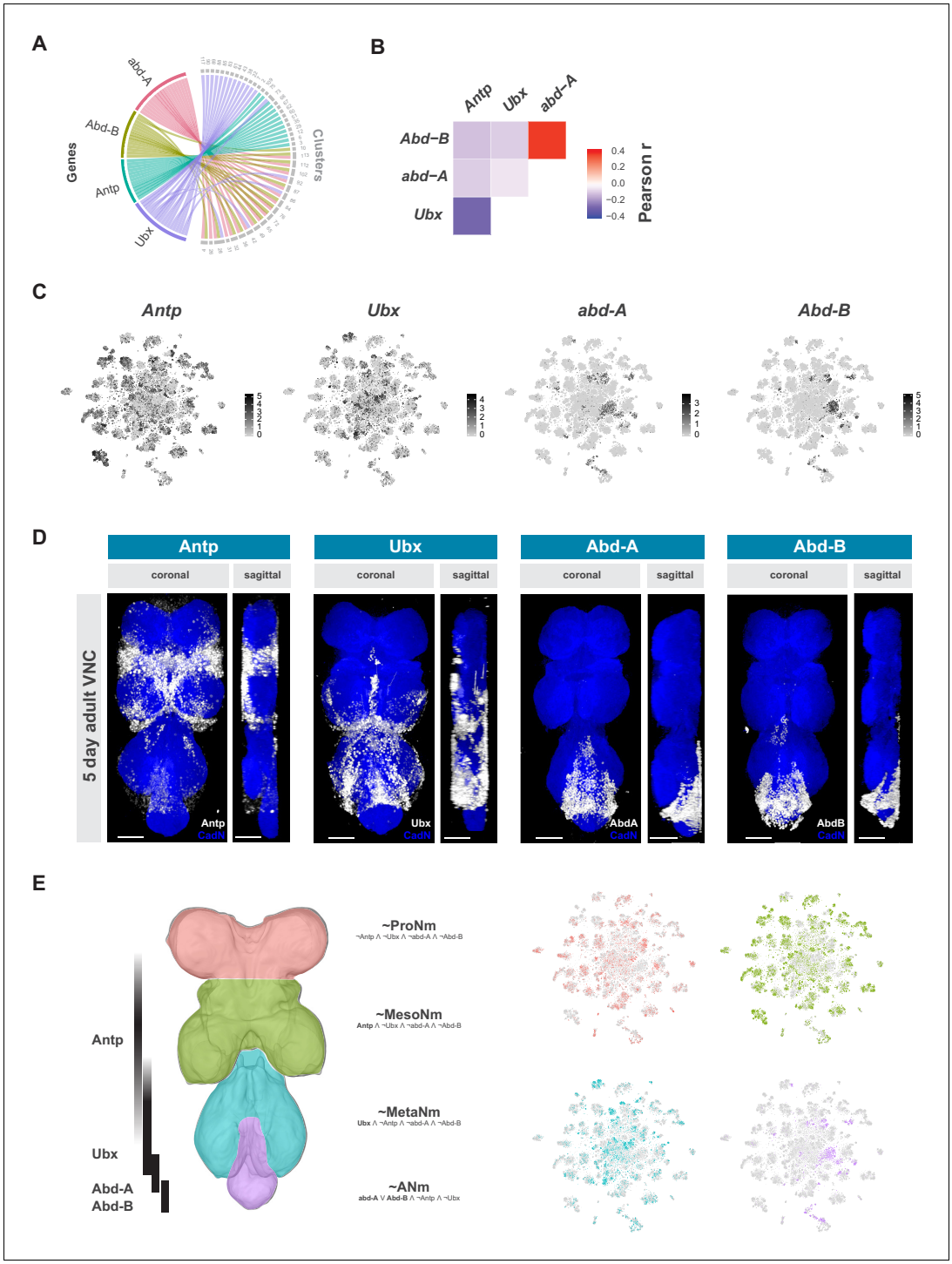


Figure 3. Neuromeres are defined by Hox gene expression in the adult VNC. (**A**) Chord diagram comparing the relationship between Hox genes found as cluster markers (left) and the clusters in which they are significantly enriched (right), based on data available in ***Figure 1—source data 1***. (**B**) Heatmap of Pearson correlation coefficient for Hox gene expression in all single cells. We saw consistent patterns of correlation at the cluster-level expression, and between replicates. (**C**) t-SNE plot of Hox gene expression distribution. Expression shown in black, intensity is proportional to the log-normalized expression levels. (**D**) Visualization of Hox protein expression in the 5 day old adult VNC. Coronal and sagittal views of anti-Antp, -Ubx, -Abd-A, and -Abd-B (white). Neuropil counterstained with anti-Cad-N (blue). Scale bars = 50 µm. (**E**) Schematic representing bands of Hox expression along the anterior-posterior axis of the VNC (left). Cells assigned to approximate neuromeres in VNC t-SNE plots (right) based on differential Hox gene expression: ProNm (pink), MesoNm (green), MetaNm (blue) and ANm

Figure 3 continued on next page

Figure 3 continued

(purple). Expressed genes used in defining each neuromere are in bold. (symbols represent the following: '¬' not, '∧' and, '∨' or, '~' approximate).

Abd-A expressed more anteriorly than Abd-B. By combining the neuromere enriched expression revealed by antibody staining, with the high level of anti-correlation in the single-cell data, we can separate the single-cell clusters into approximations for each neuromere (***Figure 3E***).

Neuroblast lineage identity

Each neuromere contains a repeating set of neuroblast lineages. We used prior knowledge of biomarkers defining the development of post-embryonic lineages (***Bossing et al., 1996***; ***Schmid et al., 1999***) to further annotate the 5 day old adult VNC single-cell atlas. All known lineage biomarkers showed continued expression in the adult VNC, and most were found as significantly enriched cluster markers (***Figure 4A and B***; ***Figure 4—figure supplement 1***; ***Figure 1—source data 1***). These established markers therefore allowed us to assign potential hemilineage identities to many of the cell clusters (***Figure 4B***; ***Figure 1—source data 2***). We also used fast-acting neurotransmitter (FAN) identity as an additional marker to specify potential hemilineages, since FAN usage is acquired in a lineage-dependent manner in post-embryonic neurons in the VNC (***Lacin et al., 2019***). For example, cluster 54 has enriched expression of lineage markers *Lim3*, *tailup* (*tup), abnormal chemosensory jump 6* (*acj6*), and the glutamatergic neuron marker *Vesicular glutamate transporter* (*VGlut*), which collectively suggests this cluster is hemilineage 9B. The correspondence between cluster markers and hemilineage markers shows that cell cluster identity is closely linked to hemilineage identity. We observed a strong concordance between the number of cells within our predicted hemilineages and previous cell counts (***Figure 4—figure supplement 2A***; ***Birkholz et al., 2015***; ***Lacin et al., 2019***). This correlation of cell counts further supports our atlas having an approximately 1.5x coverage of the VNC.

Many of the established lineage markers are expressed at low levels and in few cells (e.g., *B-H1/2*, *exex*, *tup*), while others are robustly expressed (e.g., *acj6*, *toy*, *Lim3*) (***Figure 4—figure supplement 1***). Nevertheless, the restricted cluster-specific expression of these established markers and co-expression with more robust novel markers (***Figure 4C***; ***Figure 4—source data 1***) allowed us to assign cells to putative hemilineages. At present some hemilineages cannot be defined with certainty due to the lack of specific sets of established markers. For instance, there is no marker for hemilineage 2A, other than *VGlut*, and hemilineage 3B's *Dbx* expression does not persist past the pupal stage (***Lacin and Truman, 2016***; ***Lacin et al., 2019***). Similarly, *unc-4* is currently the sole established marker for many cholinergic hemilineages, including 7B, 12A, 18B and 19B. We observe multiple cholinergic clusters enriched for *unc-4* (***Figure 1—source data 1***; ***Figure 4—figure supplement 1***). To assign these clusters to known hemilineages, additional cluster specific enriched genes can be investigated by immunohistochemistry. Conversely, clusters 6, 69, and 110 express markers for both hemilineages 8A (*ems*, *ey*, *VGlut*) and 24B (*ems*, *toy*, *VGlut*), suggesting that these lineages may exhibit very similar transcriptional profiles and resolving them will likely require more cells (***Figure 4C***).

We have identified several novel potential hemilineage marker genes (***Figure 4—figure supplement 1***). Of the 59 most highly enriched hemilineage-defining genes, shown in ***Figure 4C***, 34 are TFs, approximately half of which contain homeodomains. Interestingly, 5 of these novel marker genes encode lncRNAs. lncRNAs have been shown to have complex spatiotemporal regulation throughout the embryo (***Wilk et al., 2016***; ***Karaiskos et al., 2017***) and specifically during embryonic neurogenesis (***Scruggs et al., 2015***), suggesting their functional importance during development. Two of these lncRNAs, *MRE23* and *CR45141*, are adjacent to and transcribed divergently to their nearest neighbour protein coding genes, *fork head* (*fkh*) and *vestigial (vg)*, respectively. These gene pairs show clear positive correlations of expression within the hemilineages (***Figure 4C***), suggesting they share a cis-regulatory landscape. Such divergent transcription has been shown to enhance transcriptional output (***Scruggs et al., 2015***; ***Schor et al., 2018***). As the roles of many of these novel marker genes have yet to be studied either during development or in the adult nervous system, our VNC analysis provides a valuable resource for future investigation of hemilineage cell types.

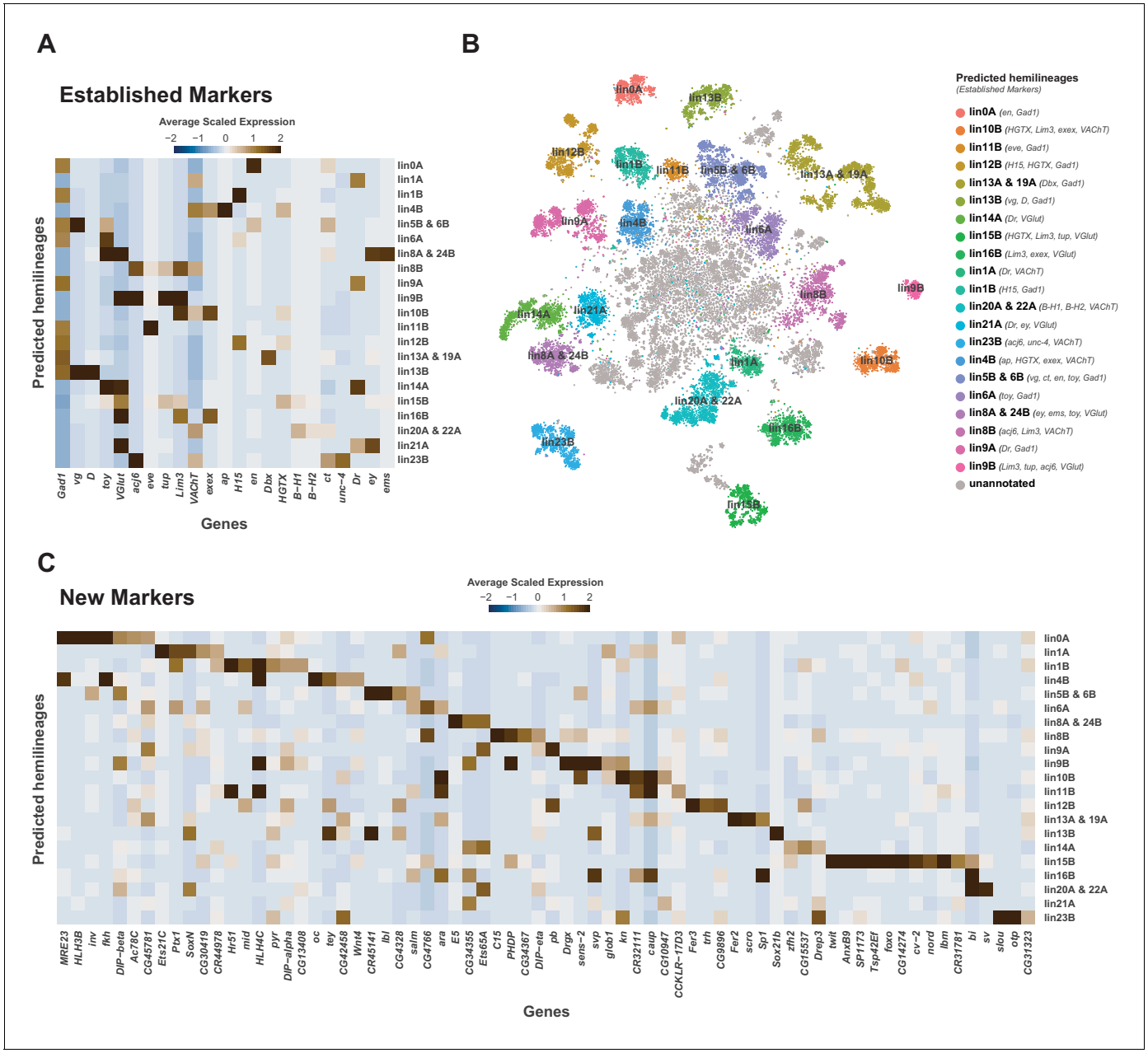


Figure 4. Cellular identities defined by developmental lineages. (**A**) Heatmap of the mean scaled log-normalized expression of established lineage markers genes (bottom) within predicated neuroblast hemilineage (right). (**B**) t-SNE plot of neuronal cells highlighting predicted hemilineages based on the expression of established hemilineage marker genes (shown beside list of predicted hemilineages on the right). (**C**) Heatmap of the mean scaled log-normalized expression of potential new hemilineage markers genes (bottom) within predicated neuroblast hemilineage (right). Full list of predicted hemilineage markers can be found in ***Figure 4—source data 1***.

The online version of this article includes the following source data and figure supplement(s) for figure 4:

Source data 1. List of marker genes for predicated hemilineages shown in ***Figure 4B***.
Figure supplement 1. Neuroblast lineage marker gene expression in the VNC.
Figure supplement 2. Neurodevelopment gene expression.

Many predicted hemilineages contain multiple clusters (***Figure 4B***; ***Figure 1—source data 2***). We propose that these clusters represent distinct variations within a hemilineage, partly informed by birth order and partly related to thoracic vs. abdominal anatomical position. Consistent with a subdivision based on birth order, many of our predicted hemilineages contain a distinct subgroup of

broad (*br*) expressing cells (***Figure 4—figure supplement 2B***). A specific isoform of *br* continues to be expressed in the adult and marks cells born around L2-L3 ecdysis (***Zhou et al., 2009***). Other hemilineage sub-division was evident when visualizing expression of neurodevelopmental genes such as *prospero* (*pros*), *datilografo* (*dati*), *maternal gene required for meiosis* (*mamo*), and *IGF-II mRNA-binding protein* (*Imp*) (***Figure 4—figure supplement 2C***). Expression of *pros* and *Imp* in the VNC were negatively correlated (***Figure 4—figure supplement 2D***), as previously observed in the adult brain (***Davie et al., 2018***). In contrast, expression of *pros* was highly correlated with that of *dati*, while *Imp* and *mamo* were highly correlated, suggesting these pairs of genes may be co-regulated in the adult (***Figure 4—figure supplement 2D***). *pros* and *Imp* also showed an anatomical bias, with *pros* being negatively correlated with abdominally expressed *abd-A* and *Abd-B*, while *Imp* is positively correlated with these markers (***Figure 4—figure supplement 2E***). However, we did not observe a striking physiological bias, as neither *pros* nor *Imp* showed correlation with fast-acting neurotransmitter markers (***Figure 4—figure supplement 2F***). We saw consistent patterns of correlation at the cluster-level expression, and between replicates. It has previously been shown that *pros* labels late-born motor neurons from hemilineage 15B (***Baek et al., 2013***), and that a trade-off of *Imp* and *pros* expression is essential for cell cycle exit of post-embryonic type I neuroblasts in the VNC (***Maurange et al., 2008***; ***Yang et al., 2017***). Studies investigating the transcriptomes of mature (embryonic) vs. immature (post-embryonic) neurons in the larval VNC have shown that *Imp* is enriched in embryonic neurons, and *pros* is enriched in post-embryonic neurons (***Etheredge, 2017***). We propose that the continued expression of *Imp* in the adult VNC labels early-born and embryonic neurons while *pros* labels late-born post-embryonic neurons.

Gene expression defines neuronal subtypes within a hemilineage

Hemilineage 23B marked by *acj6*, *unc-4*, and *VAChT* has four distinct clusters that can only be partially explained by potential birth order and neuromere identity (***Figure 3***; ***Figure 4***). *acj6* is also a marker for the cholinergic hemilineage 8B and the glutamatergic hemilineage 9B (***Figure 4A***; ***Lacin et al., 2019***). We examined *acj6* expressing cells in more detail by visualizing the anatomical distribution of *acj6* expressing cell bodies and neuronal projections in the VNC using $acj6^{GAL4}$ to express the GAL4-responsive dual reporter *UAS-Watermelon* (WM), which simultaneously marks the plasma membrane with GFP and the cell nucleus with mCherry (***Figure 5A***; ***Lee et al., 2018***). We also validated the accuracy of $acj6^{GAL4}$ using an anti-Acj6 antibody (***Clyne et al., 1999***; ***Figure 5—figure supplement 1***). Consistent with previous reports, $acj6^{GAL4}$ labeled distinct clusters of cells in the three thoracic neuromeres of the VNC, which predominantly innervate the leg neuropils. (***Figure 5A***; ***Harris et al., 2015***; ***Shepherd et al., 2019***). The *acj6* expressing neurons were confirmed as hemilineages 8B, 9B, and 23B (D Shepherd, pers. comm.) on the basis of their primary neurite projections (***Shepherd et al., 2016***).

