Differential Expression of SNMPS and Their Underlying Transcriptional Regulation in Chemosensory Pathways of Drosophila

Richard A. Fandino

University of South Carolina - Columbia

Follow this and additional works at: https://scholarcommons.sc.edu/etd

Part of the Biology Commons

Recommended Citation

This Open Access Dissertation is brought to you by Scholar Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact dillarda@mailbox.sc.edu.
DIFFERENTIAL EXPRESSION OF SNMPs AND THEIR UNDERLYING TRANSCRIPTIONAL REGULATION IN CHEMOSENSORY PATHWAYS OF DROSOPHILA

by

Richard A. Fandino

Bachelor of Science
Temple University, 2006

Submitted in Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy in
Biological Sciences
College of Arts and Sciences
University of South Carolina
2014

Accepted by:

Richard G. Vogt, Major Professor
Richard Showman, Committee Member
Jeff Dudycha, Committee Member
W. Stephen Kistler, Committee Member
Rekha Patel, Committee Member
Lacy Ford, Vice Provost and Dean of Graduate Studies
DEDICATION

I dedicate this effort to my mother.

Her selflessness and dedication is my inspiration.
ACKNOWLEDGEMENTS

My love for scientific research would not have been possible without the encouragement, support, and training from Dr. Jaqueline C. Tanaka. Thank you so much for your continuing interest and moral support. I arrived to the University of South Carolina with the goal to become an integrative scientist, one who could view nature’s curiosities from both a macro and microscopic level. I would like to thank the biology department’s faculty and graduate students for sharing their diverse research and encouraging casual conversations over a pint of brew. Many thanks to my advisor Dr. Dick Vogt for taking a chance on me, you have introduced me to a scientific field that both inspires and fulfills my need to continue exploring.

Thank you to all who have encouraged and emotionally supported me during this long journey of academic training. Many thanks to my partner Troy for filling my life with love and laughter; you make everything brighter. Finally, I would like to thank my mother and my aunt whose boundless love and support have been constant encouragement to rise above.
ABSTRACT

Insects possess complex and diverse chemosensory pathways which have specific phenotypes determined by the regulated expression of specific combinations of genes. While many of these genes (e.g. ORs and GRs) have very narrow expression patterns, associating with only a limited number of chemosensory sensilla, SNMPs show a very broad expression pattern. In *Drosophila*, SNMP1 and SNMP2 associate with the majority of olfactory and gustatory sensilla, but express in different cell types. In olfactory sensilla, SNMP1 and SNMP2 associate with trichoid or coeloconic neurons respectively, and both additionally express in a variety of sensilla support cells; in gustatory sensilla, SNMP2 expresses in neurons while SNMP2 expresses in support cells. We have characterized the SNMP1 and SNMP2 promoters to identify minimal promoter units, to identify regulatory mechanisms directing expression to specific cells or tissues (e.g. wings, legs, neurons, support cells) (3) and to determine if SNMPs and ORs/GRs have independently acquired common regulatory elements.
For both SNMP1 and SNMP2 genes, we identified relatively short regions (~350bp) directly upstream of the transcriptional start sites that directed the majority of expression observed for longer promoters (>3kb). Pertinent to specific tissues, we demonstrated the presence of negative and positive regulatory elements both upstream and downstream of these regions. And we demonstrated the existence of common regulatory elements shared by SNMPs and specific ORs. In general, expression in different tissues or cell types appeared to require specific combinations of regulatory elements. The identification of common cis-regulatory elements in evolutionarily unrelated SNMP and OR/GR genes suggests that chemosensory gene expression is evolutionarily malleable.
# Table of Contents

DEDICATION ................................................................. iii

ACKNOWLEDGEMENTS ...................................................... iv

ABSTRACT ................................................................. v

LIST OF TABLES ......................................................... ix

LIST OF FIGURES ......................................................... x

LIST OF ABBREVIATIONS .................................................. xiii

CHAPTER 1: ORGANIZATION, FUNCTION, AND EVOLUTION OF INSECT CHEMOSENSATION ......................................................... 1

1.1 INTRODUCTION ......................................................... 1

1.2 DROSOPHILA CHEMOSENSATION .................................. 2

1.3 EVOLUTION OF INSECT CHEMOSENSATION ...................... 13

1.4 GENE CHOICE ........................................................ 18

1.5 SUMMARY ........................................................... 23

CHAPTER 2: CHEMOSENSORY SNMP1 AND SNMP2 EXPRESSION IN ADULT DROSOPHILA SENSORY CELLS ......................................................... 29

2.1 INTRODUCTION ......................................................... 29

2.2 RESULTS ............................................................. 31

2.3 DISCUSSION .......................................................... 46

2.4 EXPERIMENTAL PROCEDURE ..................................... 52

CHAPTER 3: CHARACTERIZATION OF SNMP1 GENE REGULATION BY ENHANCER DRIVER DISSECTION ......................................................... 68

3.1 INTRODUCTION ......................................................... 68
3.2 Results ................................................................. 73
3.3 Discussion ......................................................... 79
3.4 Experimental Procedure ........................................... 81

Chapter 4: Characterization of SNMP2 Gene Regulation by Promoter Dissection .................................... 95
4.1 Introduction ............................................................ 95
4.2 Results ................................................................. 99
4.3 Discussion ............................................................. 104
4.4 Experimental Procedure ............................................ 107

Chapter 5: Evolution Through Cis–Regulatory Change ............... 119
5.1 Introduction ............................................................ 119
5.2 Transcriptional Regulation: Mechanism ......................... 120
5.3 Functional Role of Cis–Regulation ................................. 122
5.4 Cis–Regulatory Evolution ............................................. 124
5.5 Summary: Insect Chemosensory Evolution Revisited .......... 128

References ..................................................................... 134
LIST OF TABLES

Table 2.1 SNMP1 expression in antennal lobe............... 67
Table 3.1 SNMP1 upstream dissection primers............... 94
Table 4.1 SNMP2 upstream dissection primers............... 118
LIST OF FIGURES

Figure 1.1 Drosophila olfaction ........................................ 25
Figure 1.2 Olfactory processing in insects ......................... 26
Figure 1.3 Drosophila gustation ....................................... 27
Figure 1.4 Chemosensory projections to the fly CNS ............ 28
Figure 2.1 Snmp1-Gal4 and Snmp2-Gal4 expression
in larvae ................................................................. 59
Figure 2.2 Peripheral expression of GFP in
Drosophila adult head ............................................. 60
Figure 2.3 Peripheral expression of GFP in
Drosophila adult wing and haltere............................ 61
Figure 2.4 Peripheral expression of GFP in
Drosophila prothoracic legs and tarsi....................... 62
Figure 2.5 Immunohistochemistry showing neuronal
projections of SNMP1 and SNMP2
in Drosophila CNS .................................................. 63
Figure 2.6 Immunohistochemistry showing neuronal
projections to the glomeruli of the AL in
Snmp1-Gal4 and Snmp2-Gal4 .................................. 64
Figure 2.7 Immunohistochemistry showing neuronal
projections to the SOG of Snmp1-Gal4 and
Snmp2-Gal4 adult flies ........................................... 65
Figure 2.8 Immunohistochemistry showing neuronal
projections to the TG in Snmp2-Gal4
adult male and female ............................................. 66
Figure 3.1 Comparitive approach between
D.melanogaster and D.pseudoobscura identifying
conserved elements in 5’ upstream region ................... 89
Figure 3.2 Neuronal expression of SNMP1-Gal4 dissected regulatory region showing differences in expression................................. 90

Figure 3.3 RSAT oligo-analysis output of S1.4 upstream sequence compared to orthologous and co-expressed genes................................. 91

Figure 3.4 Mutagenesis of GCAATTA element in S1.4-Gal4 driver sequence........................................ 92

Figure 3.5 Expression of SNMP1 driven mCD8::GFP in different enhancer drivers revealing changes in glomeruli expression.............. 93

Figure 4.1 Comparitive approach between D.melanogaster and D.pseudoobscura identifying conserved elements in 5’ upstream region of SNMP2 ........................................ 114

Figure 4.2 Neuronal expression of SNMP2 dissected regulatory region show minimal differences in expression........................................ 115

Figure 4.3 MEME motif-finder analysis of S2.2(0.90kb)upstream sequence motif compared to orthologous and non-related co-expressed chemosensory genes........................................ 116

Figure 4.4 Mutagenesis of TTTAATCG element in S2.2-Gal4 driver sequence........................................ 117

Figure 5.1 Cis-transcriptional Regulation........................................ 131

Figure 5.2 Proposed mechanism of TF binding........................................ 132

Figure 5.3 Mechanism of Regulatory Evolution........................................ 133
LIST OF ABBREVIATIONS

ACCf.............................. Anterior Cerebrocervical Fascicle
AMMC............................. Antennal Mechanosensory and Motor Center
AMS................................. Anterior Maxillary Sensory
CD36................................. Cluster of Differentiation 36
CNS.................................. Central Nervous System
CRE.................................. Cis-Regulatory Element
cVA................................. Cis-Vaccenyl Acetate
DCSO................................. Dorsal Cibrial Sense Organ
GR.................................. Gustatory Receptor
GSN.................................. Gustatory Sensory Neuron
IR..................................... Ionotropic Receptor
LbN.................................... Labial Nerve
LSO.................................... Labral Sense Organ
OBP.................................... Odorant Binding Protein
ODE.................................. Odorant Degrading Enzyme
OR..................................... Olfactory Receptor
ORCO................................. Olfactory Receptor Co-receptor
OSN.................................... Olfactory Sensory Neuron
PGC.................................. Primary Gustatory Center
PMS.................................. Primary Maxillary Sensory
SNMP................................. Sensory Neural Membrane Protein
SOG.................................... Sub-oesophageal Ganglion
SR-B1.................. Scavenger Receptor class B member 1
TF.................................. Transcription Factor
TG.................................. Thoracic Ganglion
VCSO.............................. Ventral Cibrial Sense Organ
VPS............................... Ventral Pharyngeal Sensory
CHAPTER 1
ORGANIZATION, FUNCTION, AND EVOLUTION
OF INSECT CHEMOSENSATION

1.1 INTRODUCTION
Organismal communication with the environment and conspecifics is mediated through chemosensory perception. In an environment permeated with a great diversity of chemical signals, an organism’s fitness may be measured by its capacity to selectively sense biologically relevant chemicals (1-3). Relevant chemical signals are those that modulate behaviors towards an appropriate substrate for feeding or oviposition, choosing an adequate mate, or avoiding environmental toxins and predators. In insects a sensory relevant chemical is processed within porous sensory hairs, named sensilla. Within sensilla the often hydrophobic molecule is solubilized and transported to the sensory cell membrane through soluble transport proteins and is delivered to the chemoreceptor, which in turn activates a cascade of events depolarizing the cell, sending signals to the brain where the message is
integrated resulting in a behavioral output. The vast diversity of sensory chemicals in the environment is processed by specific organization of expansive multigene chemosensory families forming a physical and functional sensory map.