The *acj6* expressing clusters in the single-cell data set are highlighted in ***Figure 5B***. Based on co-expression of additional markers, we could assign *acj6*-expressing hemilineages within the VNC data; lin8B expresses *acj6*, *Lim3*, and *VAChT*; lin9B *acj6*, *Lim3*, *VGlut*, and *tup*; Lin23B *acj6*, *unc-4*, and *VAChT* (***Figure 5C***). Of the 4 distinct clusters within predicted hemilineage 23B, the genes *knot* (*kn/collier*) and *tiwaz* (*twz*) (***Crozatier et al., 1996***; ***Williams et al., 2014***) are significantly enriched in two independent clusters; 51 and 93, respectively (***Figure 5D***; ***Figure 1—source data 1***). We validated the co-expression evident in the single-cell data using GAL4 lines that represent *kn* and *twz* expression, and co-stained these VNCs with anti-Acj6 antibody (***Figure 5E and F***; ***Figure 5—figure supplement 2***). We also used a reporter for the established marker *Lim3*, which is co-expressed with *acj6* in hemilineages 8B and 9B, but not 23B (***Figure 5B***; ***Lacin and Truman, 2016***). Hemilineage 23B is located dorsolaterally at the posterior edge of each of the three thoracic neuromeres, whereas 8B and 9B are both located laterally at the anterior side of each thoracic neuromere with 8B located near the ventral surface and 9B located more dorsally (***Figure 5E***; ***Lacin et al., 2019***; ***Shepherd et al., 2019***). Focusing on the ProNm-MesoNm border, which encompasses prothoracic 23B and mesothoracic 8B and 9B, we found that a subset of 23B co-expressed *kn* and a subset co-expressed *twz* (***Figure 5F***; ***Video 1***; ***Video 2***). Of the 40 dorsolateral $Acj6^{+}$ 23B cells, 14 were kn^{+} and 9 were twz^{+} (***Figure 5—source data 1***; ***Figure 5—figure supplement 3***). These proportions (34% and 21%) are roughly consistent with that seen in the single-cell data (31% and 18% for *kn* and *twz,* respectively). *Lim3* expression was restricted to 8B and 9B, with 90% of $Acj6^{+}$ 8B and 9B cells expressing *Lim3,* and 0% co-expression in hemilineage 23B (***Figure 5F***; ***Figure 5—source data 1***;

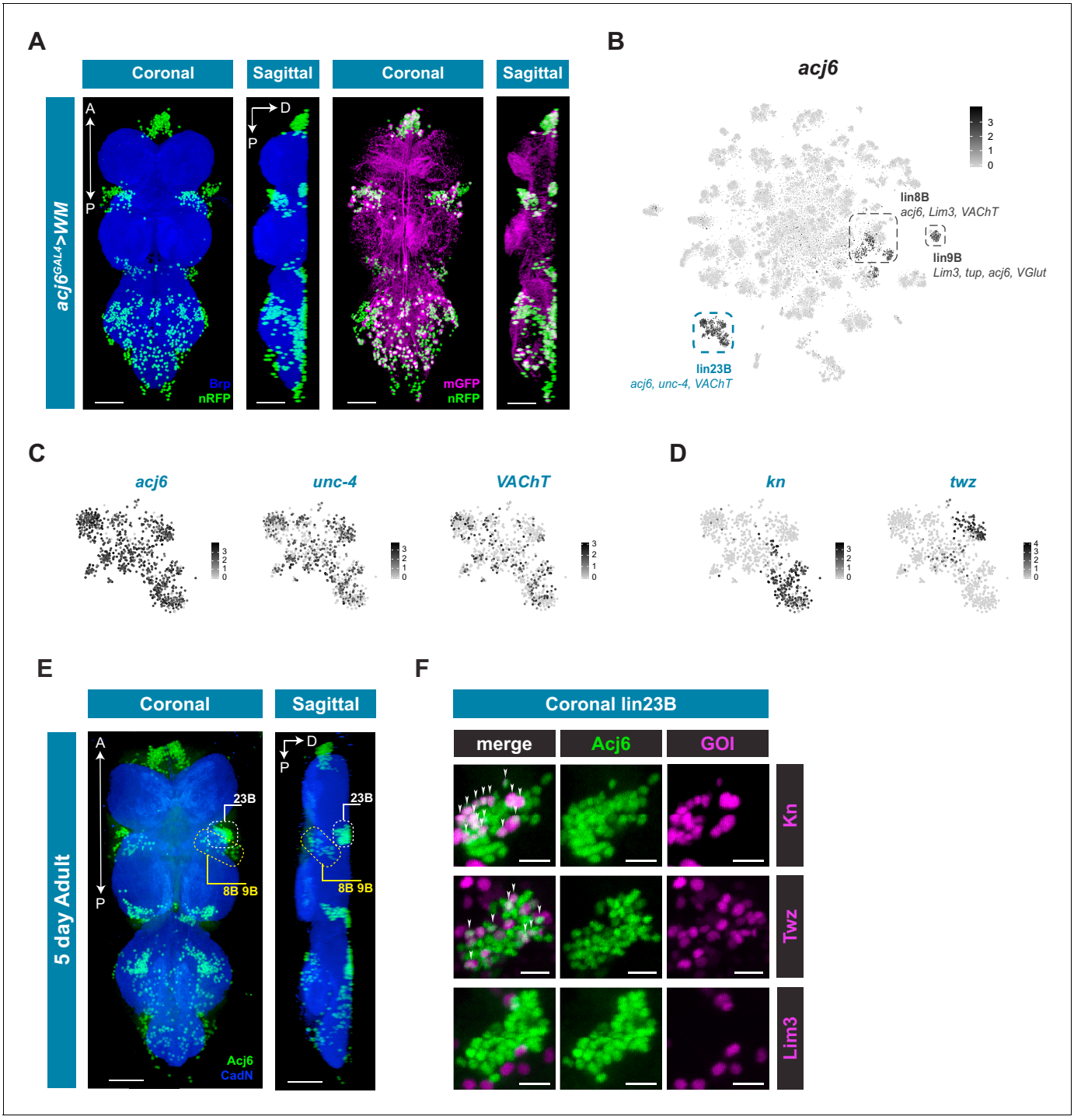


Figure 5. *acj6* expression, hemilineage identity, and novel sub-lineage marker co-expression in the VNC. (**A**) Visualization of *acj6* expressing cells in the 5 day old adult VNC. Maximal coronal projection, and hemi-maximal sagittal projectionof $acj6^{GAL4}$ driving expression of *UAS-WM*, enabling the visualization of cell nuclei (nRFP; green) and neuronal morphology (mGFP; magenta). Neuropil is counterstained (Brp; blue). A, Anterior, P, posterior, D, dorsal; scale bars = 50 µm. (**B**) t-SNE plot of *acj6* expression in the VNC. Predicted hemilineage 8B, 9B, and 23B, based on known markers (shown below) are highlighted. (**C**) t-SNE plots of predicted hemilineage 23B, showing expression of markers *acj6*, *unc-4*, and *VAChT*. (**D**) t-SNE plots of predicted hemilineage 23B, showing expression of novel sub-type markers *kn* and *twz*. (**B–D**) Expression shown in black, intensity is proportional to the log-normalized expression levels. (**E**) Maximal coronal view (left) and hemi-maximal sagittal view (right) of Acj6 expression (green) in the 5 day old adult VNC. The posterior prothoracic hemilineages 23B (dashed white box) and the anterior mesothoracic 8B and 9B (dashed yellow box) are marked. A, Anterior, P, posterior, D, dorsal; Neuropil is counterstained (Cad-N; blue) scale bars = 50 µm. (**F**) Close-up maximal coronal views of dorsolateral hemilineage 23B showing the co-expression of kn^{GAL4} and twz^{GAL4} driven *UAS-Stinger* expression (nGFP; magenta) with Acj6 (green). Co-positive cells

Figure 5 continued on next page

Figure 5 continued

(white) are indicated with arrow heads. *Lim3*^{GAL4} driven expression did not colocalize with Acj6 in hemilineage 23B. GOI = gene of interest, scale bars = 10 μm.

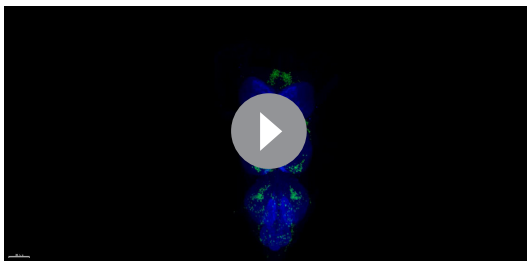
The online version of this article includes the following source data and figure supplement(s) for figure 5:

Source data 1. Cell counts of novel sub-lineage marker co-expression in the VNC.
Figure supplement 1. *acj6*^{GAL4} recapitulates Acj6 protein expression in the adult VNC.
Figure supplement 2. Novel sub-lineage marker expression in the VNC.
Figure supplement 3. Single slices of novel sub-lineage marker co-expression in the VNC.

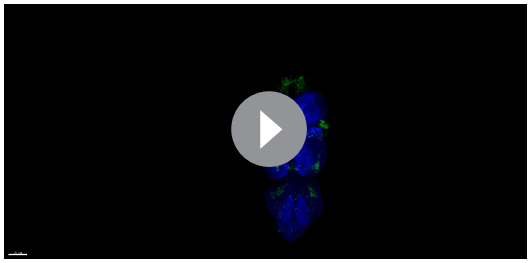
Video 3; ***Figure 5—figure supplement 3***). *kn* was also expressed in a subset of predicted hemilineage 9B in the single cell data (***Figure 5—figure supplement 2***), and co-expression was seen between *kn*^{GAL4} and Acj6 in the 8B/9B cluster *in vivo* (***Figure 5—figure supplement 3***; ***Figure 5—source data 1***). Full expression patterns in the VNC are shown in ***Figure 5—figure supplement 2***. These findings highlight the predictive power of the single-cell transcriptome to identify new markers that refine our understanding of hemilineage subtypes which contribute to neuronal cell diversity in post-embryonic lineages.

Classification using fast-acting neurotransmitters

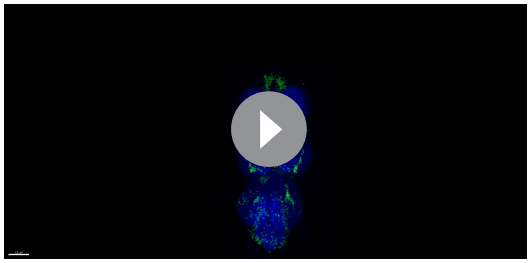
A recent comprehensive map of fast-acting neurotransmitter (FAN) usage across the VNC found that neurons within a post-embryonic hemilineage use the same neurotransmitter; either acetylcholine, gamma-aminobutyric acid (GABA), or glutamate, unifying their shared developmental and functional identities (***Lacin et al., 2019***). We examined FAN usage in the VNC using expression of established biosynthetic and vesicular loading markers: *Choline acetyltransferase* (*ChAT*) and *Vesicular acetylcholine transporter* (*VAChT*) for acetylcholine, *VGlut* for glutamate, and *Glutamic acid decarboxylase 1* (*Gad1*) and *Vesicular GABA Transporter* (*VGAT*) for GABA (***Figure 6A***). Clusters showed mutually exclusive enrichment of *VAChT*, *VGlut*, and *Gad1* (***Figure 6B***). However, this exclusivity partially breaks down at the cell-by-cell level. Although expression levels of *VAChT*, *VGlut*, and *Gad1* were negatively correlated (***Figure 6—figure supplement 1A***), we also observed significant levels of co-expression, with 31% of cells expressing at least two of these markers (***Figure 6—figure supplement 1B***). A similar pattern of co-expression was observed in single-cell data from the adult brain (***Croset et al., 2018***; ***Davie et al., 2018***). However, immunostaining of the intact adult VNC suggested that cytoplasmic co-expression of these markers does not occur (***Lacin et al., 2019***). Nuclear transcriptomic profiles of neuronal subtypes from both the visual system and the mushroom body found no evidence for FAN co-release (***Davis et al., 2020***; ***Shih et al., 2019***). Co-expression in the VNC may therefore represent contamination due to ambient RNA present in the cell suspension (discussed below). Since the cells cluster due to hemilineage identity and hemilineages are restricted to a single FAN, we assigned FAN identity by comparing the average expression of these FAN markers at the cluster level (***Figure 6—figure supplement 1C***; ***Figure 1—source data 2***). Amongst the co-expressing cells, the expression of FAN makers is negatively correlated, and cells assigned to one FAN identity tended to show higher expression of the corresponding marker (***Figure 6—figure supplement 1D***). This is, however, not always the case, presumably due to the much higher expression levels seen for *VGlut* and *Gad1* (***Figure 6A***; ***Figure 6—figure supplement 1D***). With these criteria (see methods), we estimate that the VNC is 40% cholinergic, 38% GABAergic, and 18% glutamatergic. The remaining 4% of neurons did not show a strong signature for any FAN (***Figure 6C and D***). These proportions differ to those in the adult brain (***Croset et al., 2018***; ***Davie et al., 2018***), with inhibitory GABAergic neurons being much more prominent in the VNC (38% vs 15% in the adult



Video 1. Video of 3D volume showing the co-expression of Acj6 (anti-Acj6; green) with *kn*^{GAL4} driven expression of *UAS-Stinger* (nGFP; magenta) in the 5 day adult VNC, with neuropil counterstained (CadN; blue).
https://elifesciences.org/articles/54074#video1



Video 2. Video of 3D volume showing the co-expression of Acj6 (anti-Acj6; green) with *twz*^{GAL4} driven expression of *UAS-Stinger* (nGFP; magenta) in the 5 day adult VNC, with neuropil counterstained (CadN; blue).
https://elifesciences.org/articles/54074#video2



Video 3. Video of 3D volume showing the co-expression of Acj6 (anti-Acj6; green) with *Lim3*^{GAL4} driven expression of *UAS-Stinger* (nGFP; magenta) in the 5 day adult VNC, with neuropil counterstained (CadN; blue).
https://elifesciences.org/articles/54074#video3

midbrain). The inhibitory glutamate receptor, *GluClα* (**Cully et al., 1996**), was broadly expressed in

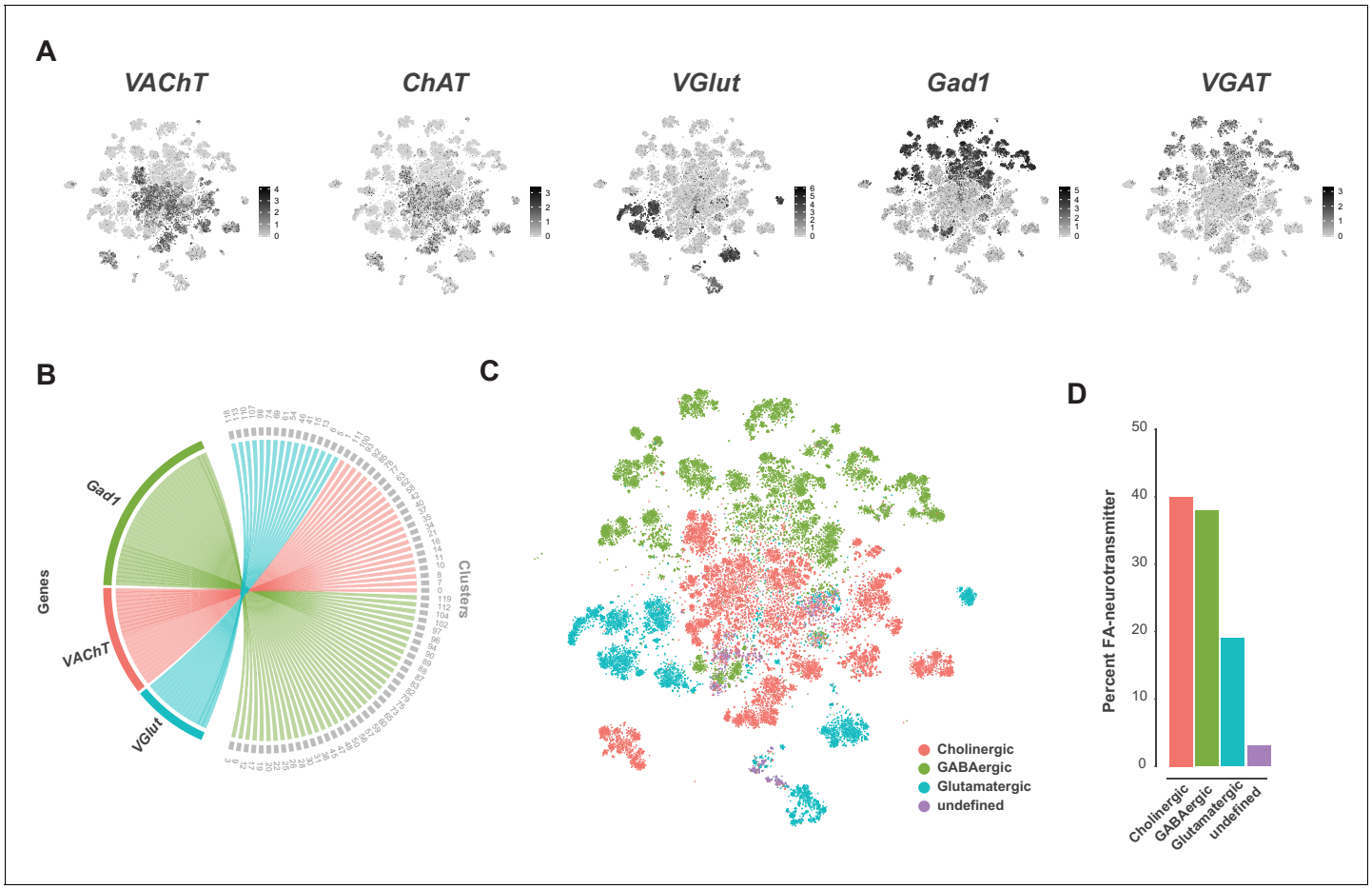


Figure 6. Fast-acting neurotransmitter usage in the VNC. (**A**) t-SNE plots showing distribution of cells expressing biomarkers for the fact-acting neurotransmitters acetylcholine (*VAChT*, *ChAT*), glutamate (*VGlut*) and GABA (*VGAT*, *Gad1*). Expression shown in black, intensity is proportional to the log-normalized expression levels. (**B**) Chord diagram comparing the relationship between fast-acting neurotransmitter cluster markers (left, genes) and the clusters in which they are significantly enriched (right, clusters). (**C**) t-SNE plot colored according to fast-acting neurotransmitter usage based on assigned cluster identity. (**D**) Percentage of cells in the VNC assigned as releasing distinct fast-acting neurotransmitters.
The online version of this article includes the following figure supplement(s) for figure 6:

Figure supplement 1. Expression of fast-acting neurotransmitter biomarkers in the VNC.

these data and is particularly enriched in cluster 103 (see ***Figure 1—source data 1***), suggesting that a portion of glutamatergic signaling is also inhibitory. We expect this relative abundance of inhibitory signaling reflects the importance of inhibition in the initiation, maintenance, and termination of complex motor programs (***Burrows, 1992***; ***Grillner, 2006***), especially in coordination between bilateral and intersegmental circuits (***Gowda et al., 2018***).

Monoaminergic neurons

Cells that produce and likely release monoamines in the VNC can be identified based on expression of the *Vesicular monoamine transporter* (*Vmat*) gene (***Greer et al., 2005***). To examine the anatomical distribution of *Vmat* expressing nuclei and projections in the VNC we combined $Vmat^{GAL4}$ with *UAS-Watermelon* (WM) (***Figure 7A***). $Vmat^{GAL4}$ labels 115 ± 6.3 neurons distributed across every neuromere of the VNC (n = 9, data not shown). Each neuromere has a main cluster of cells located on the ventral surface, near the midline, consistent with known developmental origins of monoaminergic neurons. Serotonin- and dopamine-producing neurons are born from the ventromedially located paired neuroblast 7–3, and octopamine-producing neurons are born from the ventral midline unpaired median neuroblast (***Bossing et al., 1996***; ***Schmid et al., 1999***). Projections of *Vmat*-labeled neurons formed thick fascicles of fibers projecting from each cluster to the dorsal surface of the VNC (***Figure 7A***, sagittal view). Each thoracic neuromere has one main fiber tract, whereas the ANm has 8 parallel tracts, one for each abdominal segment. Despite the relatively small number of *Vmat* expressing cells, they project throughout the VNC and densely innervate the entire neuropil. Additional *Vmat* nuclei associated with the VNC are consistent with its reported expression in perineural surface glia (***DeSalvo et al., 2014***). Surface glial *Vmat* expressing cells can be seen in cluster 98 of our primary t-SNE plot (***Figure 1—figure supplement 4***; ***Figure 1—source data 1***).

Neurons that express *Vmat* co-cluster in the single-cell data (Clusters 72 and 84, ***Figure 1—figure supplement 4***; ***Figure 7B***). To define the specific monoaminergic identity of *Vmat* expressing neurons, we sub-clustered these $Vmat^{+}$ cells from clusters 72 and 84 (***Figure 7C***) and identified distinct groups synthesizing specific monoamines, as defined by expression of known biosynthetic and transporter markers (***Figure 7D***; ***Martin and Krantz, 2014***). All replicates contributed to all clusters and showed high correlation of cluster level expression (***Figure 7—figure supplement 1***).

Histaminergic neurons

Histamine (HA) is well established as the primary fast neurotransmitter of adult photoreceptors in many insects, including *Drosophila* (***Hardie, 1987***; ***Nässel et al., 1988***). The VNC contains 18 ventral HA-immunoreactive neurons with extensive axons and projections, six in the thoracic neuromeres and twelve in the abdominal neuromere (***Buchner et al., 1993***; ***Nässel et al., 1990***). Potential new markers for HA neurons, distinguishing them from other monoaminergic neurons, include *Frequenin 1* (*Frq1*), a Ca^{2+}-binding protein that regulates neurotransmitter release (***Dason et al., 2009***), and *CG43795*; an uncharacterized GPCR predicted to have Glutamate/GABA receptor activity (***Agrawal et al., 2013***).

Tyraminergic and octopaminergic neurons

Tyramine (TA) and Octopamine (OA) neurons innervate and modulate many tissues throughout the fly, including female and male reproductive systems, skeletal muscles, and sensory organs (***Pauls et al., 2018***; ***Rezával et al., 2014***). *Tdc2* encoded tyrosine decarboxylase catalyzes the synthesis of TA (***Cole et al., 2005***), which can also be converted to OA by *Tbh* encoded Tyramine beta-hydroxylase (***Monastirioti et al., 1996***). All *Tdc2* expressing neurons in the VNC also express *Tbh* suggesting that they are likely to be octopaminergic, consistent with a previous report using $Tdc2^{GAL4}$ in the VNC (***Pauls et al., 2018***). Prior work in the larval VNC suggested that the TA:OA ratio might be altered in a state-dependent manner (***Schützler et al., 2019***).

Dopaminergic neurons

Dopamine (DA) is a critical neuromodulator controlling learning and state-dependent plasticity in the fly brain (***Cognigni et al., 2018***). Dopamine and DA neurons in the VNC have been implicated in motor behaviors such as grooming and copulation (***Crickmore and Vosshall, 2013***; ***Yellman et al., 1997***). Novel DA markers in the VNC include *beat-Ib* and *beat-Ic*, a likely recent duplication in the

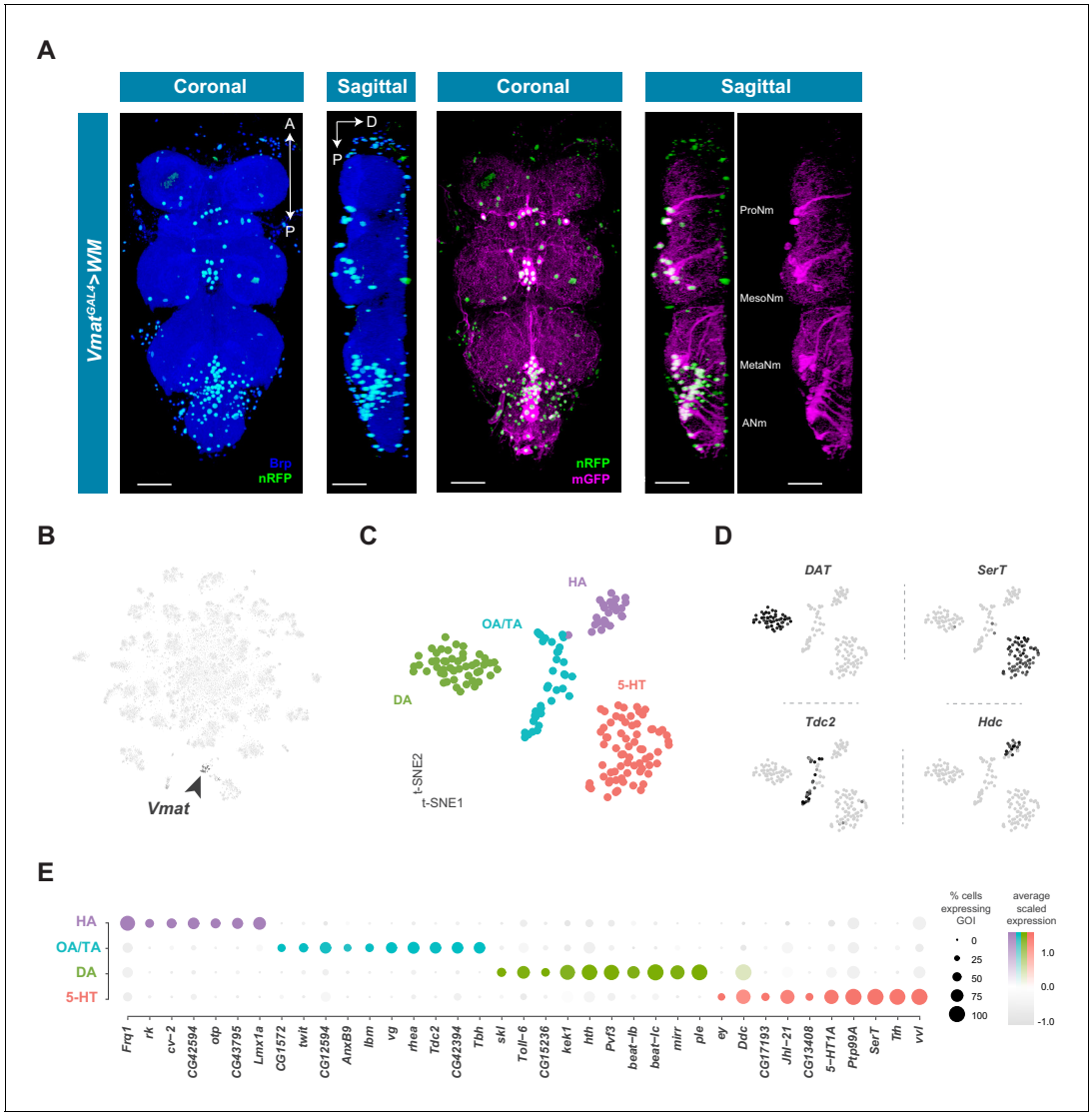


Figure 7. Identification and characterization of monoaminergic cell sub-types. (**A**) Visualization of *Vmat* expressing cells in the 5 day old adult VNC. Coronal and sagittal views of $Vmat^{GAL4}$ driving expression of *UAS-WM*, enabling the visualization of cell nuclei (nRFP; magenta) and neuronal morphology (mGFP; green). Neuropil is counterstained (Brp; blue). A, Anterior, P, posterior, D, dorsal; scale bars = 50 µm. (**B**) t-SNE plot of *Vmat* expression in black, intensity is proportional to the scaled log-normalized expression level. *Vmat* enriched Cluster 72 is highlighted with arrowhead. (**C**) t-SNE plot showing sub-clustering analysis of $Vmat^{+}$ cells from clusters 72 and 84. Four sub-clusters are identified representing dopaminergic (DA), octopaminergic/tyraminergic (OA/TA), histaminergic (HA), and serotonergic neurons (5-HT). (**D**) Expression of established monoaminergic subtype-specific biomarkers used to determine cluster identity. *Histidine decarboxylase* (*Hdc*) labels Histamine (HA) neurons, *Tyrosine decarboxylase 2* (*Tdc-2*) labels Tyramine (TA) and Octopamine (OA) neurons, *Dopamine transporter* (*DAT*) labels Dopamine (DA) neurons and *Serotonin transporter* (*SerT*) labels Serotonin (5-HT) neurons. (**E**) Dot plot of the top genes in each monoaminergic sub-type based on fold-enrichment (***Figure 8—source data 1***). Size of dots represent percent of cells in cluster expressing gene of interest (GOI); intensity of color reflects average scaled expression. ***Figure 7—source data 1*** contain the full list of significantly enriched genes.

The online version of this article includes the following source data and figure supplement(s) for figure 7:

Source data 1. List of marker genes for the $Vmat^{+}$ sub-clusters shown in ***Figure 7C***.
Figure supplement 1. Individual replicate contribution to $Vmat^{+}$ sub-clusters.

beat family of genes which act as heterophilic cell-cell adhesion molecules (***Pipes et al., 2001***). VNC DA neurons overexpress *PDGF- and VEGF-related growth factor* (*Pvf3*) and *kekkon 1* (*kek1*) which encodes a transmembrane protein that binds the EGF receptor and controls the activity of this pathway. Adult midbrain DA neurons also overexpress these genes (***Croset et al., 2018***), suggesting a possible common role for *Pvf3* and *kek1* in DA neuron function. The *Toll-6* neurotrophin-like

receptor implicated in neuronal survival and motor-axon targeting (**McIlroy et al., 2013**), appears specifically enriched in VNC DA neurons.

Serotonergic neurons

Serotonin (5-hydroxytryptamine, 5-HT) releasing neurons within the thoracic neuromeres of the VNC modulate walking speed in a context-independent manner as well as in response to startling stimuli (**Howard et al., 2019**). In the abdominal ganglion, two clusters of sexually dimorphic neurons expressing 5-HT and *fruitless* (approximately ten cells per cluster in males) innervate the internal reproductive organs. These abdominal clusters control transfer of sperm and seminal fluid during copulation (**Billeter et al., 2006**; **Lee and Hall, 2001**; **Lee et al., 2001**). 5-HT neurons in the VNC express the 5-HT receptor *5-HT1A*, suggesting auto-regulatory/autocrine control of 5-HT neurons. VNC 5-HT neurons also express the amino acid transporter encoded by *juvenile hormone inducible 21 (JhI-21)* which is required for 5-HT-dependent evaluation of dietary protein (**Ro et al., 2016**). Finding *JhI-21* expression in 5-HT neurons in the VNC, therefore suggests that some 5-HT neurons may directly sense circulating amino acids.

A transcription factor code for monoaminergic neurons

The top group of genes defining each subclass of monoaminergic neurons types also included unique expression of Homeobox-containing TFs (**Bürglin and Affolter, 2016**; **Figure 7E**). *orthopedia* (*otp*) and the *LIM homeobox transcription factor 1 alpha* (*Lmx1a*) define the HA cluster. DA neurons are enriched for the Hox co-factor *homothorax* (*hth*) as well as the Iroquois homeobox TF *mirror* (*mirr*). Lastly, 5-HT neurons express the homeobox TFs *ventral veins lacking* (*vvl*), *eyeless* (*ey*) and *Lim3*. Although these different combinations of homeobox genes likely act in these cells to specify position-specific patterning decisions and wiring specificity during development, it is also possible that they contribute to the function of mature monoaminergic neurons.

Peptidergic neurons

Neuropeptides are the largest family of signaling molecules in the nervous system and act as important regulators of development, physiology, and behavior (**Nässel and Zandawala, 2019**). We analyzed neuropeptide expression in the adult VNC by looking at the expression of genes encoding neuropeptide precursors, from which active neuropeptides are derived. We identified 28 neuropeptides that are expressed in at least one cell, at a level of 10 or more transcripts per cell (**Figure 8A**). Some neuropeptides are known to be co-expressed with FANs, where they serve to increase signaling flexibility within neural networks (**Croset et al., 2018**; **Nässel, 2018**; **Nusbaum et al., 2017**). We found that in the VNC some neuropeptide genes are predominantly co-expressed with particular FANs, e.g. *sNPF* and *spab* in cholinergic neurons, *MIP* and *CCHa2* in GABAergic neurons, and *Ilp7* and *Proc* in glutamatergic neurons (**Figure 8A**; **Figure 6**). A few neuropeptides were also co-expressed in the same cells, e.g. *AstC* and *CCHa2* in GABAergic cluster 94. Some of these associations are similar to what was observed previously in mid-brain single-cell data (**Croset et al., 2018**), e.g. *sNPF* and *spab* in cholinergic neurons, but others are strikingly different. *AstC* and *CCHa2* are robustly expressed in GABAergic neurons in the VNC but are notably absent from *Gad1* expressing neurons in the mid-brain.