Increasing complexity of chemosensory reception can be traced through major episodes of insect evolution (4-6). The majority of chemosensory proteins in insects have come about by convergent evolution resulting in a similarly organized sensory system as that of vertebrates. However, the method of sensory coding is subject to much evolutionary quandary because of the level of expanding and contracting gene families reflecting olfactory behaviors specific to the ecological niche of the insects (7). Moths, for example, have relatively large and finely tuned antennae, which are able to interpret spatial and temporal information from the stochastic distribution of pheromonal filaments in the atmosphere facilitating long distance mate location (8). The honeybee is able to preform complex associative learning through sharing of gustatory signals from newly foraged food sources (9). The fruit fly has a complex and stereotyped mating routine involving the integration of olfactory, gustatory, and mechanosensory inputs accessing mate fitness and inducing female
receptivity (10). To facilitate an understanding of the evolution and diversity of chemosensory driven behaviors in insects, the behavior is broken down to the relevant molecular and physiological processes. What is found is a hierarchical network of proteins and sensory circuits eliciting complex behaviors modified through evolutionary time resulting in ecological and behavioral adaptations.

1.2 Drosophila Chemosensation

The molecular and genetic model organism Drosophila melanogaster has provided a richness of information expanding our knowledge on the biochemical and neuronal interface between chemistry and behavior.

Chemosensory modalities are broadly separated into smell and taste, much like the vertebrate nose and mouth, insects have specialized organs with focal functions. The adult fly is able to discriminate a vast array of sensory information through a highly organized network of stereotypically distributed sensory hairs, called sensilla, distributed throughout the surface of the chemosensory organs. Sensilla are porous hair-like structures with the purpose of encapsulating the sensory cells from environmental exposure (11). Each sensilla houses one or
many sensory neurons and support cells contributing to an individual sensory profile.

**D. melanogaster Olfaction**

*Drosophila* olfaction is localized to the antennae and maxillary palps. Both antennae and palps have olfactory sensilla arranged in a stereotyped pattern (12, 13) (Fig 1.1A). The pattern is further characterized as exhibiting groupings of morphologically distinct sensilla. Olfactory sensilla are porous cuticular hairs on the surface of the olfactory epithelia belonging to three morphological types: club-shaped large / small basiconic sensilla, spine-shaped trichoid sensilla, and small, cone-shaped coeloconic sensilla (14) (Fig 1.1B). Each sensilla is also characterized by broad functional role grouping together in distinct zones on the antennae. Large basiconic sensilla, sensitive to food odors, cluster at the medial-proximal side of the antennae and the maxillary palps. Trichoid sensilla, pheromonal, cluster at the lateral-distal edge, the number of trichoid sensilla is dependent on sexual dimorphic developmental constraints, with males exhibiting a greater number of trichoid sensilla. Small basiconic and coeloconic (amine and acid sensing) are interspersed in the middle region of the antennae (15-17).
Olfactory sensilla house one to four OSNs adding to a total population of ~1200 OSNs in antennae and ~120 in maxillary palps (12). The OSNs are bipolar cells extending an apical dendrite into the sensillum shaft and an axon afferent to the olfactory centers of the brain. The support cells are made up of three distinct glial-like cell types: tricogen (shaft), thormogen (socket), and thecogen (sheath) cells (18) (Fig 1.1C). Thecogen cells have a small, flattened cell body attached to the surface of the neuron cell body and dendrite. Both trichogen and tormogen cells are large cells bodies that lie adjacent to the neuron on opposite sides, trichogen is close to the surface of the epidermis, while thecogen orient below the trichogen cells (19). The support cells function as a physical and physiological separation of individual neurons in addition to secreting the extracellular lymph. The OSNs and support cells express a network of chemosensory proteins involved in granting each sensillum with odorant selective functionality.

The support cells express a small (13.4 – 28kDa) soluble protein, first discovered in moth antennae, the odorant binding proteins (OBPs) (20). The role of most OBPs have not been fully characterized, however it is broadly interpreted that they are involved in direct
binding of an odorant, possibly modulating the accessibility of a signaling molecule to its cognate receptor (21) (Fig1.1D). The olfactory receptor (OR) proteins are a membrane bound receptor that bears no sequence or topological homology the vertebrate ORs (22) (23). The ORs form functional heterodimer cation channels on the neuronal surface through a OR chaperone and co-receptor ORCO (24). Neuronal and support cells express the sensory neural membrane protein (SNMPs), a two-transmembrane receptor homologous to the human CD36 family of fatty-acid transporters are believed to be involved in modulating the ligand bound OR activity by serving as an additional co-receptor in OSNs of trichoid (25), basiconic (26), and intermediate sensilla (27). SNMP is broadly expressed in a number of chemosensory cells. An additional member to the grouping of olfactory proteins are the soluble odorant degrading enzymes (ODEs), these form the least characterized of the chemosensory proteins, thought to function in modulating or degrading the odorant signal in the sensillar space (28). Recent studies have found another family of chemoreceptors expression coeloconic sensilla, the ionotropic receptors (IRs) (29). These receptors are related to the vertebrate ionotropic glutamate receptors (IGluR), however they have
convergently evolved in invertebrates. The IRs associate with coeloconic sensilla OSNs and have been characterized in sensing specific amines and acids (30). The olfactory biochemical network functions in defining the odorant specificity for each cell.

The antenna is central to the perception of a wide array of odorants. The basiconic sensilla house OSNs selective to a diversity of food related odorants, trichoid is generalized to pheromonal functions, and coeloconic to organic acids (31). Palp basiconic OSNs mediate the sensory enhancement of gustatory signals through olfactory signaling (32). Many significant odorants in Drosophila are perceived in a combinatorial fashion, with broad groups of sensilla OSNs sending signals to the brain where information is further processed. OSNs have been largely classified by the OR expressed and the functional role of that OR. OSNs express only one, rarely two functionally distinct ORs per neuron (33-35). Axonal projections from OSNs expressing the same receptor synapse onto specific olfactory glomeruli of the antennal lobe. The hierarchical mapping of olfactory detection permits the integration and modulation of signals from ~1300 OSNs in 550 - 600 sensilla, expressing any number of ~62 ORS, to 50 discrete glomeruli (36) (Fig 1.2). The glomeruli integrate sensory
cell synapses with interneurons, involved in “fine-tuning” of the olfactory signal through sensory signal inhibition and projection neurons involved in sending the signal to higher order processes (Fig1.2A).

**D. melanogaster Gustation**

Gustation is broadly represented in the adult fly (Fig1.3A). The main taste organ is the most distal portion of the proboscis, the labellum, analogous to the human tongue. Additionally, gustatory sensilla have been identified in the pharynx, legs, wings, and female genitalia. The labellum consists of two palps with four rows of thirty-one stereotypically distributed sensilla on both sides (Fig1.1B). Much like the olfactory organs, the labellar palps are covered with sensilla of different morphological types; 12 short (s) sensilla, 10 intermediate (i), and 9 long (l), housing between 2 (s) or 4 (i and l) gustatory sensory neurons (GSNs) accompanied by 1 mechanosensory neuron. The labial palps expose during feeding and reveal a region of minimally characterized sensilla, classified as pegs because of their short nub-like morphology, taste pegs (37). Taste pegs are dual sensing, housing one GSN and one mechanosensory neuron, these sensilla make contact with the food as it enters the pharynx. The pharynx has three separately distributed
clusters of taste sensilla, the labral sense organ (LSO), dorsal and ventral cibirial sense organs (DCSO and VCSO). The pharyngeal sensory cells are thought to be involved in a level of internal quality control of food shortly prior to being ingested. Gustatory sensilla on the legs and wings are stereotypically distributed amongst many mechanosensory sensilla, distinguishable by a long / curved phenotype. Gustatory sensilla in the legs house two to four GSNs, accompanied by one mechanosensory neuron. Distribution on the tarsal segment of the foreleg is sexually dimorphic; ~50 chemosensory sensilla in males, ~37 in females. The middle legs and hind legs have 30 and 32 sensilla respectively. The wing gustatory sensilla are organized in regular intervals on both dorsal and ventral sides of the wing within range of the anterior wing margin. GSNs of gustatory sensilla in the labellum, legs, and wings are all responsive to sugar and salt perception, the proboscis has sensitivity to bitter as well (14, 38, 39). The dimorphic patterning and number of sensilla in the legs is thought to function in mate selection and copulation, suggesting an additional functionality in these GSNs. The functional nature of GSNs in the wing are still largely unknown, but are believed to be involved in qualifying tastants in an
environment before ingesting and/or routine gustatory functions like grooming (40).

Gustatory sensilla are characterized as having a single pore permitting access of tastants into an extracellular environment similar to previously described olfactory biochemical network. GSNs are enveloped within the three glial support cells, tricogen, thecogen, and thormogen, which secrete lymph into the sensillar space. The biochemical network of proteins involved in coding for gustatory signals are less understood than of olfactory neurons, however a similar array of ligand specific proteins are co-expressed mediating taste perception. The gustatory receptors (GRs) are a gene family of 68 divergent genes coding for receptor involved in mediating taste (41-45). GRs are expressed in the GSN membrane and may express up to five different GRs per neuron (14). Most GRs are receptors for a wide range of soluble ligands that mediate attractive (sweet), aversive (bitter), and pheromone signals (44, 46). It is assumed that GRs group within a subfamily based on structural and response profile similarities based on the similar nature of mammalian GRs and insect ORs (47, 48). For example, Gr5a, a trehalose receptor, forms part of a subfamily of GRs encoding sensitivity to diverse sugars (44, 49). Additionally, the
Gr68a, a pheromone receptor crucial for efficient mating, group within a subfamily of pheromone receptors (50). Bitter sensitive receptors are vastly more represented and form various groups, suggesting the necessity of highly specific bitter receptors (51).