Some neuropeptides are also expressed in 'neurosecretory' neurons that lack FANs. Peptides released from these cells into circulation provide neuroendocrine signals, whereas those released into the nervous system have neuromodulatory function. Cells in cluster 84 exhibit several characteristics of neurosecretory cells. The TF *dimmed* (*dimm*) is required for the differentiation of neurosecretory cells (**Hamanaka et al., 2010**; **Hewes et al., 2003**; **Park et al., 2008**). Although *dimm* is detected at very low levels and in few cells in our dataset, it showed a bias to cluster 84 (10.7% of cells express *dimm* vs. 0.03% elsewhere). Cluster 84 is also particularly enriched for neuropeptide expression (**Figure 8A and B**) and lacking strong expression of FAN or monoaminergic neuron markers (**Figure 1—figure supplement 7**; **Figure 6C**; **Figure 7B**). They also express several genes involved in neuropeptide processing (see **Figure 8C and D**; **Figure 1—source data 1**). Most neuropeptides are expressed in non-overlapping subsets of the cluster (**Figure 8E**), suggesting that a particular neuropeptide does not drive the unifying identity of the cluster, but rather that it is generically peptidergic. The *Glycoprotein hormone alpha 2* (*Gpa2*) and *Glycoprotein hormone beta*

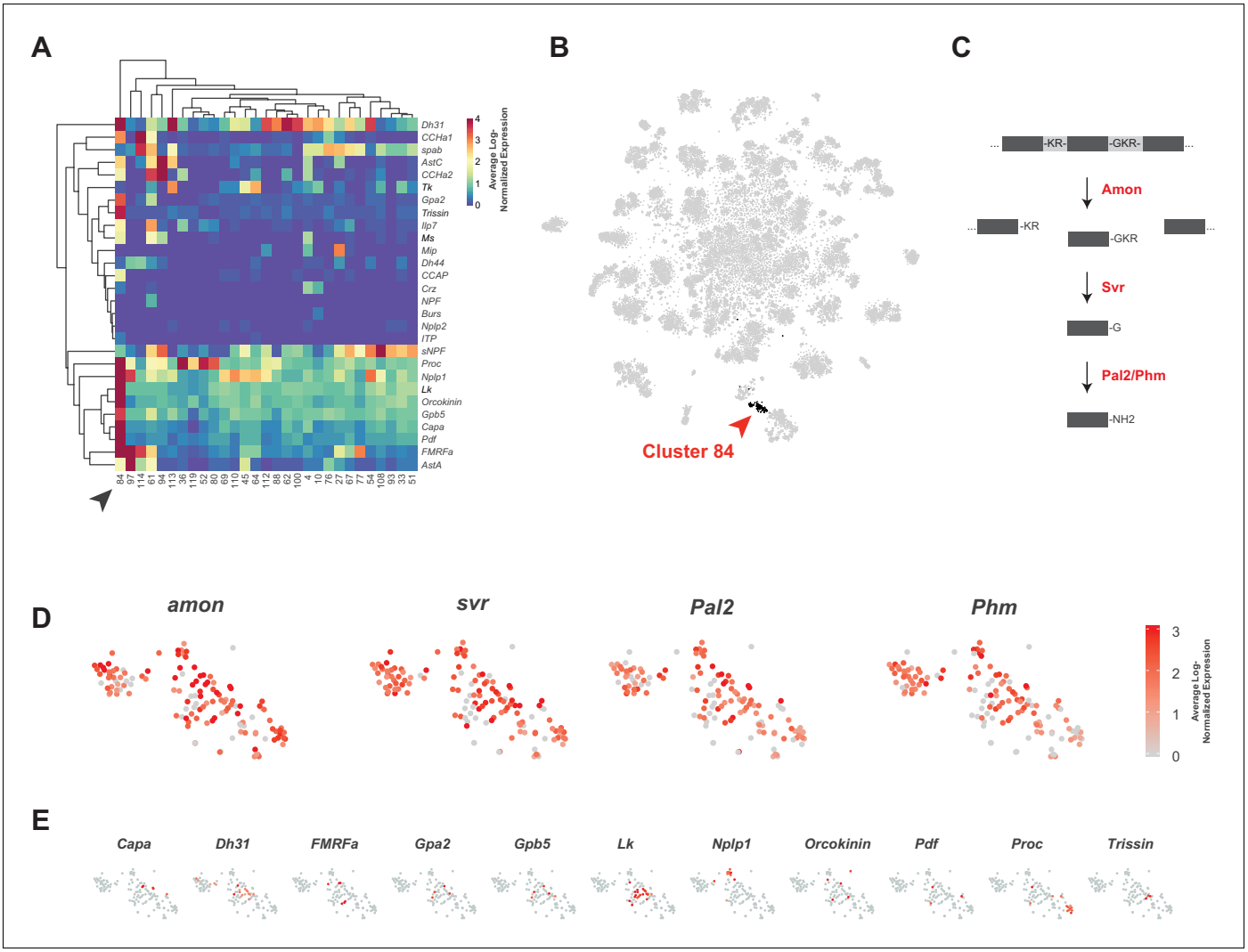


Figure 8. Identification of neuropeptide expressing cells. (**A**) Heatmap of the mean log-normalized expression (0–4), by cluster (bottom), of neuropeptide genes (right). Only clusters with at least one neuropeptide gene expressing in excess of mean log-normalized value of 10 are shown. Only neuropeptide genes that were expressed in at least one cell at a level >10 transcripts per cell are shown. Arrowhead highlights cluster 84. Dendrograms represent hierarchical clustering. (**B**) t-SNE plot highlighting neuropeptide expressing cluster 84. (**C**) Schematic of neuropeptide processing steps highlighting enzymes (in red) identified as enriched markers for cluster 84. Propeptides are cleaved by the pro-hormone convertase *Amontillado* (*Amon*); the carboxypeptidase *Silver* (*Svr*) then removes the C-terminal cleavage sequence. C-terminal amidation occurs through the combined actions of *Peptidyl-α-hydroxyglycine-α-amidating lyase 2* (*Pal2*) and *Peptidylglycin-α-hydroxylating monooxygenase* (*Phm*) (reviewed in ***Pauls et al., 2014***). (**D, E**). Expression of neuropeptide processing enzymes (**D**) and multiple neuropeptide genes (**E**) in t-SNE cluster 84. Intensity of red is proportional to the log-normalized expression level.

The online version of this article includes the following source data and figure supplement(s) for figure 8:

Source data 1. Neuropeptide expression levels.
Figure supplement 1. Neuropeptide gene expression levels.
Figure supplement 2. Analysis of VNC data removing upper limit on transcript number (UMI) per cell.
Figure supplement 3. *Orcokinin* expression analysis.

5 (*Gpb5*) genes, whose peptides form heterodimers, are co-expressed in the same cells in this peptidergic cluster (***Sellami et al., 2011***).

High neuropeptide gene expression reveals important technical considerations

When investigating neuropeptide expressing cluster 84 (NP cluster) we noticed that the median number of transcripts in the cluster was 5277, which was far greater than that observed in the data set as a whole, 2497 (***Figure 8—figure supplement 1A***). This large number of transcripts is due, in part, to high expression levels of the neuropeptide genes themselves. For example, the maximum expression level of *Leucokinin* (*Lk*) was 2629 transcripts in a single cell (***Figure 8—figure supplement 1B***), whereas the maximum expression level of an average gene was 11 transcripts per cell. Amongst neuropeptide genes, *Lk* is not an anomaly, as 18 of the top 40 most highly expressed genes per cell encode neuropeptides (***Figure 8—figure supplement 1C***). All 28 of the expressed neuropeptide genes had a maximum observed expression above the 93rd percentile (see methods), with 23 of the 28 above the 99th percentile (***Figure 8—source data 1***). In some cases, these genes represented 40% of a cell's captured transcriptional output (***Figure 8—figure supplement 1D***). Similar patterns of high neuropeptide precursor gene expression have also been seen in the mouse cortex (***Smith et al., 2019***).

For all previous data analyses, we used an upper limit of 10,000 total transcripts (nUMI) per cell to remove potential outliers. Given the high expression seen for neuropeptide genes, we repeated our analysis without this cut-off, revealing that many cells in the neuropeptide (NP) and *Proctolin+* motor neuron (*Proc*$^{+}$ MN) clusters express more than 10,000 transcripts (***Figure 8—figure supplement 2***). It is worth considering the filtering cut-off when studying cells with high transcriptional output, and the genes expressed therein, whose expression can exceed the threshold. For example, the maximum observed expression of *Orcokinin* was 23,092 transcripts.

Another important consideration when investigating cells expressing genes at very high levels, as many neuropeptide cells do, is the consequence of these cells rupturing during the dissociation process. Once ruptured, large numbers of neuropeptide transcripts could be present in the ambient solution, leading to background levels of neuropeptide transcripts being 'picked up' with non-expressing cells. Our data for the neuropeptide *Orcokinin* illustrates this point (***Figure 8—figure supplement 3***). Orcokinin is expressed in just 5 neurons in the adult VNC, two pairs of neurons in thoracic neuromeres, with additional expression in one unpaired neuron in the abdominal neuromere (***Chen et al., 2015***). In our data, 5 cells in the NP cluster show *Orcokinin* expression at a level of more than 100 transcripts (***Figure 8—figure supplement 3***). However, almost 8000 cells, across all clusters, also appear to express *Orcokinin* at a level of just 1–10 transcripts (***Figure 8—figure supplement 3***). We speculate that the low-level expression outside the NP cluster reflects background levels, due to rupture of *Orcokinin* expressing cells, while the expression above 100 transcripts within the NP cluster is *bona fide*.

Non-neuronal cell types: Glia

Glia are key regulators of nervous system physiology maintaining the concentration of chemicals in the extracellular environment. We identified glia cells in our VNC data using the established glial markers *reversed polarity* (*repo*) (***Xiong et al., 1994***) and the long non-coding RNA *MRE16* (***Davie et al., 2018***; ***Figure 1—figure supplement 7***). Five clusters, representing 3.6% of the total cells, were highly enriched for *repo* and *MRE16*. In addition, these clusters lacked expression of neuronal markers, such as *elav*, *nSyb*, and *noe* (***Figure 1—figure supplement 7***). Two of these clusters contained cells that were positive for both glial markers and cells positive for neuronal markers, suggesting mixed populations. To define specific glial cell types, we performed a sub-clustering analysis of the five glial clusters (***Figure 9A***), while removing any cell with a neuronal signature (see methods). This sub-clustering revealed four distinct glial subpopulations that could be classified based on enriched expression of established markers (***Figure 9B***): *astrocytic leucine-rich repeat molecule* (*alrm*) for astrocytes (***Doherty et al., 2009***), *I'm not dead yet* (*Indy*) for surface glia (***Knauf et al., 2002***), *wrapper* for cortex glia (***Noordermeer et al., 1998***), and *Excitatory amino acid transporter 2* (*Eaat2*) for ensheathing glia (***Stahl et al., 2018***).

Surface glia

Consistent with the established metabolic role of glia in the brain, many solute carrier (SLC) membrane transporters are enriched in glial cell types (***Figure 9D***; ***Figure 9—source data 1***). Surface

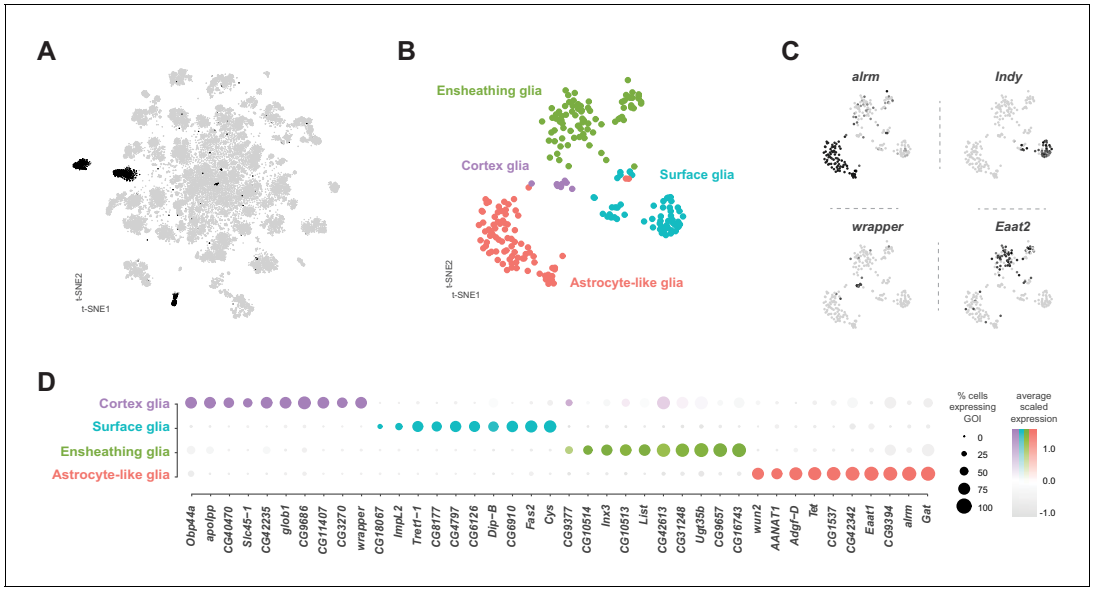


Figure 9. Identification and characterization of glial cell sub-types. (**A**) t-SNE plot highlighting glial cell clusters. Clusters 23, 24, 70, 98, and 106 are shown in black, as they were determined to be glia based on differential expression of established glial and neuronal biomarkers (***Figure 1—figure supplement 7***). (**B**) t-SNE plot showing sub-clustering analysis of glial cells. Distinct clusters are color-coded and named. (**C**) Expression of established glial subtype-specific biomarkers used to determine cluster identity. *astrocytic leucine-rich repeat molecule* (*alrm*), for astrocytes; *I'm not dead yet* (*Indy*) for surface glia; *wrapper* for cortex glia; *Excitatory amino acid transporter 2* (*Eaat2*) for ensheathing glia. Expression shown in black, intensity is proportional to the log-normalized expression levels. (**D**) Dot plot of the top 10 genes in each glial sub-type based on fold-enrichment (***Figure 9—source data 1***). Size of dots represent percent of cells in cluster expressing gene of interest (GOI); intensity of color reflects average scaled expression. ***Figure 9—source data 1*** contain the full list of significantly enriched genes.

The online version of this article includes the following source data and figure supplement(s) for figure 9:

Source data 1. List of marker genes for the glial sub-clusters shown in ***Figure 9D***.

Figure supplement 1. Identification of neuronal and glia clusters.

glia, which form the blood-brain barrier (BBB), metabolically insulate the nervous system from the hemolymph. The BBB also subserves the considerable energetic demands of the nervous system (***Laughlin et al., 1998***), by efficiently transporting sugars, ions, and metabolites between the hemolymph and brain (***Limmer et al., 2014***). Surface glia strongly express the previously described surface glia marker SLC2 *Trehalose transporter 1–1* (*Tret1-1*), which as the name suggests transports trehalose, the main carbohydrate in the hemolymph, into the nervous system (***Volkenhoff et al., 2015***). Surface glia also contain high levels of a putative SLC2 sugar transporter *CG4797*, previously reported to be expressed in the perineural glia of the optic lobe (***Figure 9D***; ***Konstantinides et al., 2018***). Both of these sugar transporters are proton-dependent and therefore, efficient sugar transport requires glial cells to have lower H^+ concentration than the hemolymph (***Kikuta et al., 2012***). We found enriched expression of multiple V-type ATPase H^+ pumping genes (p-value=3.9×10^{-9} based on DAVID analysis; ***Figure 9—source data 1***; ***Chintapalli et al., 2013***) in surface glia. Expression of the SLC bicarbonate transporter *CG8177*, which has been shown to reduce extracellular pH (***Overend et al., 2016***), and the *Ecdysone-inducible gene L2* (*ImpL2*) were also enriched in surface glia (***Figure 9D***). *ImpL2* antagonizes *insulin-like peptide 2* (*Dilp2*) and inhibits insulin/insulin-like growth factor signaling (***Honegger et al., 2008***) suggesting that the BBB is responsive to the metabolic demands of the fly.

Astrocytes and ensheathing glia

Over 70% of the VNC astrocyte cell-specific markers are shared with those in astrocytes in the midbrain (***Croset et al., 2018***). This similarity suggests that there is minimal regional specialization of astrocyte identity and function within the CNS. All astrocytes and ensheathing glia in the thoracic VNC are derived from the same lineages as leg motor neurons (***Enriquez et al., 2018***). However, we did not find evidence of a lineage-specific transcriptional code in these glial cell types (***Figure 9—***

source data 1). Astrocytes instead, strongly express the paired-like homeobox TF *CG34367*, previously reported to be expressed in astrocytes throughout development (**Huang et al., 2015**). In contrast, ensheathing glia express the POU homeobox TF *ventral veins lacking* (*vvl*) (**Anderson et al., 1995**). These TFs can, therefore, be considered as novel markers for these adult VNC cell types. The rarity of TF markers is in accord with the finding that the morphological specificity of motor neurons depends on a unique TF code, whereas astrocytes and ensheathing glia show plasticity in their morphologies dependent on their position, rather than a distinct TF code (**Enriquez et al., 2018**).

Non-neural 'contamination'

We identified two clusters of cells that represent contamination during the dissection process (**Figure 1—figure supplement 7B**). Cluster 99 was determined to be salivary gland tissue, since the cluster markers for cluster 99 contains 8 of the top 10 genes enriched in salivary gland tissue in the FlyAtlas 2 dataset (**Figure 1—source data 1**; **Leader et al., 2018**). Cluster 116 appears to be sperm as many of the markers for this cluster are unique to the testis and sperm (**Figure 1—source data 1**; **Witt et al., 2019**).

Discussion

As the field tries to categorize nervous systems at higher resolution, the question arises what precisely constitutes a cell type? The goal itself seems straightforward in principle, to find a way to define different groups of cells that carry out distinct tasks. Single-cell mRNA sequencing techniques provide a possible route to answering this question by allowing the neuronal transcriptome of thousands of individual cells within a complex nervous system to be collected in parallel. To this end, we generated a single-cell transcriptional atlas that reveals extensive cellular diversity in the adult *Drosophila* ventral nerve cord. In combination with previous single-cell data from the antennal lobe, optic lobe, and brain (**Croset et al., 2018**; **Davie et al., 2018**; **Konstantinides et al., 2018**; **Li et al., 2017**), our data contribute towards a comprehensive cell atlas representative of the entire adult central nervous system.

Distinguishing between some cell types such as neurons and glia is relatively straightforward, but the extent of neuronal diversity provides a real challenge. Different neuronal types transmit particular neurotransmitters, neuropeptides and monoamines, and they also respond to a variety of these signals using their complement of cell-surface receptors. Specific neurons might also express unique ion-channels and cell-signaling cascades that provide the cell with a range of electrophysiological characteristics, and potential mechanisms of plasticity. In principle, we can also define neuronal cell types by characterizing their neuroanatomy - where the neurons are located and to which neurons they are pre- and post-synaptically connected. Since neurons acquire their anatomy through developmental programs, it was not known whether this information would remain accessible in any form in the snapshot of the transcriptome of mature fully differentiated adult VNC. However, one of the most evident cluster-defining features of our VNC data is the abundance of transcription factors and cell-adhesion molecules that are classically thought of as being developmental.

The cluster-defining TFs are particularly useful for annotating the cell types in the VNC. Decades of work has investigated the development and structural organization of the VNC and has described roles for these genes in developmental specification of hemilineages (**Baek and Mann, 2009**; **Birkholz et al., 2015**; **Bossing et al., 1996**; **Lacin et al., 2019**; **Lacin and Truman, 2016**; **Prokop and Technau, 1991**; **Schmid et al., 1999**; **Shepherd et al., 2016**; **Shepherd et al., 2019**; **Truman and Bate, 1988**; **Truman et al., 2004**). Hemilineages have been proposed to be the functional units of the VNC, representing fundamental organizational principles for connectivity (**Harris et al., 2015**). Our data are entirely consistent with the hemilineages as functional units with each hemilineage made up of a population of neurons that share morphological, transcriptional, and neurochemical features. They represent a familial unit and our data clearly in an unbiased way, pulls them out as separate and identifiable genetic units, reinforcing the idea that the hemilineages are functional groups that share molecular/genetic identity as well as morphology and function. Moreover, we can see that hemilineages are not all homogenous in their composition; there are distinct subtypes evident from the single-cell data. For example, we documented that *kn* and *twz* are expressed in distinct subsets of hemilineage 23B. The fact that subtypes are seen in our data, as well

as the number of subtypes for a given hemilineage, is consistent with what others have observed anatomically (D. Shepherd, pers. comm).

We also found continued expression in hemilineages of neurodevelopment genes, such as *br*, *Imp*, and *pros*, which may signify cell birth order and therefore further distinguish cell types. Expression levels of *pros* indicate the birth order of post-embryonic leg motor neurons belonging to hemilineage 15B, that precedes muscle-specific innervation (**Baek et al., 2013**; **Baek and Mann, 2009**). Relative levels of *Imp* and *pros* in our predicted hemilineage 15B, can therefore be used to infer muscle-specific innervation along the leg. *Imp* and *pros* expression may also distinguish between embryonic born and post-embryonic born neurons. Mature (embryonic) neurons of the larval VNC show elevated *Imp* expression, whereas immature (post-embryonic) neurons have higher levels of *pros*, relative to each other (**Etheredge, 2017**). Dopaminergic neurons in the PAM1 cluster in the larva are the only dopaminergic neurons to be born post-embryonically (**Hartenstein et al., 2017**). So, in some instances, embryonic vs. post-embryonic identity can also be used to identify neurons within a given class.

Fast-acting neurotransmitter identity is acquired in a hemilineage dependent manner, amongst post-embryonic neurons (**Lacin et al., 2019**), but the underlying transcriptional programs appear to be complex (**Estacio-Gómez et al., 2019**). We did not identify single TFs that globally defined FAN identity. Different neurons have been reported to use different TFs to define the same FAN identity in the adult fly optic lobes (**Konstantinides et al., 2018**) and in *C. elegans* (**Hobert and Kratsios, 2019**). In addition, a recent gene-profiling study of cholinergic, glutamatergic, and GABAergic neurons across development found that many FAN-specific TFs are transiently expressed at particular times, and only a few remain constant (**Estacio-Gómez et al., 2019**). As noted previously (**Lacin et al., 2019**), we found some TFs that are restricted to individual FAN-specific clusters. Some GABAergic clusters express *Dbx*, *vg*, *D*; some cholinergic neurons express *unc-4* and some glutamatergic neurons express *ems*. We also found *Lim3* to be expressed in cholinergic and glutamatergic neurons, but not in GABAergic neurons (**Lacin et al., 2019**). This is in contrast to what is seen in the optic lobes, where *Lim3* specifically regulates GABAergic cellular identity (**Konstantinides et al., 2018**). Yet, consistent with the optic lobes, *Lim3* is only expressed in Notch^{off} hemilineages in the VNC (**Li et al., 2017**). Therefore, the TF code specifying FAN identity may change across development and also differ depending on the regional context.

Hox gene expression (*Antp*, *Ubx*, *abd-A*, and *Abd-B)* developmentally define the neuromeres of the VNC. Finding that established neuromere- and hemilineage-enriched developmental markers are expressed in the 5 day old adult, suggests they may continue to act in these differentiated cells to maintain the distinct transcriptional profile underlying neuronal identity and function (**Deneris and Hobert, 2014**). It will therefore be interesting to investigate functional consequences of disrupting these potential maintenance programs.

Not all cells group by their developmental lineage, for example, monoaminergic cells formed a distinct cluster, despite originating from multiple different neuroblasts (**Schmid et al., 1999**). We also identified a cluster of neurosecretory cells that includes cells with different developmental origins (**Park et al., 2008**). High expression levels of neuropeptides in this cluster is consistent with features of neurosecretory biology. Neuropeptides are often produced by small numbers of cells, yet they can be broadly released and by volume transmission and/or secretion into the hemolymph act at a distance to modulate disparate neural circuits, physiological functions and behaviors (**Nässel, 2018**). The extraordinarily high-level expression of pre-pro-neuropeptide genes in these cells suggests that producing these molecules presents a considerable burden. The TF *dimmed* (*dimm*), which marks neurosecretory cells of this type, has been shown to promote neurosecretory identity and suppress FAN identity (**Hewes et al., 2003**; **Park et al., 2008**).

In this study, we have demonstrated the predictive power of the single-cell atlas of the VNC. Despite the caveats associated with this technique, such as sparse sampling of a cell's transcriptome and ambient RNA contamination, cellular identity signals, especially developmental programs, are surprisingly robust. Defining cell types based on clustering should be viewed as an exploratory rather than a confirmatory process (**Crow and Gillis, 2019**). It is worth noting that the genes identified using this technique that most robustly define cell clusters do not necessarily reflect their importance for cell type function, however many will no doubt be useful for the generation of tools to study and ultimately define cell phenotypes. Our results suggest many new directions for further investigation. For example, understanding the correspondence and potential causal relationships

between transcriptomic signatures and specific anatomical, physiological and functional properties, and how these relationships change depending on cell state.

Materials and methods

Fly strains and husbandry

The fly strain used for VNC analysis (+/w*; *UAS-Stinger*, *13XLexAop2-IVS-tdTomato.nls*/+; dsx^{GAL4}/+) was a genetic cross between *w*; UAS-Stinger*, *13XLexAop2-IVS-tdTomato.nls* males and +; +; dsx^{GAL4} virgin females (**Rideout et al., 2010**). All flies were reared at 25˚C in a 12:12 hr light: dark cycle on standard food at 40–50% relative humidity. Virgin males and females were collected and stored individually. Flies were aged 5 days post-eclosion at 25˚C prior to dissection. Additional fly strains used in this study include: (1) wild-type *Canton-S* (2) y^1w*; Mi{Trojan-GAL4.0}Vmat [MI07680-TG4.0]* (BDSC:66806) (3) y^1w*; Mi{Trojan-GAL4.1}Lim3[MI03817-TG4.1]/SM6a* (BDSC:67450) (4) y^1w*; {Mi{Trojan-GAL4.2}kn[MI15480-TG4.2]/SM6a* (BDSC:67516) (5) y^1w*; Mi{Trojan-GAL4.2}twz[MI14153-TG4.2]/SM6a* (BDSC:76758) (6) y^1w* Mi{Trojan-GAL4.0}acj6[MI07818-TG4.0]/FM7c* (BDSC:77788) (7) *w*; {UAS-Stinger}* denoted as *UAS-nGFP* (**Barolo et al., 2000**) (8) *w*; UAS-myr-GFP-V5-P2A-H2B-mCherry-HA/TM3, Ser* (aka *UAS-WM*; **Chang et al., 2019**).

VNC single-cell sample preparation

The VNC dissociation protocol was carried out as described previously (**Croset et al., 2018**). 40 Male and 40 female VNCs (20 per sexed replicate) were individually dissected in toxin-supplemented ice-cold calcium- and magnesium-free DPBS (Gibco, 14190–086 + 50 µM D(−)−2-Amino-5-phosphonovaleric acid, 20 µM 6,7-dinitroquinoxaline-2,3-dione and 0.1 µM tetrodotoxin). Each replicate was then washed in 1 mL ice-cold toxin-supplemented Schneider's medium (tSM: Gibco, 21720–001 + toxins, as above). VNCs were then incubated for 30 min in 0.5 mL of tSM containing 1 mg/mL papain (Sigma, P4762) and 1 mg/mL collagenase I (Sigma, C2674). VNCs were washed once more with tSM and subsequently triturated with flame-rounded 200 µL pipette tips. Dissociated VNCs were resuspended into 1 mL PBS + 0.01% BSA and filtered through a 10 µm CellTrix strainer (Sysmex, 04-0042-2314).

Data processing

Libraries were made using the Chromium Single Cell 3' v2 kit from 10x Genomics. Cells were loaded in accordance with 10x Genomics documentation, with the aim of 5000–8000 cells per sample. The samples were sequenced with 8 lanes of Illumina HiSeq4000 by Oxford Genomics Centre. We obtained a mean of 20,550 reads per cell and mean sequence saturation of 71%. The fastq data were processed with Drop-seq_tools v2.1.0 (**Macosko et al., 2015**) and aligned to the *D. melanogaster* genome (R6.13) to generate digital expression matrices for each sample.

Data analysis with seurat

The digital expression matrices were analyzed with the Seurat 2.3.4 R package (**Satija et al., 2015**). Genes expressed in fewer than 3 GEMs (Gel Bead-In EMulsions) were removed. GEMs with more than 10,000 UMI (1.4% of GEMs, 3.6 standard deviations away from the median) were removed as outliers (unless otherwise mentioned). GEMs with fewer than 1,200 UMI were removed. The lower limit was determined using a histogram of the number of UMI per GEM. This distribution has a local minimum between 1000 and 1,200 UMI. GEMs with fewer than 200 genes, or more than 15% mitochondrial derived UMI were removed. This resulted in 26,768 cells in total and 2590, 9060, 6522, 8596 cells for the four replicates, respectively. Sex-specific replicates were merged with the '*MergeSeurat*' function and normalized with the '*NormalizeData*' function, using default parameters. The merged sex-specific groups were then scaled with the '*ScaleData*' function while regressing out variation due to replicate, nUMI, and proportion mitochondrial transcript. Low expression level, and low dispersion cut-offs of 0.001 were used to identify variable expressed genes in each sex. The intersection of these genes was used to perform a canonical correlation analysis (CCA) with the first 45 dimensions. t-distributed stochastic neighbor embedding (t-SNE) was performed, with perplexity of 30, theta of 0.05, and 20,000 iterations, on the data to reduce the dimensionality to 2 for visualization. Clusters were defined using the '*FindClusters*' function with the default Louvain algorithm and

using 45 dimensions and a cluster resolution of 12 (unless otherwise mentioned). Comparison of different cluster resolutions was evaluated with the 'clustree' package (**Zappia and Oshlack, 2018**). Cluster resolutions above 12 yielded few new clusters and increased the between cluster exchanges. Positively enriched cluster markers were identified using the *'FindAllMarkers'* function with a negative binomial distribution test and fold enrichment of at least 0.5 and a Bonferroni adjusted p-value of less than 0.05. Neuronal identity was determined by having an average scaled expression (calculated with the *'AverageExpression'* function in *'Seurat'*) of at least 0.15 for any of the following neuronal markers (*elav*, *nSyb*, *para*, *VAChT*, *ChAT*, *Gad1*, *VGAT*, *VGlut*, *noe*). All other clusters were deemed to be non-neuronal, with the exception of neuropeptide-expressing cluster 84. Glial identity was determined in a similar manner with the following glial markers (*repo*, *alrm*, *wrapper*, *Indy*). Functionally related gene enrichment analysis on neuronal cluster markers was performed using DAVID (**Huang et al., 2009**).

Comparisons to bulk sequence data and between replicates

FlyAtlas 2 bulk RNA-seq data of the adult VNC was obtained from http://flyatlas.gla.ac.uk/FlyAtlas2/index.html (**Leader et al., 2018**). The convert IDs tool from Flybase was used to convert gene symbols and Flybase ID between release versions (http://flybase.org/convert/id). Pseudo-bulk normalized expression from the filtered single-cell VNC data was then compared to the pooled female and male FPKM from the FlyAtlas 2 data set. Pearson correlation coefficients were calculated with the *'stat_cor'* function from the *'ggpubr'* package. Pseudo-bulk normalized expression from each replicate were compared to each other, and correlations were calculated with *'stat_cor'*. Alignment of the replicates was assessed with the *'CalcAlignmentMetric'* function in *'Seurat'*. The *'AverageExpression'* function in *'Seurat'* along with the *'correlate'* function in *'corrr'* were used to determine the correlations of gene expression between the replicates at the cluster level.

Comparison with brain Single-Cell data

Loom files for the mid-brain (**Croset et al., 2018**) and the whole (**Davie et al., 2018**) were downloaded from 'scope.aertslab.org'. Digital expression matrices of filtered cells were extracted from the loom files and reanalyzed with similar parameters to those used for the VNC. The data were normalized with the *'NormalizeData'* function with default parameters. The data were scaled with the *'ScaleData'* function while regressing out variation due to nUMI and proportion mitochondrial expression. Variably expressed genes were identified with the *'FindVariableGenes'* function with the following cutoffs: x.low.cutoff = 0.001, x.high.cutoff = Inf, y.cutoff = 0.001. A principal component analysis was performed using the identified variable genes. Clusters were identified with the *'FindClusters'* function using the default Louvain algorithm. A cluster resolution of 2.5 and the first 50 PCAs were used for the mid-brain data (**Croset et al., 2018**), and a cluster resolution of 2 and the first 82 PCAs were used for the brain data (**Davie et al., 2018**). t-SNE was performed with perplexity of 30, theta of 0.1, and 20,000 iterations, on the data to reduce the dimensionality to two for visualization. Significantly enriched genes for each cluster were identified using the *'FindAllMarkers'* function using the negative binomial test with a log fold-change threshold of 0.5 while repressing out variation due to replicate, and a Bonferroni adjusted p-value of less than 0.05. The resulting cluster markers for the mid-brain and brain data sets were compared to the VNC data cluster markers (excluding the salivary gland cluster and the sperm cluster). The overlap of markers was visualized with an Euler plot using the 'eulerr' package in R.

Predicting hemilineage identity

The average expression of previously established hemilineage markers (**Lacin et al., 2019**; **Venkatasubramanian and Mann, 2019**) across all neuronal clusters was calculated with the *'AverageExpression'* function in *'Seurat'*. Predicted hemilineage identities were deduced from these expression patterns. With these assignments, we used the *'FindAllMarkers'* function (with the above settings) to identify novel biomarkers for each predicted hemilineage. Hemilineage specific cell counts (inferred from **Birkholz et al., 2015** and **Lacin et al., 2019**) were compared to the cell counts from our predicted labeled hemilineages. To approximate post-embryonic thoracic cells, only *pros*$^{+}$, *abd-A*$^{-}$, *Abd-B*$^{-}$ cells were considered.