A few principles arise in gustatory coding mediated by the GRs. GRNs may express a large number of GRs conferring single ligand specificity. Some GRs express in a large number of GSNs and may function as a co-receptor with other narrowly tuned GRs (52) (53-57). OBPs were originally described in olfactory tissue, however several OBPs have been found to express in support cells of gustatory sensilla conferring signal specificity or modulation. Characterized OBPs have been implicated in pheromonal perception in GSNs expressing with neuronal Gr32a in tarsal gustatory sensilla (58). Many foods are comprised of a combination of sweet and bitter, stimulating opposing feeding behaviors. A recent study found that OBP49a is involved in mediating sugar (attractive) and bitter (aversive) stimulations, by inhibiting an attractive response in presence of aversive signals (21). Other taste sensitivities have been recently found to affect GSN activity or Drosophila feeding behaviors; water sensitive GSNs (59), CO2 sensing GSNs, salt sensing, and fatty acid
sensing (60). In this study both SNMP1 and SNMP2 are shown to express in support cells and neuronal cells of gustatory sensilla. The broad gustatory expression of SNMP2 in GSNs suggests its role in coding gustatory neuron sensitivity. In mammals CD36 expressing in gustatory oral tissue is believed to be involved in fatty acid taste (61, 62). The highly conserved SNMP belonging to the CD36 family might reveal to have a role in fatty acid taste perception.

GSNs send multiple overlapping neuronal projections to gustatory centers (Fig1.3B,C). Difficulty in characterizing the GSNs is partially a result of overlapping neuronal projections to the brain and thoracic ganglia (TG). GSNs from proboscis, labellum, and tarsi all project to the suboesophageal ganglion (SOG), located in the tritocerebrum ventrally to the ALs in the Drosophila brain. Studies using the indirect cell labeling Gal4/UAS expression system, have characterized projections based on GR expression profiles or selection of transgenic lines with broad non-specific expression in the labellum.

The SOG is made up of a neuropil with no clear defined boundaries, receiving projections from sensory neurons from the pharynx, labial palps, and tarsi. Defining a gustatory map has proven tedious because of the lack of physical boundaries and the colocalization of many gustatory
projections, making the mapping of taste signals dependent on the individual characterization of single or groups of sensilla (37, 63). Detailed characterizations of the SOG projections have identified zonal segregation within the primary gustatory center (PGC) of the SOG (Fig1.3B). The PGC receives inputs from the labellar sensilla and is subdivided into 3 broadly categorized projections: Primary maxillary sensory (PMS 1-3), anterior maxillary sensory (AMS 1-3), and the Lateral sensory (LS 1-3). Forming nine distinct zones receiving projections from bitter sensing (Gr66a) or sweet sensing (Gr5a) GSNs from the labial lateral sensilla and medial taste pegs, in addition to 3 zones from the labellum receiving inputs from water sensing (PMS 4-5) and CO2 sensing sensilla (AMS1). Additional zones, the vental pharyngeal sensory (VPS 1-3) and anterior cerebrocervical fascicle (ACCF 1-3), receive inputs from the pharyngeal nerves and the tarsi respectively (63).

The TG receives a number of chemosensory and mechanosensory projections from legs and wings (64) (Fig1.3C, Fig1.4). The TG is separated into the prothoracic, midthoracic, and metathoracic neuromeres, receiving afferents from the frontal, mid, hindlegs respectively (Fig1.3C). The most ventral neuropil receives projections from the female genitalia (63). Projections
form the wings are located between the prothoracic and mesothoracic neuromeres. Chemosensory projections from the appendages characteristically synapse close to the central midline of the neuromeres, the mechanosensory projections synapse closer to distal lateral borders with ipsilateral chiral projections extending dorso-laterally in the TG (65).

1.3 Evolution of Insect Chemosensation

Most insect chemosensory proteins belong to large and highly divergent gene families. Gene expansions occur through repeated gene duplications events either by transposition or interchromosomal translocations. The chemosensory protein repertoire presents a fascinating evolutionary laboratory exhibiting increasing complexity throughout evolutionary time, gene expansion and contractions at different rates, and species-specific ecological adaptation in chemoreceptor evolution (66). Insect evolutionary and ecological success may be aided by adaptive responses from chemosensory gene subfunctionalization, elevating success rates in diverse chemical environments.
The insect chemoreceptor family is comprised of convergently evolved receptor families, the ORs, GRs, and IRs (4, 67–69). Olfactory receptor gene evolution has been extensively studied in terms of a highly divergent nature and critical role in odorant perception. The ORs were first identified in vertebrates and subsequently in insects. In insects ORs are found to be a highly expanded and diverged gene family, between orders and even species. Often making the interspecies mapping of ORs difficult. The OR protein family is thought to follow a birth-and-death model of gene expansion.

In evolutionary history insects mark, to this day, the most diverse species on earth, establishing populations through diverse ecosystems. Insect terrestrialisation occurred at the point where crustacea led way to the terrestrial hexapoda. It is speculated that the hexapoda arthropods diversification coincided either before or during the diversification of land plants. The strictly land dwelling hexapoda gave way to the pterygota (flying insects), it is thought that flight had an important role in long distance mobility, vastly exploiting diverse environments. The flying insects shortly, in a relative sense, developed the capacity for metamorphosis (Holometabola), the evolutionary successful holometabola
gave way to the formation of complex societies, the
eusocial insects. These points mark the major
morphological changes in insects and coincide with the
major adaptations to the chemosensory system.

The transition from crustacea to hexapoda must have
signified a transition from chemoreception from an aqueous
to a gaseous environment (70, 71). The evolution of the
ORs and co-receptor ORCO are thought to indicate an
explosion of volatile semiochemicals, therefore the early
arthropods and crustacean would exemplify a ‘taste’ based
chemoreceptor family. In crustacea, the Daphnia pulex was
found to representative of earlier aquatic based organism
that led to the hexapoda terrestrialization have an expanded
repertoire of GRs but no ORs. The IRs are an ancient
chemoreceptor family with two members found in Daphnia (4)
and are thought to be the earliest form of chemoreception
in insects. Transcriptome analysis of apterygote hexapoda
(wingless terrestrial insects) modern representatives such
as the Archaeognatha (jumping bristtails) and Zygentoma
(silverfish and firebrat), reveal their chemoreceptor
repertoire to be highly representative of the IR gene
family (6). In addition to IR chemoreceptor family the
Zygentoma Thermobia domestica was found to express three
ORCO like OR genes. GRs were found for both but in L. y-
signata there was no evidence of GRs represented in gustatory epithelia.

The phasmatodea order of insects, popularly known members of this order are the walking sticks, these are representative of early Pteryogota species. Transcriptomic data from the Phylium siccifolium identified expanded OR family with co-receptor ORCO representing the chemoreceptor repertoire for this order of insects. This data is of a high level of importance in understanding the evolution of chemoreception in early insects demonstrating that ORs began to emerge past the Zygentoma family and went through extensive divergence upon the appearance of early Pterygota insects. This signifies a chemosensory transition from aquatic to terrestrial with transition from GRs to the early and possibly broadly functioning IRs, to olfactory complexity with the convergently evolved function of flight. This progress in chemosensation leads to the chemosensory evolution of the Drosophilid species and the advanced and chemosensory neurocomplexity of the eusocial insects.

Comprehensive analysis of the twelve Drosophila species has given insight into the fine-scale tuning of the OR repertoire as many of these species are representative of distinct ecological adaptations (72). The ORs and GRs in
Drosophila species appear to be scattered throughout the genome, at times making the identification of diverged ORs a troublesome feat. However, clustering of these genes are more evident in other holometabulous species: *Apis mellifera* (73), *Aedes aegypti* (74), and *Tribolium castaneum* (75). Comparison of gene gain and loss events between Drosophila species demonstrates a rather conserved structure of genes compared to distant holometaboulous species, for example the genome of *Apis mellifera* reveals an extensive expansion of ORs, while the GRs have been greatly reduced. It is believed that the chemosensory proteins arise through a birth and death model where a gene duplication event leads to the gene copies evolving separately and diverging either through gain of function or are lost through pseudogenization. Within the Drosophila species the most dramatic example of gene gain and loss is found in the *Drosophila grimshawi* has undergone twenty-seven OR gains and fourteen OR loss, compared to *D. melanogaster* which has gone through four OR gains and 12 OR loss. The dramatic difference between *D. grimshawi* and *D. melanogaster* is possibly a result of chemosensory adaptation through positive selection, since the *grimshawi* family of Drosophila are endemic to the Hawaiian islands,
therefore they possess an expected expanded chemosensory repertoire.

The chemoreceptor genes are subject to repeated expansion through inter-chromosomal translocations then other chemosensory genes (76). Chemoreceptors are distinct in their highly specified expression pattern; each sensory neuron expresses one to a few chemoreceptor genes out of a large repertoire. The spaced translocations possibly give way for diversity in regulatory elements controlling the spatial distribution of individual chemoreceptors. Therefore, the subfunctionalization of a newly duplicated gene could possibly be a function of a newly adopted regulatory code.

1.4 Gene Choice

Most chemosensory proteins form part of diverse families of functional genes, each with different recognition sites and as a singular mechanism able to confer ligand specificity and selectivity in one neuronal sensory cell out of many. This raises a very important question, how is a regulatory network able to selectively choose one out of numerous genes to express in one particular cell from the same developmental precursor?
The Drosophila sensory appendages and sensilla develop from a third larval instar and early pupal stage grouping of primordial cells referred to as the imaginal disc. The proneural genes, encoding the achaete–scute family of helix-loop-helix (bHLH) transcription factors specify the sensory hairs located in the antennae (77). These express in the group of cells, the proneural cluster providing the TF network for the initial cellular identity of the precursor cells. The proneural genes Amos and Atonal provide the sensillar identity in the precursor cells, Atonal controls the fate of sensory cells including the coeloconic cells (Gupta and Rodrigues, 1997), while Amos specifies the two remaining morphological types basiconic and trichoid (Goulding et al., 2000). Two to Three cells are selected from the precursors to become sensory cells forming the pre-sensillum cluster, further division of the PSC leads to recruitment of neighboring cells granting the cellular identity of the neuronal and support cells (18).

The spatial organization of sensilla types is a result of specific recruitment pattern of amos and atonal throughout the developing antennal imaginal disc (78). The developed OSN is functionally specified by the chemosensory proteins and synaptic wiring to the antennal lobe. A few
transcription factors have been implicated in functional identity of the OSN, it is believed that the POU-domain TF Acj6 and Notch mediating signaling are involved in the recruitment of specific OR to certain OSNs (79-81). Unlike the vertebrate system where specific wiring to the olfactory bulb is dependent on OR expression in the sensory cell, Drosophila glomerular targeting of the OSNs seems to involve an intrinsic process based on dual functionality of cellular identity transcription factors like Acj6 and Notch. Although OR expression does not play a decisive role in OSN targeting, the regulation of OR gene expression and OSN targeting are integrated during functional development of the OSN and further studies are identifying a number of TF involved in specifying the expression pattern of ORs (81, 82).

Developmental programs affecting the pattern and distribution of sensory cells and sensilla, coupled with observations of specific TFs in directing OR expression suggest that OR gene choice is a result of a diverse network of TFs triggered by specific proneural genes. Gene regulation occurs through many pathways, specific spatial and temporal control is coded within the non-transcribed regulatory regions flanking a gene. Suggesting, that the upstream regulatory elements flanking the coded OR gene
hold specific nucleotide binding sites that specify the code for “choosing” a chemosensory gene out of an array of multiple genes. The precise expression of chemosensory genes in one sensory neuron provides the functional identity for that neuron.

Fuss and Ray 2009, proposed mechanisms of gene choice are either deterministic or stochastic. Deterministic models suggest two possible methods of gene choice: 1) A “one-to-one” model suggest that the regulatory region for each chemosensory gene, in this case the ORs, has an enhancer for a specific TF. The model suggests that each OSN is expressing a specific or a grouping of TFs for each OR. 2) In a similar fashion the combinatorial method suggest that gradients of TF expression throughout the sensory cell detail the chemosensory gene identity for each neuron, that is, overlapping gradients would express alternate genes. The stochastic models are based on post-transcriptional modification of expression: 1) Feedback regulation suggest that each sensory cell expresses a group of chemosensory genes that by means of a functional protein is able to produce a negative feedback mechanism limiting the expression of other genes. 2) The locus control region model depends on an observed correlation between functionally similar genes and organization in the genome,
by which a single regulatory locus can act either through cis- or trans- mechanism to stochastically select the expression of one gene out of a number of grouped chemosensory genes. However, this does not explain a highly organized mapping of similar versus distinct sensory neurons and functions. Previous data on the question of OR gene choice in Drosophila has discovered that co-expressed genes in the same neuronal type are under the control of identical enhancer elements (83). Furthermore, comparative analysis of regulatory control in twelve Drosophila orthologues revealed that these enhancer sequences are conserved, mutation to certain enhancer elements at time led to misexpression of the gene, suggesting that TFs specific for the spatial expression of each gene may very well be broadly represented. Allowing specific gene choice to arise from a combinatorial code of enhancer / repressor elements. These studies support the hypothesis that Drosophila chemosensory gene choice is a result of a diverse cis-regulatory code furthering the functional identity of each sensory cell.
1.5 Summary

Insect chemosensation is mediated by a highly diverse and adaptive network of proteins, which are a result of a gradual adaptive specialization to a chemical environment. The Drosophila chemosensory system is organized in a stereotyped zonal distribution of sensilla on the sensory epithelia, which each sensillum housing one or two sensory neurons, greater than two for gustatory sensilla. Each sensory neuron asymmetrically develops with three support cells and the neuron is pre-wired to the antennal lobe. The sensory neuron is functionally tuned to either a range of related sensory chemicals by a network of chemosensory proteins that aid in specifying and modulating the signal from the chemoreceptor. Most chemosensory genes belong to expansive gene families that have gone through gain and loss events characterizing the functional sensory code for each insect reflecting its life history and ecological niche. Amongst a diversity of genes each neuron should be able to specify a precise network of chemosensory proteins as to specifically tune each neuron to a specific chemical signal profile. The function of choosing a repertoire of proteins in a specific grouping of sensory cells poses an interesting model for understanding how cis-regulatory
mechanisms have evolved in possibly being a primary source of altering the sensory code during speciation.

In this study we have focused on the *Drosophila* melanogaster SNMP chemosensory protein because of its unique evolutionary and functional properties. SNMPs are a highly conserved chemosensory protein family belonging to the vertebrate type B scavenger receptor CD36 (SR-B1 / CD36) family of two transmembrane fatty acid transporter proteins. The expression pattern for this protein is unique in that it is found to associate between sensory modalities (olfactory and gustatory) and sensory cell type (neuronal and support cell). The broad expression pattern for SNMPs promoted the necessity to understand the question of chemosensory gene choice. It was found that the SNMPs are under a similar mechanism of cis-regulatory control as previously characterized ORs, suggesting that enhancer and repressor elements are conserved in the formation of the biochemical network of proteins characterizing each neuron within each sensilla in the formation of an evolutionarily selected chemosensory code.
Figure 1.1: *Drosophila* olfaction.

A. Electron microscopy image of Drosophila head, depicting chemosensory sensilla covering the antennae and palps (Image courtesy of John Janeike, Rochester University). B. Zonal distribution of basiconic, coeloconic, and trichoid sensilla on antennal surface. C. Olfactory sensory neuron (OSN) and support cells (Tricogen, Tormogen, and Thecogen) bordering the OSN cell body. D. Proposed pathway of chemosensory proteins aiding in chemoreceptor function: Odorant binding proteins (OBP), OR (Olfactory Receptor), ORCO (OR Co-receptor), SNMP (Sensory neural membrane protein).
Figure 1.2: Olfactory processing in insects.

A. Topographic map of sensilla organize OSNs of similar functions. OSNs expressing the same OR send projections to one glomerulus. The signal is integrated through inhibitory synapses by the lateral neurons (LN), the integrated signals travels to the center of higher order processing and integration in the Mushroom body/Lateral horn by projection neurons (PN). Image adapted from Pelligrino and Nakagawa, 2009. B. Confocal image of Brain 20x the AL is highlighted in Red. C. 100x confocal image of the glomeruli in the antennal lobe depicts clear boundaries between glomerular neuropil.
Figure 1.3: *Drosophila* Gustation.

A. *Drosophila* gustation occurs in gustatory sensilla located in proboscis, legs, and wing. Identified by red circles. B. Gustatory sensory neurons (GSN) from the labellum, pharynx, and frontal leg tarsi send projections to the sub-oesophageal ganglion (SOG) (highlighted in blue). The central region of convergence of gustatory projections from the primary gustatory center (PGC). The SOG has no physical boundaries between functionally distinct neuropil. C. The thoracic ganglia (TG) feature distinct neuromeres. The prothoracic neuromeres receives inputs from the frontal leg, followed by projections from the wings. The mesothoracic neuromere receives projections from mid-legs, metathoracic from the hing legs. Inputs from the female genetalia synapse onto the most ventral portion of the TG.
Figure 1.4: Chemosensory projections to the fly CNS

Image adapted from Stocker, 2004. Olfactory sensilla (Olf) are multi-porous cuticular structures housing the one to four OSNs surrounded by support cells. Projections from antennae (ANT) and palps (MP) (blue, orange) synapse on the Brain. Gustatory sensilla (Gus) have a single pore and often contain four GSNs and one mechanosensory. Gustation is more broadly represented. Projections from the labellum (84) (red) and pharyngeal GSNs (PHAR) (green) synapse on the SOG. Projections from the tarsi to the SOG are not depicted in this illustration. Gustatory projections from the wing synapse between frontal and mid leg projections to the TG. Frontal leg (red), Mid leg (blue), Hind leg (red), genital (GEN) chemosensory (yellow) send projections to distinct neuromeres in the thoracic ganglia.
CHAPTER 2
CHEMOSENSORY SNMP1 AND SNMP2 EXPRESSION IN ADULT
DROSOPHILA SENSORY CELLS

2.1 INTRODUCTION

Sensory neural membrane proteins (SNMPs) are a two-transmembrane insect chemosensory proteins belonging to the broadly functional family of Class B type I Scavenger Receptors (SR-B1). SNMPs were first described in Lepidoptera and found to associate with a sexual dimorphic pheromone specific sensilla in the antennae (85). The SNMPs split into two orthologues; SNMP1 and SNMP2 (86, 87). In Lepidoptera SNMP2 associated with the accessory support cells of pheromone sensory neurons in moth antennae(88). Subsequently, the SNMPs have been further characterized in other Lepidopteran species, uncovering a broader expression pattern for both SNMPs in chemosensory tissue in a number of moth pest species(89-91). The broadly expressed pattern and cell-type association suggest a broader functional role for SNMPs in chemoreception.
Both SNMPs are highly conserved within the insect holometabola (87). The holometaboulous insects emerged 300+ Mya and constitute one of the most proliferous groups of insects, making up 85% of all known species. Insect SNMPs are related to the vertebrate CD36/SR-B1, belonging to an ancient class of broadly functioning and expressed proteins extending through both deuterostome and protostome lineages (48, 68). In Drosophila this family is grouped amongst similar SR-B1 type proteins with broad functional roles (86, 87). SNMPs are homologous to the human CD36 class of SR-B1 receptors involved in the signaling, transport, and internalization of a versatile array of ligands (61, 92-94). In Drosophila SNMP1 has been characterized as an olfactory receptor (OR) cofactor required for normal response kinetics of the OR (25, 27, 95).

Drosophila SNMP1 was first noticed to express in antennal and fly body tissue. In the antennal, olfactory sensory neurons (OSN) expression was characterized at the lateral-distal population of sensilla projecting to trichoid specific glomeruli in the antennal lobe (AL). SNMP1 was also noted in non-neuronal support cells of the antennae and proboscis (25). This expression pattern was further expanded by the confirmation of SNMP1 and the first description of SNMP2 cDNA reported in head, leg and wing
tissue (87). This study characterizes the expression pattern of SNMPs in D. melanogaster using indirectly expressed membrane bound green fluorescent protein (mCD8::GFP) mediated by a UAS-Gal4 enhancer trap system. SNMP1 and SNMP2 exhibit broad chemosensory expression pattern between olfactory and gustatory tissue. Labeled sensory neural projections to the antennal lobes (AL), suboesophageal ganglion (SOG) and thoracic ganglia (TG) suggest expression in either neuronal or support cells in the majority of chemosensory cells. SNMP2 expression is broadly associated with gustatory sensory neurons (GSN) and is found in sexual dimorphic sensilla and GSN projections. The characterization of broadly expressed SNMPs between olfactory and gustatory tissue and expression in both neuronal and support cell, suggest a broader functional role for SNMPs, which goes in line with reported multiple functional roles for members of the SR-B1/CD36 protein family.