Assigning fast-acting neurotransmitter identity

The following genes were used to assign fast-acting neurotransmitter (FAN) identity; *Vesicular acetylcholine transporter* (*VAChT*) and *Choline acetyltransferase* (*ChAT*) for acetylcholine, *Vesicular glutamate transporter* (*VGlut*) for glutamate, and *Vesicular GABA Transporter* (*VGAT*) and *Glutamic acid decarboxylase 1* (*Gad1*) for GABA. Cells from clusters with average scaled expression above 0.5 for *Gad1* or *VGAT*, and average scaled expression less than 0 for the other markers were all assigned to be GABAergic. Glutamatergic and cholinergic cells were assigned similarly. For clusters with an average expression less than 0.5 for all markers, cells were assigned FAN identity at a cell-by-cell level. For these remaining cells, we established different criteria for assigning cholinergic vs. glutamatergic or GABAergic since the cholinergic markers were expressed at much lower levels. If both *VAChT* and *ChAT* had a log-normalized expression greater than 0, then the cells were assigned to be cholinergic. Additionally, if either *VAChT* or *ChAT* had a log-normalized expression greater than 2, the cells were assigned to be cholinergic. Otherwise, if *VGlut* had a log-normalized expression greater than 2 and greater than either *VAChT* or *Gad1*, then the cells were assigned to be glutamatergic. Otherwise, if *Gad1* had a log-normalized expression greater than 2 and greater than either *VAChT* or *VGlut*, the cells were assigned to be GABAergic. All remaining cells were labelled undefined.

Sub-clustering of monoaminergic cells

Vmat expressing cells from clusters 72 and 84 were sub-clustered. Identification of variable genes and canonical correlation analysis were performed as above. t-SNE was performed using the first 7 dimensions and a cluster resolution of 1.2 was used to identify clusters. Positively enriched cluster markers were identified as above.

Sub-clustering of glial cells

Clusters 23, 24, 70, 98, and 106 were identified as being predominantly glial cells. Cells from these clusters were sub-clustered. Any cell expressing *elav*, *nSyb*, *VAChT*, *VGlut*, or *Gad1* was removed. Identification of variable genes and canonical correlation analysis were performed as above. t-SNE was performed using the first 6 dimensions and a cluster resolution of 0.9 was used to identify clusters. Positively enriched cluster markers were identified as above.

R packages and plotting

Expression heatmaps were drawn using the 'pheatmap' package, correlation heatmaps were drawn using the 'ggcorplot' package. Chord diagrams were drawn using the 'circlize' package (***Gu et al., 2014***), and represent the relationship between cluster markers and the clusters in which they are significantly enriched, the data is taken from ***Figure 1—source data 1***. No weighted relationships are inferred in chord diagrams. Scatter plots and bar charts were drawn using ggplot2 in R. A full list of packages used, with their version numbers can be found at https://github.com/aaron-allen/VNC_scRNAseq/blob/master/sessionInfo.txt. All plots were edited in Adobe Illustrator.

Phylogeny

The multiple sequence alignment (not shown) and the phylogenetic tree were created with Clustal Omega (***Sievers and Higgins, 2018***) and FigTreeV1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/), respectively.

Immunohistochemistry

Flies were reared at 25˚C and aged for 5 days prior to dissection and staining as per ***Rideout et al. (2010)*** with the following modifications: VNC samples were fixed in 4% PFA (40 mins at room temperature) immediately following dissection, to maintain tissue integrity and minimize cell loss. Samples were pre-incubated in 5% NGS overnight at 4˚C. Samples were then incubated with primary antibodies for 3 days at 4˚C, followed by an overnight incubation in secondary antibodies at 4˚C. Primary antibodies used were: mouse mAb nC82 (1:50), mouse anti-AbdA (1:100, C-11), mouse anti-AbdB (1:10, 1A2E9-s), mouse anti-Acj6-s (1:20), mouse anti-AntP (1:20, 8C11-s), mouse anti-Ubx (1:20, FP3.38-s), rat anti-CadN (1:30, DN-Ex #8;) from DSHB, Univ. of Iowa. Additional primary antibodies used were chicken anti-GFP (1:1200, Abcam) and anti-AbdA (1:100, C-11 Santa Cruz

Biotechnology). Secondary antibodies used included: anti-chicken Alexa Fluor488, anti-mouse Alexa Fluor633, anti-rat Alexa Fluor546, (1:300, Invitrogen Molecular Probes, Carlsbad, CA). Samples were left in 70% Glycerol/30% PBT overnight at 4˚C prior to mounting with Vectashield (Vector Labs) and imaged with a Leica SP5 Microscope. Stacks of optical sections were generated at 0.5 µm intervals. Images were processed in Imaris 8.2.1 (Bitplane Scientific, AG, Zürich).

Data and code availability

Raw sequencing files (fastq) and digital expression matrices from each replicate are available from the Gene Expression Omnibus (GSE141807). Code used in this analysis is available from GitHub (https://github.com/aaron-allen/VNC_scRNAseq; ***Allen, 2020***; copy archived at https://github.com/elifesciences-publications/VNC_scRNAseq).

Acknowledgements

We thank Tetsuya Nojima and Annika Rings for dissection assistance. We thank David Shepherd, Devika Agarwal, Deniz Erezyilmaz, Julian Dow, and members of the Goodwin lab for helpful discussions and critical reading of the manuscript. We thank the Bloomington Stock Center for flies. We thank Josh Dubnau for the *UAS-WM* line. Sequencing was carried out by The Oxford Genomics Centre. We thank Deborah J Andrew and Caitlin D Hanlon for assistance generating a Figure of the GPCRs, and Andreas Prokop for the adult *Drosophila* illustration. SW is supported by a Wellcome Principal Research Fellowship (200846/Z/16/Z) and ERC Advanced Grant (789274). This work was funded by a Wellcome Investigator Award (106189/Z/14/Z) to SFG, and a Wellcome Collaborative Award (209235/Z/17/Z) to SFG and SW.

Additional information

Funding

Funder	Grant reference number	Author
Wellcome Trust	Wellcome Trust Investigator Award (106189/Z/14/Z)	Aaron M Allen Megan C Neville Stephen F Goodwin
ERC Advanced Grant	789274	Vincent Croset Christoph Daniel Treiber Scott Waddell
Wellcome Trust	Wellcome Principal Research fellowship (200846/Z/16/Z)	Scott Waddell
Wellcome Trust	Wellcome Collaborative Award (209235/Z/17/Z)	Scott Waddell Stephen F Goodwin
MRC	Graduate Student Fellowship	Sebastian Birtles

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Aaron M Allen, Conceptualization, Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing - original draft, Writing - review and editing; Megan C Neville, Conceptualization, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing - original draft, Writing - review and editing; Sebastian Birtles, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing - original draft, Writing - review and editing; Vincent Croset, Methodology, Writing - review and editing; Christoph Daniel Treiber, Methodology; Scott Waddell, Conceptualization, Supervision, Funding acquisition, Writing - review and editing; Stephen F Goodwin, Conceptualization, Supervision, Funding acquisition, Writing - original draft, Project administration, Writing - review and editing

Author ORCIDs

Aaron M Allen https://orcid.org/0000-0002-7961-4392
Megan C Neville https://orcid.org/0000-0001-8506-9944
Sebastian Birtles https://orcid.org/0000-0003-4617-4092
Christoph Daniel Treiber http://orcid.org/0000-0002-6994-091X
Scott Waddell http://orcid.org/0000-0003-4503-6229
Stephen F Goodwin https://orcid.org/0000-0002-0552-4140

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.54074.sa1
Author response https://doi.org/10.7554/eLife.54074.sa2

Additional files

Supplementary files

- Transparent reporting form

Data availability

Sequencing data have been deposited in GEO under accession code GSE141807.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Allen AM, Neville MC, Birtles S, Croset V, Treiber C, Waddell S, Goodwin SF	2019	A single-cell transcriptomic atlas of the adult Drosophila ventral nerve cord	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE141807	NCBI Gene Expression Omnibus, GSE141807

The following previously published datasets were used:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Leader DP, Krause SA, Pandit A, Davies SA, Dow JAT	2018	FlyAtlas 2: a new version of the Drosophila melanogaster expression atlas with RNA-Seq, miRNA-Seq and sex-specific data	https://www.omicsdi.org/dataset/omics_ena_project/PRJEB22205	Omnics DI, PRJEB22205

References

Agrawal T, Sadaf S, Hasan G. 2013. A genetic RNAi screen for IP_3/Ca^{2+} coupled GPCRs in *Drosophila* identifies the PdfR as a regulator of insect flight. *PLOS Genet* **9**:e1003849. DOI: https://doi.org/10.1371/journal.pgen.1003849, PMID: 24098151

Allen AM. 2020. Analysis of scRNA-seq of the *Drosophila melanogaster* adult ventral nerve cord from females and males. 569bc02. GitHub. https://github.com/aaron-allen/VNC_scRNAseq

Anderson MG, Perkins GL, Chittick P, Shrigley RJ, Johnson WA. 1995. *Drifter*, a *Drosophila* POU-domain transcription factor, is required for correct differentiation and migration of tracheal cells and midline Glia. *Genes & Development* **9**:123–137. DOI: https://doi.org/10.1101/gad.9.1.123, PMID: 7828848

Angelini DR, Kaufman TC. 2005. Insect appendages and comparative ontogenetics. *Developmental Biology* **286**:57–77. DOI: https://doi.org/10.1016/j.ydbio.2005.07.006, PMID: 16112665

Baek M, Enriquez J, Mann RS. 2013. Dual role for hox genes and hox co-factors in conferring leg motoneuron survival and identity in *Drosophila*. *Development* **140**:2027–2038. DOI: https://doi.org/10.1242/dev.090902, PMID: 23536569

Baek M, Mann RS. 2009. Lineage and birth date specify motor neuron targeting and dendritic architecture in adult *Drosophila*. *Journal of Neuroscience* **29**:6904–6916. DOI: https://doi.org/10.1523/JNEUROSCI.1585-09.2009

Barolo S, Carver LA, Posakony JW. 2000. GFP and beta-galactosidase transformation vectors for promoter/enhancer analysis in *Drosophila*. *BioTechniques* **29728**:726–732. DOI: https://doi.org/10.2144/00294bm10

Bates AS, Janssens J, Jefferis GS, Aerts S. 2019. Neuronal cell types in the fly: single-cell anatomy meets single-cell genomics. *Current Opinion in Neurobiology* **56**:125–134. DOI: https://doi.org/10.1016/j.conb.2018.12.012, PMID: 30703584

Bidaye SS, Machacek C, Wu Y, Dickson BJ. 2014. Neuronal control of *Drosophila* walking direction. *Science* **344**: 97–101. DOI: https://doi.org/10.1126/science.1249964, PMID: 24700860

Billeter JC, Villella A, Allendorfer JB, Dornan AJ, Richardson M, Gailey DA, Goodwin SF. 2006. Isoform-specific control of male neuronal differentiation and behavior in *Drosophila* by the *fruitless* gene. *Current Biology* **16**: 1063–1076. DOI: https://doi.org/10.1016/j.cub.2006.04.039, PMID: 16753560

Birkholz O, Rickert C, Nowak J, Coban IC, Technau GM. 2015. Bridging the gap between postembryonic cell lineages and identified embryonic neuroblasts in the ventral nerve cord of *Drosophila melanogaster*. *Biology Open* **4**:420–434. DOI: https://doi.org/10.1242/bio.201411072, PMID: 25819843

Bossing T, Udolph G, Doe CQ, Technau GM. 1996. The embryonic central nervous system lineages of *Drosophila melanogaster* I Neuroblast lineages derived from the ventral half of the neuroectoderm. *Developmental Biology* **179**:41–64. DOI: https://doi.org/10.1006/dbio.1996.0240, PMID: 8873753

Buchner E, Buchner S, Burg MG, Hofbauer A, Pak WL, Pollack I. 1993. Histamine is a major mechanosensory neurotransmitter candidate in *Drosophila melanogaster*. *Cell and Tissue Research* **273**:119–125. DOI: https://doi.org/10.1007/BF00304618, PMID: 8103421

Bürglin TR, Affolter M. 2016. Homeodomain proteins: an update. *Chromosoma* **125**:497–521. DOI: https://doi.org/10.1007/s00412-015-0543-8, PMID: 26464018

Burrows M. 1992. Local circuits for the control of leg movements in an insect. *Trends in Neurosciences* **15**:226–232. DOI: https://doi.org/10.1016/0166-2236(92)90040-F, PMID: 1378667

Card G, Dickinson MH. 2008. Visually mediated motor planning in the escape response of *Drosophila*. *Current Biology* **18**:1300–1307. DOI: https://doi.org/10.1016/j.cub.2008.07.094, PMID: 18760606

Carrillo RA, Özkan E, Menon KP, Nagarkar-Jaiswal S, Lee PT, Jeon M, Birnbaum ME, Bellen HJ, Garcia KC, Zinn K. 2015. Control of synaptic connectivity by a network of *Drosophila* IgSF Cell Surface Proteins. *Cell* **163**:1770–1782. DOI: https://doi.org/10.1016/j.cell.2015.11.022, PMID: 26687361

Cavodeassi F, Modolell J, Gómez-Skarmeta JL. 2001. The iroquois family of genes: from body building to neural patterning. *Development* **128**:2847–2855. PMID: 11532909

Chang YH, Keegan RM, Prazak L, Dubnau J. 2019. Cellular labeling of endogenous retrovirus replication (CLEVR) reveals de novo insertions of the gypsy retrotransposable element in cell culture and in both neurons and glial cells of aging fruit flies. *PLOS Biology* **17**:e3000278. DOI: https://doi.org/10.1371/journal.pbio.3000278, PMID: 31095565

Chen J, Choi MS, Mizoguchi A, Veenstra JA, Kang K, Kim YJ, Kwon JY. 2015. Isoform-specific expression of the neuropeptide orcokinin in *Drosophila melanogaster*. *Peptides* **68**:50–57. DOI: https://doi.org/10.1016/j.peptides.2015.01.002, PMID: 25600008

Chen CL, Hermans L, Viswanathan MC, Fortun D, Aymanns F, Unser M, Cammarato A, Dickinson MH, Ramdya P. 2018. Imaging neural activity in the ventral nerve cord of behaving adult *Drosophila*. *Nature Communications* **9**: 4390. DOI: https://doi.org/10.1038/s41467-018-06857-z, PMID: 30348941

Chintapalli VR, Wang J, Herzyk P, Davies SA, Dow JA. 2013. Data-mining the FlyAtlas online resource to identify core functional motifs across transporting epithelia. *BMC Genomics* **14**:518. DOI: https://doi.org/10.1186/1471-2164-14-518, PMID: 23895496

Clyne PJ, Certel SJ, de Bruyne M, Zaslavsky L, Johnson WA, Carlson JR. 1999. The odor specificities of a subset of olfactory receptor neurons are governed by Acj6, a POU-domain transcription factor. *Neuron* **22**:339–347. DOI: https://doi.org/10.1016/S0896-6273(00)81094-6, PMID: 10069339

Clyne JD, Miesenböck G. 2008. Sex-specific control and tuning of the pattern generator for courtship song in *Drosophila*. *Cell* **133**:354–363. DOI: https://doi.org/10.1016/j.cell.2008.01.050, PMID: 18423205

Cognigni P, Felsenberg J, Waddell S. 2018. Do the right thing: neural network mechanisms of memory formation, expression and update in *Drosophila*. *Current Opinion in Neurobiology* **49**:51–58. DOI: https://doi.org/10.1016/j.conb.2017.12.002, PMID: 29258011

Cole SH, Carney GE, McClung CA, Willard SS, Taylor BJ, Hirsh J. 2005. Two functional but noncomplementing *Drosophila* tyrosine decarboxylase genes: distinct roles for neural tyramine and octopamine in female fertility. *The Journal of Biological Chemistry* **280**:14948–14955. DOI: https://doi.org/10.1074/jbc.M414197200, PMID: 15691831

Court RC, Armstrong J, Börner J, Card G, Costa M, Dickinson M, Duch C, Korff W, Mann R, Merritt D. 2017. A systematic nomenclature for the *Drosophila* ventral nervous system. *bioRxiv*. DOI: https://doi.org/10.1101/122952

Crickmore MA, Vosshall LB. 2013. Opposing dopaminergic and GABAergic neurons control the duration and persistence of copulation in *Drosophila*. *Cell* **155**:881–893. DOI: https://doi.org/10.1016/j.cell.2013.09.055, PMID: 24209625

Croset V, Treiber CD, Waddell S. 2018. Cellular diversity in the *Drosophila* midbrain revealed by single-cell transcriptomics. *eLife* **7**:e34550. DOI: https://doi.org/10.7554/eLife.34550, PMID: 29671739

Crow M, Gillis J. 2019. Single cell RNA-sequencing: replicability of cell types. *Current Opinion in Neurobiology* **56**:69–77. DOI: https://doi.org/10.1016/j.conb.2018.12.002, PMID: 30654233

Crozatier M, Valle D, Dubois L, Ibnsouda S, Vincent A. 1996. Collier, a novel regulator of *Drosophila* head development, is expressed in a single mitotic domain. *Current Biology* **6**:707–718. DOI: https://doi.org/10.1016/S0960-9822(09)00452-7, PMID: 8793297