2.2 RESULTS

Rapid progress in understanding the chemosensory code in Drosophila has been possible by the relative simplicity in neuronal population and tractability of projections from the sensory organs, i.e. direct wiring of peripheral
chemosensory neurons to primary processing centers in the central nervous system (CNS) of fly. Many of these intellectual feats have been possible through an array of genetic tools available for the fruit fly. Indirect methods of labeling have greatly assisted the characterization of many chemosensory cells and proteins. The Gal4-UAS system uses a putative gene enhancer/promoter to drive the expression of the yeast transcriptional regulator Gal4 (enhancer – Gal4), which binds and tightly regulates the expression of a downstream UAS-reporter (96). This technique has the distinct advantage of facilitating the visualization of the complete temporal and spatial expression pattern of a gene in vivo. GFP may be specifically targeted to different cellular structures, in this study the CD8 transmembrane glycoprotein is fused to GFP directing the expression to the cellular membrane. It should be noted that even though the Gal4-UAS replicates endogenous expression accurately, occasionally the driver is not representative of the endogenous transcriptional control, affecting the expression pattern of the reporter.
Snmp1-Gal4 and Snmp2-Gal4 are expressed in chemosensory organs of Drosophila larvae and adults.

Enhancer driver sequences were 3.1 Kb upstream of the SNMP1 transcriptional start site and 5.0 Kb from the start ATG of SNMP2 (Fig2.1A), they were evaluated for any neighboring genes possibly affecting the expression pattern from the annotated FlyBase gene database (97). Transgenic flies positive for Snmp1-Gal4 and Snmp2-Gal4 were evaluated based on expression of UAS-mCD8::GFP in head, legs, and wings (Fig2.1B). In lepidoptera SNMP1 was seen to begin expression at about 40 % development with an increase shortly before full adult eclosion and several days post-eclosion (85). In Drosophila it was observed that mCD8::GFP expression turned on during pupal stage coinciding with expression of other chemosensory proteins. However, observations of third and fourth instar larvae showed expression of both SNMP1-Gal4 and SNMP2-Gal4 in chemosensory organs (Fig2.1C). The anterior larval head has three sensory organs: Dorsal organ (DO), Ventral Organ (VO), and Terminal Organ (TO) (98). Amongst these the principal olfactory organ, the DO, is composed of 6 peripheral sensilla comprising several sensory neurons and three support cells the majority of which are larval olfactory sensory neurons (OSN) specific. On the other
hand the TO is composed of several types of sensilla and is mostly innervated by gustatory sensory neurons (GSN) (47, 98, 99). Snmp1-Gal4 expression in the larvae exhibits broad internal expression in the DO and TO, since there are no distinguishable projections it is thought that these are localized in the support cell (Fig2.1C). Snmp2-Gal4 expression is seen in several neuronal projections tracking from the periphery to the TO and VO (Fig2.1C).

**SNMP1-Gal4 and SNMP2-Gal4 express in both olfactory and gustatory sensilla in the head of adult fly.**

SNMP expression in the adult fly is exhibited in the olfactory and gustatory organs of the head (Fig2.2A). The antennae and palps are the primary olfactory organs in *Drosophila*, both are composed of morphologically distinct and zonally organized sensilla (12, 23). *Drosophila* olfactory sensilla are characterized under three morphological subtypes, the sensilla coeloconic, basiconic, and trichoid, in addition to basiconic sensilla covering the maxillary palps reviewed Vosshall and Stocker, 2007). The main gustatory center in the head is mediated by morphologically distinct sensilla on the labial palps and in the pharynx; the external labellar sensilla are characterized as short (s), long (l), and intermediate (i). The s and l sensilla house 4 GSNs whereas the i house 2
GSNs the majority of one or groups of GSNs are accompanied by a mechanosensory neuron. All chemosensory neurons are accompanied by three accessory support cells; providing the intrasensillar lymph and expression of soluble chemosensory proteins (OBPs).

**SNMP expression is broadly associated with many sensilla types.**

The Snmp1-Gal4 and Snmp2-Gal4 expression is exhibited in both antennal and maxillary palp sensilla, with subtle differences between Snmp1 and Snmp2 expression patterns (Fig 2.2 B,C,F,G). Expression in the labellum exhibited noticeable differences between SNMPs. SNP1-Gal4 expression was observed between the pseudotrachae suggesting expression in the taste pegs (Fig2.2D). The taste pegs are a peculiar chemoreceptor unit in that each peg holds one mechanoreceptor and one chemoreceptor, stimulation of the taste pegs leads to an innate ingestion behavior (100). Snmp2-Gal4 expression in the labellum shows the labial nerve (LbN) projections to the ventral ends of both labial palps (Fig2.2H). Both Snmp1-Gal4 and Snmp2-Gal4 express in the pharyngeal cells the lateral sensory organ (LSO) and ventral cibrial sense organs (VCSO) with unclassified axonal processes travelling in parallel between these sensilla groups. (Fig2.1 A, E). The
pharyngeal sensilla are thought to function in assessing food quality before ingestion.

**SNMPs drive expression of mCD8::GFP in sensilla on the wing.**

S1-Gal4 and S2-Gal4 both express in gustatory sensilla along the anterior wing margin. The Drosophila wing has three rows of morphologically distinct sensilla of chemosensory and mechanosensory types on dorsal and ventral sides of the anterior wing margin (101). Snmp1-Gal4 displays expression in both chemosensory and in a number of mechanosensory hairs (Fig 2.3). Chemosensory sensilla are characterized by a slender-curved phenotype amongst stout-straight or peg-like mechanosensory sensilla (Fig2.3). Snmp1-Gal4 in gustatory sensilla on the wing margin are characterized by mCD8::GFP expression in equally slender-curved sensilla phenotype. Mechanosensory association of SNMP1 is characterized by expression of mCD8::GFP in the wing hinge, campaniform sensilla of the wing blade, and in a number of stout-straight sensilla in the wing margin (Fig1.2). Additionally, Snmp1-Gal4 is noticed in the proximal, distal radii, and wing blade of the haltere (Fig1.2D), the sensilla of the haltere are poorly characterized, the majority of which are mechanosensory (102). The proposed role for mechanosensory sensilla in
the halteres is thought to be involved in equilibrium and reflexes during flight (103). Snmp2-Gal4 associates with the majority of chemosensory sensilla in the wing margin (Fig1.2 D,E). In contrast with the Snmp1-Gal4 pattern, Snmp2 is strictly co-localized with chemosensory curved bristles on both dorsal and ventral rows of sensilla in the anterior wing margin. The functional role of chemosensory sensilla in the wing is not well characterized, however gustatory reflex behavior is established through the proboscis extension reflex (PER) mediated by a droplet of an attractive substance is placed on the wing margin.

**SNMP1-Gal4 and SNMP2-Gal4 show sexual dimorphic expression pattern from gustatory sensilla in the leg.**

Gustatory centers in Drosophila legs are involved in various chemosensory processes, tasting quality of a substrate for feeding or ovipositing and mating. Chemosensory projections between the taste centers of the proboscis and legs are functionally integrated, as in the wing GSNs, the leg GSNs respond physiologically and behaviorally (PER) upon delivery of an attractive substance (104). There are a number of mechanosensory sensilla distributed amongst approximately thirty gustatory sensilla, which house four GSNs and one mechanosensory with corresponding support cells, distributed in a stereotyped
pattern across all three sets of legs, with frontal legs having a few additional taste sensilla in the males (38, 65). GFP driven expression by SNMP1-Gal4 is exhibited in chemosensory centers of the Drosophila legs along with a number of mechanosensory sensilla in the femoral and tibial segments (Fig2.4B). Snmp2-Gal4 expression associates consistently with chemosensory cells along the tarsal and tibia segments of the leg (Fig1.4B). Sexual dimorphism is noticed in both snmp1 and snmp2 expression, indicating that both SNMPs possibly associate with pheromonal sensitive chemoreceptive sensilla in the leg (105) (Fig1.4C). Figures depict the prothoracic legs however expression was similar in all other pairs of legs.

**Snmp1-Gal4 and Snmp2-Gal4 characterization of neuronal expression of fly central nervous system (CNS)**

Snmp1-Gal4 and Snmp2-Gal4 enhancers facilitate an indirect characterization of sensory neural projections to both olfactory and gustatory centers of the Drosophila CNS, SNMP1-Gal4 expression is highly localized in the distal portion of the AL glomeruli, narrow projections to the SOG, and no noticeable neuronal projections to the TG (Fig 2.5A). The complete Snmp1-Gal4 expression pattern is suggestive of support cell localization of SNMP1 in
gustatory and mechanosensory sensilla and neuronal localization within pheromonal trichoid sensilla.

Snmp2-Gal4 exhibits chemosensory projections to the TG neuromeres from all appendages (Fig1.5A), in addition to showing broadly expressed neuronal projections from GSNs to the SOG (Fig1.5C). SNMP2-Gal4 neuronal expression is also represented in the AL glomeruli suggesting co-expression of SNMP2 within a number of OSNs. Additionally, there are projections travelling through the antennal mechanosensory and motor centers (AMMC). This area of mechanosensory signal relay is normally represented by acoustic sensitive neurons from Johnston’s Organ (JO) located in the distal segment from the chemosensory antennae, coincidentally the expression pattern for Gr68a, a male specific pheromone receptor is expressed in the AMMC (106). Snmp2-Gal4 expression in the second antennal segment housing the JO was not noticed (Fig2.1F; Fig2.5C).