Cully DF, Paress PS, Liu KK, Schaeffer JM, Arena JP. 1996. Identification of a *Drosophila melanogaster* glutamate-gated chloride channel sensitive to the antiparasitic agent avermectin. *Journal of Biological Chemistry* **271**:20187–20191. DOI: https://doi.org/10.1074/jbc.271.33.20187, PMID: 8702744

Dason JS, Romero-Pozuelo J, Marin L, Iyengar BG, Klose MK, Ferrús A, Atwood HL. 2009. Frequenin/NCS-1 and the Ca2+-channel alpha1-subunit co-regulate synaptic transmission and nerve-terminal growth. *Journal of Cell Science* **122**:4109–4121. DOI: https://doi.org/10.1242/jcs.055095, PMID: 19861494

Davie K, Janssens J, Koldere D, De Waegeneer M, Pech U, Kreft Ł, Aibar S, Makhzami S, Christiaens V, Bravo González-Blas C, Poovathingal S, Hulselmans G, Spanier KI, Moerman T, Vanspauwen B, Geurs S, Voet T, Lammertyn J, Thienpont B, Liu S, et al. 2018. A Single-Cell transcriptome atlas of the aging *Drosophila* brain. *Cell* **174**:982–998. DOI: https://doi.org/10.1016/j.cell.2018.05.057, PMID: 29909982

Davis FP, Nern A, Picard S, Reiser MB, Rubin GM, Eddy SR, Henry GL. 2020. A genetic, genomic, and computational resource for exploring neural circuit function. *eLife* **9**:e50901. DOI: https://doi.org/10.7554/eLife.50901, PMID: 31939737

Deneris ES, Hobert O. 2014. Maintenance of postmitotic neuronal cell identity. *Nature Neuroscience* **17**:899–907. DOI: https://doi.org/10.1038/nn.3731, PMID: 24929660

DeSalvo MK, Hindle SJ, Rusan ZM, Orng S, Eddison M, Halliwill K, Bainton RJ. 2014. The *Drosophila* surface Glia transcriptome: evolutionary conserved blood-brain barrier processes. *Frontiers in Neuroscience* **8**:346. DOI: https://doi.org/10.3389/fnins.2014.00346, PMID: 25426014

DiAntonio A, Burgess RW, Chin AC, Deitcher DL, Scheller RH, Schwarz TL. 1993. Identification and characterization of *Drosophila* genes for synaptic vesicle proteins. *The Journal of Neuroscience* **13**:4924–4935. DOI: https://doi.org/10.1523/JNEUROSCI.13-11-04924.1993, PMID: 8229205

Doherty J, Logan MA, Taşdemir OE, Freeman MR. 2009. Ensheathing Glia function as phagocytes in the adult *Drosophila* brain. *Journal of Neuroscience* **29**:4768–4781. DOI: https://doi.org/10.1523/JNEUROSCI.5951-08.2009, PMID: 19369546

Enriquez J, Rio LQ, Blazeski R, Bellemin S, Godement P, Mason C, Mann RS. 2018. Differing strategies despite shared lineages of motor neurons and Glia to achieve robust development of an adult neuropil in *Drosophila*. *Neuron* **97**:538–554. DOI: https://doi.org/10.1016/j.neuron.2018.01.007, PMID: 29395908

Estacio-Gómez A, Hassan A, Walmsley E, Lee L, Southall TD. 2019. Dynamic neurotransmitter specific transcription factor expression profiles during *Drosophila* development. *bioRxiv*. DOI: https://doi.org/10.1101/830315

Etheredge J. 2017. *Transcriptional Profiling of Drosophila Larval Ventral Nervous System Hemilineages and Neuroscience*. University of Cambridge Development Physiology. DOI: https://doi.org/10.17863/CAM.17445

Gowda SBM, Paranjpe PD, Reddy OV, Thiagarajan D, Palliyil S, Reichert H, VijayRaghavan K. 2018. GABAergic inhibition of leg motoneurons is required for normal walking behavior in freely moving *Drosophila*. *PNAS* **115**:E2115–E2124. DOI: https://doi.org/10.1073/pnas.1713869115, PMID: 29440493

Greer CL, Grygoruk A, Patton DE, Ley B, Romero-Calderon R, Chang HY, Houshyar R, Bainton RJ, Diantonio A, Krantz DE. 2005. A splice variant of the *Drosophila* vesicular monoamine transporter contains a conserved trafficking domain and functions in the storage of dopamine, serotonin, and octopamine. *Journal of Neurobiology* **64**:239–258. DOI: https://doi.org/10.1002/neu.20146, PMID: 15849736

Grillner S. 2006. Biological pattern generation: the cellular and computational logic of networks in motion. *Neuron* **52**:751–766. DOI: https://doi.org/10.1016/j.neuron.2006.11.008, PMID: 17145498

Gu Z, Gu L, Eils R, Schlesner M, Brors B. 2014. Circlize implements and enhances circular visualization in R. *Bioinformatics* **30**:2811–2812. DOI: https://doi.org/10.1093/bioinformatics/btu393, PMID: 24930139

Hamanaka Y, Park D, Yin P, Annangudi SP, Edwards TN, Sweedler J, Meinertzhagen IA, Taghert PH. 2010. Transcriptional orchestration of the regulated secretory pathway in neurons by the bHLH protein DIMM. *Current Biology* **20**:9–18. DOI: https://doi.org/10.1016/j.cub.2009.11.065, PMID: 20045330

Hanlon CD, Andrew DJ. 2015. Outside-in signaling–a brief review of GPCR signaling with a focus on the Drosophila GPCR family. *Journal of Cell Science* **128**:3533–3542. DOI: https://doi.org/10.1242/jcs.175158, PMID: 26345366

Hardie RC. 1987. Is histamine a neurotransmitter in insect photoreceptors? *Journal of Comparative Physiology A* **161**:201–213. DOI: https://doi.org/10.1007/BF00615241, PMID: 2442380

Harris RM, Pfeiffer BD, Rubin GM, Truman JW. 2015. Neuron hemilineages provide the functional ground plan for the *Drosophila* ventral nervous system. *eLife* **4**:e09443. DOI: https://doi.org/10.7554/eLife.04493

Hartenstein V, Cruz L, Lovick JK, Guo M. 2017. Developmental analysis of the dopamine-containing neurons of the *Drosophila* brain. *The Journal of Comparative Neurology* **525**:363–379. DOI: https://doi.org/10.1002/cne.24069, PMID: 27350102

Hartenstein V. 2018. Development of the Nervous System of Invertebrates. In: Byrne J. H (Ed). *The Oxford Handbook of Invertebrate Neurobiology*. Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780190456757.001.0001

Hewes RS, Park D, Gauthier SA, Schaefer AM, Taghert PH. 2003. The bHLH protein dimmed controls neuroendocrine cell differentiation in *Drosophila*. *Development* **130**:1771–1781. DOI: https://doi.org/10.1242/dev.00404, PMID: 12642483

Hobert O, Kratsios P. 2019. Neuronal identity control by terminal selectors in worms, flies, and chordates. *Current Opinion in Neurobiology* **56**:97–105. DOI: https://doi.org/10.1016/j.conb.2018.12.006, PMID: 30665084

Honegger B, Galic M, Köhler K, Wittwer F, Brogiolo W, Hafen E, Stocker H. 2008. Imp-L2, a putative homolog of vertebrate IGF-binding protein 7, counteracts insulin signaling in *Drosophila* and is essential for starvation resistance. *Journal of Biology* **7**:10. DOI: https://doi.org/10.1186/jbiol72, PMID: 18412985

Howard CE, Chen CL, Tabachnik T, Hormigo R, Ramdya P, Mann RS. 2019. Serotonergic modulation of walking in Drosophila. *Current Biology* **29**:4218–4230. DOI: https://doi.org/10.1016/j.cub.2019.10.042, PMID: 31786064

Huang daW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**:44–57. DOI: https://doi.org/10.1038/nprot.2008.211, PMID: 19131956

Huang Y, Ng FS, Jackson FR. 2015. Comparison of larval and adult *drosophila* astrocytes reveals stage-specific gene expression profiles . *G3: Genes, Genomes, Genetics* **5**:551–558. DOI: https://doi.org/10.1534/g3.114.016162

Ito K, Shinomiya K, Ito M, Armstrong JD, Boyan G, Hartenstein V, Harzsch S, Heisenberg M, Homberg U, Jenett A, Keshishian H, Restifo LL, Rössler W, Simpson JH, Strausfeld NJ, Strauss R, Vosshall LB, Insect Brain Name Working Group. 2014. A systematic nomenclature for the insect brain. *Neuron* **81**:755–765. DOI: https://doi.org/10.1016/j.neuron.2013.12.017, PMID: 24559671

Jois S, Chan YB, Fernandez MP, Leung AK. 2018. Characterization of the sexually dimorphic *fruitless* Neurons That Regulate Copulation Duration. *Frontiers in Physiology* **9**:780. DOI: https://doi.org/10.3389/fphys.2018.00780, PMID: 29988589

Jonsson T, Kravitz EA, Heinrich R. 2011. Sound production during agonistic behavior of male *Drosophila Melanogaster*. *Fly* **5**:29–38. DOI: https://doi.org/10.4161/fly.5.1.13713, PMID: 20953152

Joseph RM, Carlson JR. 2015. *Drosophila* chemoreceptors: a molecular interface between the chemical world and the brain. *Trends in Genetics* **31**:683–695. DOI: https://doi.org/10.1016/j.tig.2015.09.005, PMID: 26477743

Karaiskos N, Wahle P, Alles J, Boltengagen A, Ayoub S, Kipar C, Kocks C, Rajewsky N, Zinzen RP. 2017. The *Drosophila* embryo at single-cell transcriptome resolution. *Science* **358**:194–199. DOI: https://doi.org/10.1126/science.aan3235, PMID: 28860209

Kerner P, Ikmi A, Coen D, Vervoort M. 2009. Evolutionary history of the *iroquois/Irx* genes in metazoans. *BMC Evolutionary Biology* **9**:74. DOI: https://doi.org/10.1186/1471-2148-9-74, PMID: 19368711

Kikuta S, Hagiwara-Komoda Y, Noda H, Kikawada T. 2012. A novel member of the trehalose transporter family functions as an h(+)-dependent trehalose transporter in the reabsorption of trehalose in malpighian tubules. *Frontiers in Physiology* **3**:290. DOI: https://doi.org/10.3389/fphys.2012.00290, PMID: 22934042

Kim B, Shortridge RD, Seong C, Oh Y, Baek K, Yoon J. 1998. Molecular characterization of a novel *Drosophila* gene which is expressed in the central nervous system. *Molecules and Cells* **8**:750–757. PMID: 9895130

Kim H, Kirkhart C, Scott K. 2017. Long-range projection neurons in the taste circuit of *Drosophila*. *eLife* **6**:e23386. DOI: https://doi.org/10.7554/eLife.23386, PMID: 28164781

Kimura K, Sato C, Koganezawa M, Yamamoto D. 2015. *Drosophila* ovipositor extension in mating behavior and egg deposition involves distinct sets of brain interneurons. *PLOS ONE* **10**:e0126445. DOI: https://doi.org/10.1371/journal.pone.0126445, PMID: 25955600

Knauf F, Rogina B, Jiang Z, Aronson PS, Helfand SL. 2002. Functional characterization and immunolocalization of the transporter encoded by the life-extending gene indy. *PNAS* **99**:14315–14319. DOI: https://doi.org/10.1073/pnas.222531899, PMID: 12391301

Konstantinides N, Kapuralin K, Fadil C, Barboza L, Satija R, Desplan C. 2018. Phenotypic convergence: distinct transcription factors regulate common terminal features. *Cell* **174**:622–635. DOI: https://doi.org/10.1016/j.cell.2018.05.021, PMID: 29909983

Lacin H, Chen HM, Long X, Singer RH, Lee T, Truman JW. 2019. Neurotransmitter identity is acquired in a lineage-restricted manner in the *Drosophila* CNS. *eLife* **8**:e43701. DOI: https://doi.org/10.7554/eLife.43701, PMID: 30912745

Lacin H, Truman JW. 2016. Lineage mapping identifies molecular and architectural similarities between the larval and adult *Drosophila* central nervous system. *eLife* **5**:e13399. DOI: https://doi.org/10.7554/eLife.13399, PMID: 26975248

Laughlin SB, de Ruyter van Steveninck RR, Anderson JC. 1998. The metabolic cost of neural information. *Nature Neuroscience* **1**:36–41. DOI: https://doi.org/10.1038/236, PMID: 10195106

Leader DP, Krause SA, Pandit A, Davies SA, Dow JAT. 2018. FlyAtlas 2: a new version of the *Drosophila melanogaster* expression atlas with RNA-Seq, miRNA-Seq and sex-specific data. *Nucleic Acids Research* **46**:D809–D815. DOI: https://doi.org/10.1093/nar/gkx976, PMID: 29069479

Lee G, Villella A, Taylor BJ, Hall JC. 2001. New reproductive anomalies in *fruitless*-mutant *Drosophila* males: extreme lengthening of mating durations and infertility correlated with defective serotonergic innervation of reproductive organs. *Journal of Neurobiology* **47**:121–149. DOI: https://doi.org/10.1002/neu.1021, PMID: 11291102

Lee KM, Daubnerová I, Isaac RE, Zhang C, Choi S, Chung J, Kim YJ. 2015. A neuronal pathway that controls sperm ejection and storage in female *Drosophila*. *Current Biology* **25**:790–797. DOI: https://doi.org/10.1016/j.cub.2015.01.050, PMID: 25702579

Lee PT, Zirin J, Kanca O, Lin WW, Schulze KL, Li-Kroeger D, Tao R, Devereaux C, Hu Y, Chung V, Fang Y, He Y, Pan H, Ge M, Zuo Z, Housden BE, Mohr SE, Yamamoto S, Levis RW, Spradling AC, et al. 2018. A gene-specific *T2A-GAL4* library for *Drosophila*. *eLife* **7**:e35574. DOI: https://doi.org/10.7554/eLife.35574, PMID: 29565247

Lee G, Hall JC. 2001. Abnormalities of male-specific FRU protein and serotonin expression in the CNS of *fruitless* mutants in *Drosophila*. *The Journal of Neuroscience* **21**:513–526. PMID: 11160431

Li H, Horns F, Wu B, Xie Q, Li J, Li T, Luginbuhl DJ, Quake SR, Luo L. 2017. Classifying *Drosophila* olfactory projection neuron subtypes by Single-Cell RNA sequencing. *Cell* **171**:1206–1220. DOI: https://doi.org/10.1016/j.cell.2017.10.019, PMID: 29149607

Limmer S, Weiler A, Volkenhoff A, Babatz F, Klämbt C. 2014. The *Drosophila* blood-brain barrier: development and function of a glial endothelium. *Frontiers in Neuroscience* **8**:365. DOI: https://doi.org/10.3389/fnins.2014.00365, PMID: 25452710

Littleton JT, Ganetzky B. 2000. Ion channels and synaptic organization: analysis of the *Drosophila* genome. *Neuron* **26**:35–43. DOI: https://doi.org/10.1016/s0896-6273(00)81135-6, PMID: 10798390

Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA. 2015. Highly parallel Genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**:1202–1214. DOI: https://doi.org/10.1016/j.cell.2015.05.002, PMID: 26000488

Mamiya A, Gurung P, Tuthill JC. 2018. Neural coding of leg proprioception in *Drosophila*. *Neuron* **100**:636–650. DOI: https://doi.org/10.1016/j.neuron.2018.09.009, PMID: 30293823

Martin CA, Krantz DE. 2014. *Drosophila melanogaster* as a genetic model system to study neurotransmitter transporters. *Neurochemistry International* **73**:71–88. DOI: https://doi.org/10.1016/j.neuint.2014.03.015

Maurange C, Cheng L, Gould AP. 2008. Temporal transcription factors and their targets schedule the end of neural proliferation in *Drosophila*. *Cell* **133**:891–902. DOI: https://doi.org/10.1016/j.cell.2008.03.034, PMID: 18510932

McGinnis W, Krumlauf R. 1992. Homeobox genes and axial patterning. *Cell* **68**:283–302. DOI: https://doi.org/10.1016/0092-8674(92)90471-N, PMID: 1346368

McIlroy G, Foldi I, Aurikko J, Wentzell JS, Lim MA, Fenton JC, Gay NJ, Hidalgo A. 2013. Toll-6 and Toll-7 function as neurotrophin receptors in the *Drosophila Melanogaster* CNS. *Nature Neuroscience* **16**:1248–1256. DOI: https://doi.org/10.1038/nn.3474, PMID: 23892553

Mendes CS, Bartos I, Akay T, Márka S, Mann RS. 2013. Quantification of gait parameters in freely walking wild type and sensory deprived *Drosophila Melanogaster*. *eLife* **2**:e00231. DOI: https://doi.org/10.7554/eLife.00231, PMID: 23326642