Snmp1-Gal4 is found to co-express in pheromone sensitive trichoid, basiconic, intermediate, and coeloconic OSNs: Evidence from AL expression

Snmp1-Gal4 associates within an array of trichoid specific glomeruli, one intermediate, one basiconic, and one coeloconic glomerulus (Fig1.6). Through immunolabeling
using Anti-GFP and neuropil specific Anti-nc82 individual glomeruli and Snmp-Gal4 projections were identifiable. Expression pattern was observed in all GFP positive lines of Snmp1-Gal4 (N = 5 males + 7 females). Confocal z-stacks were used to image through the dense layers of glomeruli. Individual glomeruli were distinguished by separating stacks into anterior, medial, and posterior expression. Glomeruli were identified using previous OR specific mapping: (23, 80, 107, 108). Anterior stacks show Snmp1-Gal4 neural projections to trichoid DA3, DA4m, DA4l, VA1d, VA1v, and VA6 (Fig2.6). It is worth noting that the VA6 glomerulus receives inputs from both trichoid (Or67d) and basiconic (Or82a) (47, 109). Basiconic sensilla OSNs expressing the Or82a are narrowly tuned to reception of geranyl acetate, a green-leaf volatile that may act as a “pheromone” in some diptera species (110). OSNs expressing Or47b, projecting to VA1v have been described as evoking socio-sexual behaviors (111). DA1 and VA6 glomeruli receive projections from trichoid OSNs expressing Or67d, which are narrowly tuned receptor in the cis-vacennyl acetate (cVA) ligand recognition pathway(25, 110). Basiconic OSN projections to DA2 are responsive to geosmin, a bicyclic alcohol involved in alarming the fly of possible harmful microbes on a substrate (26). Trichoid projections
visualized from the medial stacks include DC1 and DL3, which receive projections from OSNs responsive to terpenes and cVA, OSNs projecting to DL3 are thought to be involved in inhibiting cVA elicited behaviors under high concentrations (112). The intermediate (previously trichoid) sensilla housing OSNs expressing Or83c project to DC3 glomerulus. Or83c is responsive to farnesol, a terpene and interacts with SNMP1 (27). The posterior AL stacks expression reports an unexpected expression of Snmp1-Gal4 in coeloconic sensilla, suggesting a possible role of SNMP1 and IRs. Ir84a OSNs sending projections to the VL2a form part of the fruitless (fru+) circuit, a master regulator of sexual behavior, along with Valv, and DA1 (113, 114). Ir84a along with other IRs respond to phenyacetaldehyde and phenylacetic acid, which are both yeast derived volatiles thought to promote male driven courtship (108, 115).

Snmp2-Gal4 expresses in a number of coeloconic and basiconic specific glomeruli.

The Snmp2-Gal4 AL projections are less confined to a specific sensillum type. Snmp2-Gal4 is reported in a number of coeloconic specific glomeruli, VL1, VC3m, VC3l, and VM4. Palp basiconics, VA4 and VM7 and the antennal basiconics, V and DL4 glomerulii (Fig2.7).
SNMP2-Gal4 axonal projections from GSNs are broadly represented in the SOG

Axonal projections from GSNs of the proboscis, labellar palps, maxillary palps, and tarsi send axonal afferents to the SOG. The SOG is less compartmentalized than the AL with no clear neuropil borders. The neuronal projections to the SOG are delineated by topographical representation rather than functional, making the identification of a gustatory map less intuitive (63, 105, 116). However, it is possible to characterize a pattern of neuronal expression using a UAS/Gal4 system. Projections from Snmp1-Gal4 and Snmp2-Gal4 were characterized based on the thorough mapping of the SOG projections from Miyazaki and Ito, 2010. The SOG forms part of the tritocerebrum and is located ventrally to the oesophagus and ALs, bordered by the AMMC neuropil.

Snmp1-Gal4 driven mCD8::GFP expression in the periphery associates with the gustatory taste pegs in the labellum, in addition to the pharyngeal sensilla. Snmp1:Gal4 expression in the SOG reports a narrow neuronal localization of Snmp1-Gal4 in the peripheral gustatory centers (Fig1.8A) Neuronal projections targeting central area of the SOG are characterized as posterior maxillary sensory (PMS) branch. The PMS is subdivided into 3
anatomically differentiated projections, PMS1, PMS2, and PMS3. Snmp1-Gal4 mCD8::GFP labeled GSNs send a narrow ipsilateral axonal projection to the ventral medial subarea zone (PMS3), from the ventral axonal projections from the labellum (49, 63, 116). The different PMS projections are believed to be involved in different combinations of bitter-associated stimuli, this was proposed through confirmation of Gr66a-Gal4 projections colocalize with the PMS 2 and 3 structures (116, 117).

Snmp2-Gal4 driven mCD8::GFP expression exhibits a broader projections pattern to the SOG. The medial stacks section of the SOG (Fig 1.8B) exhibit neuronal projections to PMS 1 – 4, which shares an expression profile to that characterized by certain GRs. Gr32a and Gr66a project to PMS 1 – 3 while Gr47a projects to PMS 2 – 3. Projections from Gr5a-Gal4 driver localize to the PMS4 area of the SOG (117). Anterior stacks show the two most lateral projections in the SOG forming the anterior maxillary sensory (AMS) branches (Fig2.5C). The medial taste pegs responsive to CO2 project to the AMS branches while the lateral sensilla in the labellar palps respond to water and send their projections to the PMS by way of the AMS branch (118). These projections are also representative of the pharyngeal inputs, the LSO and (63). The posterior
confocal stacks of the SOG report Snmp2-Gal4 driven mCD8::GFP across long dorsal axon projections, the anterior cerebrocervical fascicle (ACCF). The ACCFs are representative of projections from tarsal neurons expressing GR66a and possibly projections from pheromone receptor Gr32a. The SOG pattern from Snmp2-Gal4 exhibits localized branches reporting expression patterns from both attractive and aversive stimuli GSNs, in addition to CO2 and water vapor sensitive GRs (37, 117, 118). The patterns exhibited by Snmp2-Gal4 are characteristic of highly co-localized GR expression patterns, increasing the difficulty of associating this expression to a definite GSN function. It can be said with assertion that the SOG in Snmp2-Gal4 flies are receiving inputs from labellum, pharyngeal, and tarsal GSNs.

**Snmp2-Gal4 exhibits broad gustatory neuronal afferents from legs and wings.**

Snmp2-Gal4 mediated mCD8::GFP expression in GSNs axons from legs and wings arborize throughout the TG neuromeres (Fig1.6). Multiple axonal chemosensory projections from the frontal legs synapse onto the medial portion of the prothoracic neuromere (Fig 1.6 A,B). The male specific projections from the frontal legs cross over, while the females synapse towards the center, both send axonal
projections to the SOG (Fig1.6A,B). The sexual dimorphic wiring of these axons is thought to be indicative of pheromone sensitive GSNs in the tarsal segments of the front legs in males (119). Proximally from the prothoracic projections are the axonal projections from the wing, these are reported as a dense bundle of axons with no sex specific dimorphism (Fig1.5). Directly below the wing projections are the mesothoracic projections from the middle legs and the ventral most hind leg projections to metathoracic neuromeres. In Figure 1.5B, there are noticeable axonal projections running dorsally from the metathoracic projections, possibly indicating the integration of chemosensory signals from the legs or mechanosensory afferents (65). Chemosensation in the appendages from the body have been poorly characterized, in exception to certain afferents from the tarsal segment, where Gr66a GSNs are seen to project to the SOG, while Gr32a GSNs from the leg bypass the SOG and project behind the ALs (116). It is worth noting that Snmp2-Gal4 drives a broad gustatory pattern labeling the majority of ipsilateral GSN projections at once, this type of expression pattern is rarely depicted from any existing chemosensory GAL4/UAS studies.
2.4 DISCUSSION

Both the Snmp1 and Snmp2 enhancer/reporter constructs exhibit a broad but patterned expression, covering both olfactory and gustatory chemosensory modalities and cell types (neuronal or support cell). Snmp1-Gal4 associates at the periphery with all chemosensory specific tissues, exhibiting high levels of expression in the antennae, maxillary palps. The majority of the expression in the antennae is localized to the OSNs in trichoid with a narrow exception in basiconic and coeloconic. Snmp1-Gal4 expression in the pharyngeal neurons, LSO and VCSO do not report neuronal inputs to the SOG, however a narrow expression pattern is noted, presumably from labial GSNs involved in bitter-tasting. This projection was not noticed in the peripheral labellum expression, mCD8::GFP expression was heavily localized in support cells of the one to one chemosensory and mechanosensory cell taste pegs. Snmp1-Gal4 expression in legs and wings is mostly support cell as reported by the absence of mCD8::GFP expression axonal projections from these appendages. Interestingly, a number of cells in the Snmp1-Gal4 transgenic flies are of specific mechanosensory function including expression to the halteres.

46
Snmp2-Gal4 peripheral expression is similar to that of Snmp1-Gal4, however SNMP2 appears to express in the GSNs throughout the body. The antennal expression was representative of SNMP2 expression in a number of coeloconic and both antennal and palp basiconic sensilla. Additionally, Snmp2-Gal4 reported expression in the V glomerulus, which has an atypical expression of GRs in OSNs responding to CO2. Projections to the SOG also showed a number of gustatory projections including inputs from CO2 and water vapor sensing GSNs in the labellum (118), including additional projections co-localizing with sweet and bitter sensing GSNs found in labellum, pharynx, and tarsi. Projections to the TG were representative of SNMP2 expressing within a number of GSNs in the legs and wings, including GSNs in the prothoracic tarsi presumably involved in mating.

**Fine tuning the chemoreceptors**

The broad expression pattern for both SNMPs raises many questions towards understanding the role of these proteins in chemosensation. Previous studies have qualified the role for SNMP1 as an obligate co-receptor for the pheromone receptor Or67d. A very recent study identified SNMP1 in mediating the response and deactivating kinetics of Or83c. Or83c is classified as a farnesol specific OR expressing in
OSNs of intermediate type sensilla of the antennae (27). In this study SNMP1 is confirmed to co-expressed in both the DA1 and VA6 glomeruli receiving signals from Or67d (25, 110), DC3 glomerulus receiving projections from OSNs expressing Or83c, the geosmin sensitive basiconic Or56a projections to DA2. Additionally, from the literature other ORs co-expressing with SNMP1 are involved in reception of cuticular hydrocarbons or terpenes (Table2.1). Most ORs in Drosophila are broadly tuned, expanding the olfactory code from only sixty-two ORs (41, 47). However, signals that are crucial towards species success and fitness i.e. pheromonal cues would be expected to be specific and encoded by narrowly tuned receptors (8, 120). However, narrow tuning is possibly a result of the networking between chemosensory proteins in the OSN and not solely based on OR specificity (121) (122). Specificity in pheromone coding is thought to be a result of pheromone specific OBPs in the sensillar environment interacting with the receptor (123) (58). The OBP LUSH is responsible and necessary for Or67d specificity for cVA pheromone (95) (124). Or67d receptor kinetics is further modulated by the co-expression with SNMP1. It is possible that most ORs are broadly tuned however co-expression with OBPs and/or SNMP1 is sufficient to confer a "narrow-tuned" receptor complex.
The co-expression of SNMP1 with the phenylacetic acid sensitive Ir84a, suggested from expression in the VL2a glomerulus poses an interesting question towards the role of SNMP1 in OR specificity. Most IRs are broadly tuned receptors but specialize in the perception of small amines and acids (125). It is worth acknowledging that VL2a is part of the fru+ circuit conferring male specific neural development in odorant coding and might be necessary to carry the signal of phenylacetic acid for downstream integration with pheromonal signals (115). SNMP1 could possibly have a role in signal integration by temporal modulation of receptor kinetics, suggesting a role for sensory integration of different odor induced signals.

SNMP2 as a fatty acid transporter

SNMP2 is likely co-expressing with the majority of GSNs in labellum, legs, and wings, also expressing in a number of antennal and palp OSNs. Interestingly, SNMP2 is possibly co-expressing with narrowly tuned atypical Gr21a, the antennal CO2 receptors, in addition to the labellar CO2 sensitive GSNs. Snmp2-Gal4 expression in the SOG and TG, suggest a functional role in the gustatory code. Gustatory receptors are broadly functioning with GR specificity of different tastants determined by co-expression with ubiquitous co-factors for bitter or sweet perception.
Recently Drosophila gustation has been subject to increasing studies in other taste categories as osmolarity, carbonation, and lipid content. Drosophila has been found to have a robust feeding response to a range of fatty acids (60). SNMP2 forms part of the CD36 family of vertebrate fatty acid transporters, which have also been characterized to play a role in vertebrate gustation. CD36 is believed to be responsible for mammalian perception of fats in a dietary substrate. Mice with CD36 knock-out are not able to qualify food on the level of fat content. The taste for fatty acid in Drosophila could possibly be a function of the broadly expressed presence of SNMP2 in labellar GSNs.

The role of SNMP2 will become clearer as the role of distinct GSNs in the head and body become more resolved. Supposing SNMP2 function is similar to that of SNMP1 it is possible for a role of fine-tuning the GR signal. The sexual dimorphic expression of SNMP2 in tarsal GSNs and projections in the prothoracic TG suggest the likely co-expression of SNMP2 with pheromone receptors, for example Gr68a, a contact pheromone receptor in Drosophila males. This pathway has a similar protein organization to the cVA pathway in that certain OBPs have been implicated in tuning the pheromone response, it would be of great interest to
see if SNMP2 has a similar role as SNMP1 in GR specificity and functionality in this pathway.

Since the discovery of SNMPs it has been clear that certain SNMPs express in the chemosensory support cells. The studied functional role for SNMP1 and speculated functionality of SNMP2 in the sensory neurons are difficult to interpret in relation to expression in the support cell. However, the possibility of SNMPs eliciting CD36-like functionality, suggest that SNMPs may also have rather diverse functions in chemosensation. A possible role for SNMPs in the support cells is signal internalization. CD36 has been implicated in internalization of oxidized low-density lipoproteins (LDL), responsible for the formation of arterial plaques in mice and humans (62, 126). It would be of great interest to understand if SNMPs are involved in clearing the sensillar space or internalizing olfactory or gustatory ligands, suggesting a possible mediator in second messenger pathways.
2.5 Experimental Procedure

Constructing SNMP1 and SNMP2 driver-Gal4

DmCG7000 (SNMP1) upstream regions were isolated from BAC clone BACR10M16; DmCG7422 (SNMP2) upstream regions were isolated from BAC clone BACR03I22; both clones were obtained from the BACPAC Resources Center of Children's Hospital Oakland Research Institute (CHORI) (RPCI-98 Drosophila melanogaster BAC Library, clone RPCI-98 10.M.16). BAC colonies were prepared for genomic isolation following CHORI protocols (127). BAC clones were obtained in stabs of DH10 E.coli, and streaked on Luria-Bertani (LB) agar plates containing 20 μg/ml chloramphenicol. Individual colonies were cultured overnight in 3 ml LB media containing 20 μg/ml chloramphenicol. Cultures were centrifuged and bacterial pellets resuspended (0.3 ml of 50 mM Tris, pH 8, 10 mM EDTA, 100 μg/ml RNase A, 4°C), lysed (0.3 ml of 0.2N NaOH, 1% SDS, room temperature, 25°C) and neutralized (0.3 ml of 3M KOAc, pH 5.5, 4°C). Lysates were centrifuged and supernatants precipitated by addition of 0.8 ml cold isopropanol (5 min. incubation on ice followed by centrifugation). DNA pellets were washed in 70% ethanol, air dried, resuspended in 40 μl Tris-EDTA (TE) buffer and stored at -20°C.
PCR primers were designed from gene scaffold AE003733 for isolation and sequencing the 5’ untranscribed upstream SNMP1 nucleotide region. Reverse primer, 5’-GGATCCGAGCTGAAGACCAGGGCTTTTC - 3’, was designed with BamHI (GGATCC) restriction endonuclease site. Reverse primer was designed to a site whose 5’ boundary was 68 bp upstream of the start ATG and 212 bp downstream of the nearest predicted consensus TATA box. Two forward primers were designed to sites whose 5’ boundaries were 3071 bp (5’-ACTAGTCACTTGGCTTTCCATCGACTACG - 3’ and 7009 bp 5’-ACTAGTAGCCATATCGCTGGGAGAAC - 3’ upstream from the 3’ primer site both with SpeI (ACTAGT) restriction endonuclease sites; these two regions are subsequently referred to as SNMP1 (3.1 kb) and SNMP1 (7.0 kb).

A 50 ul PCR analysis was used to amplify SNMP1(3.1 kb) and SNMP1(7.0 kb) from Qiagen mini-prep BAC DNA (1 ul) using PlatinumTaq HiFi (Invitrogen) in a BioRad thermocycler. Amplified DNA was cloned using a TOPO-TA cloning kit pcrII-TOPO vector (Invitrogen) and transformed into TOP10 bacterial cells (Invitrogen).

Primers for the amplification of presumed 5.0 kB promoter for SNMP2 (CG7422) were designed using Primer3 software (http://biotools.umassmed.edu/bioapps/primer3_www.cgi): (Forward: 5’ – ACTAGTGCCTATGCGTATTTTTTGTGA – 3’, SpeI)
and (Reverse: 5’ — GGATCCAGATACGATCGGATGGATTG – 3’, BamHI). These amplify exactly 5000 bp and end 6 nucleotides upstream of the start ATG. PCR product was amplified from BAC clone and blunt-end ligated into cloning vector as described above for CG7000.

P-element transformations, SNMP1 (3.1 kb) and SNMP1 (7.0 kb) were sub-cloned into VP16 transgenic vector. VP16 was derived from pCASPER4 and contains a GAL4 region with an adjacent cloning site suitable for inserting a driver sequence of interest (128). The VP16 vector was constructed by Makoto Makishima and David Manglesdorf and generously provided by Dean Smith (Southwestern Medical Center, Dallas). TOPO::SNMP1 (3.1kb) / (7.0kb), TOPO::SNMP2 (5.0 kb), and VP16 vector were digested with SpeI and BamHI restriction endonucleases and isolated from 4.0 % TAE agarose gel (Bio101 Geneclene Turbo Kit). Isolated SNMP1 and SNMP2 DNA was de-phosphorylated using Shrimp Alkaline Phosphatase (SAP; Promega) and ligated to VP16 (T4 DNA Ligase, Invitrogen) DNA was transformed into TOP10 cells and plasmid / insert ends were sequenced for confirmation of successful ligation.
Fly rearing and genetics.

*D. melanogaster* colonies are reared with a cornmeal, molasses, and yeast medium with 10% p-Hydroxy-benzoic acid methyl ester (Tegosept) dissolved in 95% ethanol at 23°C (16h:8h L:D). *Drosophila* strain w118 embryos were injected for p-element transformation at Model System Genomics of Duke University (wild-type except for mutant white eyes). Plasmid DNA preps for injection were prepared using Qiagen midi-prep, eluted in dH2O and provided to the Duke facility at >1 mg/ml. Resulting adult flies were backcrossed into W1118, and red eyed progeny crossed with yw; Cyo; Sb,TM3/ T(2;3)ap[Xa],ap[Xa] (Bloomington Stock # 2475) to determine chromosomal location: male red - eyed flies were crossed with #2475 females and screened for both loss of apterous [Ap] phenotypic marker and retention of red eye; positive males were re-crossed with #2475 virgin females and screened for loss of either curlyO (Cyo) and/or stubble (Sb,TM3) and retention of red eye to determine chromosomal location of the P-element insertion. One line each of SNMP1 (3.1 kb) and SNMP1 (7.0 kb) was preserved carrying P-element insertions on both chromosome 2 and 3; these lines identified as homozygous by loss of balancer. Single P-element insertions (on chromosome 2 or 3) were maintained against appropriate balancers. Males of the
resulting lines were crossed with virgin female yw,UAS-mCD8::GFP flies (Bloomington Stock # 1801) for expression analysis. Three independent driver lines were recovered for SNMP2 (5.0 kb) promoter, two located on chromosome 3 and one located on chromosome 2. Only one chromosome 3 insertion was used for imaging, as it provided the brightest GFP fluorescence, noting all three lines drove apparent identical expression of marker (insertion location may affect expression secondarily). Genotypes of imaged flies are as follows. For DmSNMP1:

yw, UAS:cd8GFP / X or Y; Cyo or pin/ +; SNMP1:GAL4/+.

For DmSNMP2: yw, UAS:cd8GFP/ X or Y; Cyo or pin/ +; SNMP2:GAL4/+.

**Whole body imaging.**

Three to five day old adult flies were fixed in 4.0% paraformaldehyde diluted in 1x PBS - (0.3 %) TritonX (PBS-T) for 15 minutes at RT and washed 3x for 15 min each with 1x PBS-T. Whole bodies and dissected heads, antennae, labellum, legs, and wings were mounted on 1.0 μm cover-slip slides with Fluormount-G (Southern Biotech) or VectaShield (Vector Laboratories). Third instar larvae were washed in 1x PBS-TX and held in a ~ 60°C water bath to promote larval body elongation before 15 minute fixation in 4.0%
paraformaldehyde PBST and further washed 3x in PBS-T. Larvae were mounted on slides under 1.0 μm cover-slips with Fluormount-G.