Mendes CS, Rajendren SV, Bartos I, Márka S, Mann RS. 2014. Kinematic responses to changes in walking orientation and gravitational load in *Drosophila Melanogaster*. *PLOS ONE* **9**:e0109204. DOI: https://doi.org/10.1371/journal.pone.0109204, PMID: 25350743

Monastirioti M, Linn CE, White K. 1996. Characterization of *Drosophila* tyramine beta-hydroxylase gene and isolation of mutant flies lacking octopamine. *The Neurosci* **16**:3900–3911. DOI: https://doi.org/10.1523/JNEUROSCI.16-12-03900.1996

Namiki S, Dickinson MH, Wong AM, Korff W, Card GM. 2018. The functional organization of descending sensory-motor pathways in *Drosophila*. *eLife* **7**:e34272. DOI: https://doi.org/10.7554/eLife.34272, PMID: 29943730

Nässel DR, Holmqvist MH, Hardie RC, Håkanson R, Sundler F. 1988. Histamine-like immunoreactivity in photoreceptors of the compound eyes and ocelli of the flies *Calliphora erythrocephala* and *Musca Domestica*. *Cell and Tissue Research* **253**:639–646. DOI: https://doi.org/10.1007/BF00219755, PMID: 2902924

Nässel DR, Pirvola U, Panula P. 1990. Histaminelike immunoreactive neurons innervating putative neurohaemal Areas and central neuropil in the thoraco-abdominal ganglia of the flies *Drosophila* and *Calliphora*. *Journal of Comparative Neurology* **297**:525–536. DOI: https://doi.org/10.1002/cne.902970406, PMID: 2117027

Nässel DR. 2018. Substrates for neuronal cotransmission with neuropeptides and small molecule neurotransmitters in *Drosophila*. *Frontiers in Cellular Neuroscience* **12**:83. DOI: https://doi.org/10.3389/fncel.2018.00083, PMID: 29651236

Nässel DR, Zandawala M. 2019. Recent advances in Neuropeptide signaling in *Drosophila*, from genes to physiology and behavior. *Progress in Neurobiology* **179**:101607. DOI: https://doi.org/10.1016/j.pneurobio.2019.02.003, PMID: 30905728

Niven JE, Graham CM, Burrows M. 2008. Diversity and evolution of the insect ventral nerve cord. *Annual Review of Entomology* **53**:253–271. DOI: https://doi.org/10.1146/annurev.ento.52.110405.091322, PMID: 17803455

Noordermeer JN, Kopczynski CC, Fetter RD, Bland KS, Chen WY, Goodman CS. 1998. Wrapper, a novel member of the ig superfamily, is expressed by midline Glia and is required for them to ensheath commissural axons in *Drosophila*. *Neuron* **21**:991–1001. DOI: https://doi.org/10.1016/S0896-6273(00)80618-2, PMID: 9856456

Nusbaum MP, Blitz DM, Marder E. 2017. Functional consequences of neuropeptide and small-molecule co-transmission. *Nature Reviews Neuroscience* **18**:389–403. DOI: https://doi.org/10.1038/nrn.2017.56, PMID: 28592905

Overend G, Luo Y, Henderson L, Douglas AE, Davies SA, Dow JA. 2016. Molecular mechanism and functional significance of acid generation in the *Drosophila* midgut. *Scientific Reports* **6**:27242. DOI: https://doi.org/10.1038/srep27242, PMID: 27250760

Özkan E, Carrillo RA, Eastman CL, Weiszmann R, Waghray D, Johnson KG, Zinn K, Celniker SE, Garcia KC. 2013. An extracellular interactome of immunoglobulin and LRR proteins reveals receptor-ligand networks. *Cell* **154**:228–239. DOI: https://doi.org/10.1016/j.cell.2013.06.006, PMID: 23827685

Pan Y, Robinett CC, Baker BS. 2011. Turning males on: activation of male courtship behavior in *Drosophila Melanogaster*. *PLOS ONE* **6**:e21144. DOI: https://doi.org/10.1371/journal.pone.0021144, PMID: 21731661

Park D, Veenstra JA, Park JH, Taghert PH. 2008. Mapping peptidergic cells in *Drosophila*: where DIMM fits in. *PLOS ONE* **3**:e1896. DOI: https://doi.org/10.1371/journal.pone.0001896, PMID: 18365028

Pauls D, Chen J, Reiher W, Vanselow JT, Schlosser A, Kahnt J, Wegener C. 2014. Peptidomics and processing of regulatory peptides in the fruit fly Drosophila. *EuPA Open Proteomics* **3**:114–127. DOI: https://doi.org/10.1016/j.euprot.2014.02.007

Pauls D, Blechschmidt C, Frantzmann F, El Jundi B, Selcho M. 2018. A comprehensive anatomical map of the peripheral octopaminergic/tyraminergic system of *Drosophila Melanogaster*. *Scientific Reports* **8**:15314. DOI: https://doi.org/10.1038/s41598-018-33686-3, PMID: 30333565

Pavlou HJ, Lin AC, Neville MC, Nojima T, Diao F, Chen BE, White BH, Goodwin SF. 2016. Neural circuitry coordinating male copulation. *eLife* **5**:e20713. DOI: https://doi.org/10.7554/eLife.20713, PMID: 27855059

Philippidou P, Dasen JS. 2013. Hox genes: choreographers in neural development, architects of circuit organization. *Neuron* **80**:12–34. DOI: https://doi.org/10.1016/j.neuron.2013.09.020, PMID: 24094100

Pipes GC, Lin Q, Riley SE, Goodman CS. 2001. The beat generation: a multigene family encoding IgSF proteins related to the beat axon guidance molecule in *Drosophila*. *Development* **128**:4545–4552. PMID: 11714679

Prokop A, Technau GM. 1991. The origin of postembryonic neuroblasts in the ventral nerve cord of *Drosophila Melanogaster*. *Development* **111**:79–88. PMID: 1901786

Rezával C, Nojima T, Neville MC, Lin AC, Goodwin SF. 2014. Sexually dimorphic octopaminergic neurons modulate female postmating behaviors in *Drosophila*. *Current Biology* **24**:725–730. DOI: https://doi.org/10.1016/j.cub.2013.12.051, PMID: 24631243

Rideout EJ, Dornan AJ, Neville MC, Eadie S, Goodwin SF. 2010. Control of sexual differentiation and behavior by the *doublesex* gene in *Drosophila Melanogaster*. *Nature Neuroscience* **13**:458–466. DOI: https://doi.org/10.1038/nn.2515, PMID: 20305646

Ro J, Pak G, Malec PA, Lyu Y, Allison DB, Kennedy RT, Pletcher SD. 2016. Serotonin signaling mediates protein valuation and aging. *eLife* **5**:e16843. DOI: https://doi.org/10.7554/eLife.16843, PMID: 27572262

Robinow S, Campos AR, Yao KM, White K. 1988. The elav gene product of *Drosophila*, required in neurons, has three RNP consensus motifs. *Science* **242**:1570–1572. DOI: https://doi.org/10.1126/science.3144044, PMID: 3144044

Satija R, Farrell JA, Gennert D, Schier AF, Regev A. 2015. Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology* **33**:495–502. DOI: https://doi.org/10.1038/nbt.3192, PMID: 25867923

Schmid A, Chiba A, Doe CQ. 1999. Clonal analysis of *Drosophila* embryonic neuroblasts: neural cell types, axon projections and muscle targets. *Development* **126**:4653–4689. PMID: 10518486

Schor IE, Bussotti G, Maleš M, Forneris M, Viales RR, Enright AJ, Furlong EEM. 2018. Non-coding RNA expression, function, and variation during *Drosophila* Embryogenesis. *Current Biology* **28**:3547–3561. DOI: https://doi.org/10.1016/j.cub.2018.09.026, PMID: 30393032

Schützler N, Girwert C, Hügli I, Mohana G, Roignant JY, Ryglewski S, Duch C. 2019. Tyramine action on motoneuron excitability and adaptable tyramine/octopamine ratios adjust *Drosophila* locomotion to nutritional state. *PNAS* **116**:3805–3810. DOI: https://doi.org/10.1073/pnas.1813554116, PMID: 30808766

Scruggs BS, Gilchrist DA, Nechaev S, Muse GW, Burkholder A, Fargo DC, Adelman K. 2015. Bidirectional transcription arises from two distinct hubs of transcription factor binding and active chromatin. *Molecular Cell* **58**:1101–1112. DOI: https://doi.org/10.1016/j.molcel.2015.04.006

Seeds AM, Ravbar P, Chung P, Hampel S, Midgley FM, Mensh BD, Simpson JH. 2014. A suppression hierarchy among competing motor programs drives sequential grooming in *Drosophila*. *eLife* **3**:e02951. DOI: https://doi.org/10.7554/eLife.02951, PMID: 25139955

Sellami A, Agricola HJ, Veenstra JA. 2011. Neuroendocrine cells in *Drosophila Melanogaster* producing GPA2/GPB5, a hormone with homology to LH, FSH and TSH. *General and Comparative Endocrinology* **170**:582–588. DOI: https://doi.org/10.1016/j.ygcen.2010.11.015, PMID: 21118692

Shepherd D, Harris R, Williams DW, Truman JW. 2016. Postembryonic lineages of the *Drosophila* ventral nervous system: neuroglian expression reveals the adult hemilineage associated fiber tracts in the adult thoracic neuromeres. *Journal of Comparative Neurology* **524**:2677–2695. DOI: https://doi.org/10.1002/cne.23988, PMID: 26878258

Shepherd D, Sahota V, Court R, Williams DW, Truman JW. 2019. Developmental organization of central neurons in the adult *Drosophila* ventral nervous system. *Journal of Comparative Neurology* **527**:2573–2598. DOI: https://doi.org/10.1002/cne.24690, PMID: 30919956

Shih M-FM, Davis FP, Henry GL, Dubnau J. 2019. Nuclear Transcriptomes of the Seven Neuronal Cell Types That Constitute the *Drosophila* Mushroom Bodies . *G3: Genes, Genomes, Genetics* **9**:81–94. DOI: https://doi.org/10.1534/g3.118.200726

Shirangi TR, Stern DL, Truman JW. 2013. Motor control of *Drosophila* courtship song. *Cell Reports* **5**:678–686. DOI: https://doi.org/10.1016/j.celrep.2013.09.039

Sievers F, Higgins DG. 2018. Clustal omega for making accurate alignments of many protein sequences. *Protein Science* **27**:135–145. DOI: https://doi.org/10.1002/pro.3290, PMID: 28884485

Smarandache-Wellmann CR. 2016. Arthropod neurons and nervous system. *Current Biology* **26**:R960–R965. DOI: https://doi.org/10.1016/j.cub.2016.07.063, PMID: 27780069

Smith SJ, Sümbül U, Graybuck LT, Collman F, Seshamani S, Gala R, Gliko O, Elabbady L, Miller JA, Bakken TE, Rossier J, Yao Z, Lein E, Zeng H, Tasic B, Hawrylycz M. 2019. Single-cell transcriptomic evidence for dense intracortical neuropeptide networks. *eLife* **8**:e47889. DOI: https://doi.org/10.7554/eLife.47889, PMID: 31710287

Stahl BA, Peco E, Davla S, Murakami K, Caicedo Moreno NA, van Meyel DJ, Keene AC. 2018. The taurine transporter Eaat2 functions in ensheathing Glia to modulate sleep and metabolic rate. *Current Biology* **28**:3700–3708. DOI: https://doi.org/10.1016/j.cub.2018.10.039, PMID: 30416062

Suska A, Miguel-Aliaga I, Thor S. 2011. Segment-specific generation of *Drosophila* capability neuropeptide neurons by multi-faceted hox cues. *Developmental Biology* **353**:72–80. DOI: https://doi.org/10.1016/j.ydbio.2011.02.015, PMID: 21354130

Tayler TD, Pacheco DA, Hergarden AC, Murthy M, Anderson DJ. 2012. A neuropeptide circuit that coordinates sperm transfer and copulation duration in *Drosophila*. *PNAS* **109**:20697–20702. DOI: https://doi.org/10.1073/pnas.1218246109, PMID: 23197833

Thor S, Andersson SG, Tomlinson A, Thomas JB. 1999. A LIM-homeodomain combinatorial code for motor-neuron pathway selection. *Nature* **397**:76–80. DOI: https://doi.org/10.1038/16275, PMID: 9892357

Truman JW, Schuppe H, Shepherd D, Williams DW. 2004. Developmental architecture of adult-specific lineages in the ventral CNS of *Drosophila*. *Development* **131**:5167–5184. DOI: https://doi.org/10.1242/dev.01371, PMID: 15459108

Truman JW, Moats W, Altman J, Marin EC, Williams DW. 2010. Role of notch signaling in establishing the hemilineages of secondary neurons in *Drosophila melanogaster*. *Development* **137**:53–61. DOI: https://doi.org/10.1242/dev.041749, PMID: 20023160

Truman JW, Bate M. 1988. Spatial and temporal patterns of neurogenesis in the central nervous system of *Drosophila Melanogaster*. *Developmental Biology* **125**:145–157. DOI: https://doi.org/10.1016/0012-1606(88)90067-X, PMID: 3119399

Tsubouchi A, Yano T, Yokoyama TK, Murtin C, Otsuna H, Ito K. 2017. Topological and modality-specific representation of somatosensory information in the fly brain. *Science* **358**:615–623. DOI: https://doi.org/10.1126/science.aan4428, PMID: 29097543

Tuthill JC, Wilson RI. 2016. Parallel Transformation of Tactile Signals in Central Circuits of *Drosophila*. *Cell* **164**:1046–1059. DOI: https://doi.org/10.1016/j.cell.2016.01.014

van der Maaten L. 2014. Accelerating t-SNE using Tree-Based algorithms. *Journal of Machine Learning Research* **15**:3221–3245.

Venkatasubramanian L, Mann RS. 2019. The development and assembly of the *Drosophila* adult ventral nerve cord. *Current Opinion in Neurobiology* **56**:135–143. DOI: https://doi.org/10.1016/j.conb.2019.01.013

Volkenhoff A, Weiler A, Letzel M, Stehling M, Klämbt C, Schirmeier S. 2015. Glial glycolysis is essential for neuronal survival in *Drosophila*. *Cell Metabolism* **22**:437–447. DOI: https://doi.org/10.1016/j.cmet.2015.07.006, PMID: 26235423

von Philipsborn AC, Liu T, Yu JY, Masser C, Bidaye SS, Dickson BJ. 2011. Neuronal control of *Drosophila* courtship song. *Neuron* **69**:509–522. DOI: https://doi.org/10.1016/j.neuron.2011.01.011, PMID: 21315261

Wilk R, Hu J, Blotsky D, Krause HM. 2016. Diverse and pervasive subcellular distributions for both coding and long noncoding RNAs. *Genes & Development* **30**:594–609. DOI: https://doi.org/10.1101/gad.276931.115, PMID: 26944682

Williams MJ, Goergen P, Rajendran J, Klockars A, Kasagiannis A, Fredriksson R, Schiöth HB. 2014. Regulation of aggression by obesity-linked genes *TfAP-2* and twz through octopamine signaling in *Drosophila*. *Genetics* **196**:349–362. DOI: https://doi.org/10.1534/genetics.113.158402, PMID: 24142897

Witt E, Benjamin S, Svetec N, Zhao L. 2019. Testis single-cell RNA-seq reveals the dynamics of de novo gene transcription and germline mutational Bias in *Drosophila*. *eLife* **8**:e47138. DOI: https://doi.org/10.7554/eLife.47138, PMID: 31418408

Wosnitza A, Bockemühl T, Dübbert M, Scholz H, Büschges A. 2013. Inter-leg coordination in the control of walking speed in *Drosophila*. *Journal of Experimental Biology* **216**:480–491. DOI: https://doi.org/10.1242/jeb.078139, PMID: 23038731

Xiong WC, Okano H, Patel NH, Blendy JA, Montell C. 1994. Repo encodes a glial-specific homeo domain protein required in the *Drosophila* nervous system. *Genes & Development* **8**:981–994. DOI: https://doi.org/10.1101/gad.8.8.981, PMID: 7926782

Yang CP, Samuels TJ, Huang Y, Yang L, Ish-Horowicz D, Davis I, Lee T. 2017. Imp and syp RNA-binding proteins govern decommissioning of *Drosophila* neural stem cells. *Development* **144**:3454–3464. DOI: https://doi.org/10.1242/dev.149500, PMID: 28851709

Yellman C, Tao H, He B, Hirsh J. 1997. Conserved and sexually dimorphic behavioral responses to biogenic amines in decapitated *Drosophila*. *PNAS* **94**:4131–4136. DOI: https://doi.org/10.1073/pnas.94.8.4131, PMID: 9108117

Zappia L, Oshlack A. 2018. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *GigaScience* **7**:083. DOI: https://doi.org/10.1093/gigascience/giy083

Zhou B, Williams DW, Altman J, Riddiford LM, Truman JW. 2009. Temporal patterns of broad isoform expression during the development of neuronal lineages in *Drosophila*. *Neural Development* **4**:39. DOI: https://doi.org/10.1186/1749-8104-4-39, PMID: 19883497