**Immunohistochemistry of Brain and TG.**

Brains and TG were prepared as described in (Wu and Luo, 2006). Flies were anesthetized by cooling, then immersed in cold 1x PBS-T (0.3 %) and cuticle in the head and thorax was broken using microforceps exposing the neural tissue. Flies were subsequently fixed at RT for 20 minutes in 4.0% paraformaldehyde in 1x PBS-T followed by a brief wash in PBS-T, brains and TG were dissected, washed 3x in (0.3%) PBS-T for 5 min, immersed in 5.0 % normal goat serum (NGS) in (0.3%) PBS-T and nutated for ~ 1 hour at RT. Primary antibody one wash in 0.3% PBST, brains and TGs were treated with primary antibody was added and incubated for 48 hours with gentle nutation at 4.0°C (Rabbit anti-GFP [Invitrogen], 1:2000; mouse monoclonal anti-nc82 [DSHB], 1:40). Tissue was gently washed in (0.3 %) PBST for five washes at 5 minutes each, tissue was subsequently treated with fluorescent conjugated secondary antibody for 48 hours with gentle nutation at 4.0°C (Alexa 488 conjugate Goat anti-rabbit, 1:2000; Alexa 568 conjugated Goat anti-mouse [Invitrogen], 1:2000). After five washes with (0.3%) PBST, neural tissue was immersed in a small volume (~10 μl) of
Slowfade Gold (Invitrogen) and mounted between two No. 1 (1.0 μm) coverslips using a manually constructed stage with No. 1 coverslips as spacers.

**Confocal microscopy**

Confocal images were obtained using a Nikon Eclipse TE300 fitted with a BioRad MRC 1024ES laser scanning microscope. Fluorescent images were obtained using an Olympus BX60 equipped with an imaging Micropublisher (32 0028a-118) CCD camera. All confocal images were collected within 24 hours of mounting. Confocal image stacks were processed between 0.5 – 2 μm increments. Images were processed using Image J and Adobe Photoshop CS2 software.
Figure 2.1: Snmp1-Gal4 and Snmp2-Gal4 expression in larvae and broad expression in adult

A. Design of Snmp1 and Snmp2 Gal4 driver analyzed through annotated FlyBase sequence.  B. Illustration of olfactory and gustatory cells positive for mCD8::GFP in adult fly, compared to presence of cDNA from head (H), leg (L), and wing (W), cDNA gel image from (Vogt et al., 2009).  C. Expression of mCD8::GFP from S1-Gal4 and S2-Gal4 third instar larvae.  Scale bar: 50 μm
Figure 2.2: Peripheral expression of GFP in *Drosophila* adult head.

Fluorescence microscopy images (10x) of S1-Gal4 and S2-Gal4 Adult head. Arrows point to the pharyngeal GSNs (A,E). S1-4 (B,C,D) and S2-4 (F,G,H) confocal stack images of dissected tissue at 60x. mCD8::GFP is in green, magenta is background fluorescence form the cuticle.
Figure 2.3: Peripheral expression of GFP in *Drosophila* adult wing and haltere.

Fluorescence microscopy images (10x) of S1-Gal4 and S2-Gal4 adult female wing (A). 40x fluorescence microscopy of female adult wing margin, arrowheads point to chemosensory sensilla (B,D). White arrows point to mechanosensory sensilla on the wing margin (B). 40x image of GFP expression in the halters of S1-Gal4 female fly (C).
Figure 2.4: Peripheral expression of GFP in *Drosophila* adult prothoracic legs and tarsi.

Fluorescence microscopy image composite (40x) of S2-Gal4 and S1-Gal4 adult leg (A). Arrowheads point to chemosensory sensilla. 40x fluorescence images of S2-Gal4 and S1-Gal4. Panels represent a comparison of male and female tarsi for both SNMP1 (bottom panels) SNMP2 (top panels (B)).
Figure 2.5: Immunohistochemistry showing neuronal projections of SNMP1 and SNMP2 in Drosophila CNS

Composite of merged confocal stacks (20x) of S1-Gal4 and S2-Gal4 mCD8::GFP (A). Brain showing expression to the AL (ant) and SOG (84). Asterisk points to projections from the labellum and TG. 20x confocal stacks of TG show frontal legs/prothoracic (FL), wing, mid leg/mesothoracic (ML), and hind leg/metathoracic (HL). 60x confocal stacks of S1-Gal4 AL and SOG expression, bottom panel shows S2-Gal4 olfactory projections (olf) to AL and gustatory projections to the SOG (gust).
Figure 2.6: Immunohistochemistry showing neuronal projections to the glomeruli of the AL in S1-Gal4 and S2-Gal4

100x confocal stack of S1-Gal4 (A) and S2-Gal4 (D) antennal lobe. S1-Gal4 stacks were separated into merged images of anterior, medial, and posterior projections (B). Illustration of identified glomeruli expressing GFP of S1-Gal4 (C) and S2-Gal4 (E).
Figure 2.7: Immunohistochemistry showing neuronal projections to the SOG of S1-Gal4 and S2-Gal4 adult flies.

100x confocal stack of S1-Gal4 SOG(A), illustration shows identified projections in to the PMS3. Anterior (C), medial (B), and posterior (D) merged stacks of S2-Gal4 SOG. Illustration of projections next to confocal images as identified in Miyazaki and Touhara, 2011.
Figure 2.8: Immunohistochemistry showing neuronal projections to the TG in S2-Gal4 adult male and female.

60x composite of merged confocal stacks of S2-Gal4 neuronal projections in 3-day post-eclosion adult male (A) and female (B). Lateral image of S2-4 projections to the TG of female. Arrows in (A) and (B) point to projections to the SOG. Arrow in (A) identifies the sexual dimorphic crossing over of prothoracic projections in male. Arrows in female TG point to projections from wing, mesothoracic, and metathoracic.
### Table 2.1 SNMP1 expression in antennal lobe

<table>
<thead>
<tr>
<th>Sensillum</th>
<th>Type</th>
<th>Receptor</th>
<th>Glomeruli</th>
<th>Ligand</th>
<th>Receptor Function and Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>at2</td>
<td>Trichoid</td>
<td>Or23a</td>
<td>DA3</td>
<td>alcohols</td>
<td></td>
</tr>
<tr>
<td>at3</td>
<td>Trichoid</td>
<td>Or43a</td>
<td>DA41</td>
<td>alcohols, aldehydes</td>
<td>nd</td>
</tr>
<tr>
<td>at3</td>
<td>Trichoid</td>
<td>Or2a</td>
<td>DA4m</td>
<td>esters, alcohols</td>
<td>nd</td>
</tr>
<tr>
<td>at4</td>
<td>Trichoid</td>
<td>Or88a</td>
<td>VA1d</td>
<td>pyrrolidine</td>
<td>nd</td>
</tr>
<tr>
<td>at4</td>
<td>Trichoid</td>
<td>Or47b</td>
<td>VA1v</td>
<td>cuticular hydrocarbons</td>
<td>Evoke socio-sexual behaviors (111)</td>
</tr>
<tr>
<td>Basiconic</td>
<td>Or82a</td>
<td>VA6</td>
<td></td>
<td>geranyl acetate (terpene)</td>
<td>Detection of green leaf volatile may act as a pheromone.</td>
</tr>
<tr>
<td>at3</td>
<td>Trichoid</td>
<td>Or19a, 19b</td>
<td>DC1</td>
<td>pentyl acetate, 1-hexanol, 1-octen-2-ol, 2-heptanone</td>
<td>Detection of terpenes promoting oviposition (26)</td>
</tr>
<tr>
<td>at4</td>
<td>Trichoid</td>
<td>Or65a, 65b, 65c</td>
<td>DL3</td>
<td>pyrrolidine, esters</td>
<td>Reduce aggresion through cVA pathway (112)</td>
</tr>
<tr>
<td>ab10</td>
<td>Basiconic</td>
<td>Or49a</td>
<td>DL4</td>
<td>2-heptanone, alcohols</td>
<td>Increased expression when females exposed to mating song. (129)</td>
</tr>
<tr>
<td>at1</td>
<td>Trichoid</td>
<td>Or67d</td>
<td>DA1</td>
<td>cuticular hydrocarbons (cVA)</td>
<td>Elicit social aggregation and aggression. Functional role with SNMP1 (25)</td>
</tr>
<tr>
<td>ac4</td>
<td>Coeloconic</td>
<td>Ir84a</td>
<td>VL2a</td>
<td>Phenylethylamine, Phenylacetaldehyde, Phenylacetic acid</td>
<td>Reception of phenylacetic acid, enhances male courtship (115)</td>
</tr>
</tbody>
</table>
CHAPTER 3

CHARACTERIZATION OF SNMP1 GENE REGULATION BY ENHANCER DRIVER DISSECTION

3.1 INTRODUCTION

Expansive sections of eukaryotic non-coded DNA within the vicinity of the transcribed gene are encrypted with nucleotide cis-regulatory elements (CREs). These elements are made up of diversely arranged enhancers, silencers, and insulator sequences at varying positions and binding affinities mediating transcription factor (TF) recruitment (130), (130-133). The cis-regulatory code is the basis of gene-gene networking instructing the precise pattern, timing and concentration of TFs and other genes during development (134). A long standing question in the quest to understand the gene regulatory network asks, how are multiple related and un-related genes spatially and functionally organized through regulatory processes? The Drosophila chemosensory system offers a highly tractable system providing a complex pattern of co-expressed non-orthologous genes, conferring a range of signal specificity
in the context of sensory organ and cell type (68, 125, 135, 136). Chemosensory proteins in Drosophila express in a highly stereotyped patterns, suggesting a regulatory code instructing the sensory code. This presents the possibility of CRE mediated instruction in directing one gene to a subset of neurons out of many possible choices. Enhancer - trap studies on antennal and palp ORs uncovered specific expression of OR genes was dependent on a combinatorial code between enhancer and repressors elements suggesting a highly specific regulatory code instructing an OR gene to a specific cell or subset of cells (83, 137).

Drosophila chemosensory system is divided between discrete sensory “organs”. Olfaction is mediated through structurally distinct sensilla found on the third antennal segment and maxillary palps. Sensilla on the antennal third segment are distributed in a bilateral and stereotyped pattern with large basiconic sensilla clustered at the medial-proximal side and trichoid clustered at the lateral-distal edge of the antennae. Small basiconic and coeloconic sensilla are interspersed in the middle region of the antennae. The maxillary palp is less diverse only expressing large basiconica. Each sensilla houses 1 -2 OSNs totalling, ~ 1200 OSNs for each antenna (12, 14).
Taste is broadly represented in the fly. The labellum consists of two palps with 31 stereotypically distributed sensilla with bilateral symmetry. Much like the olfactory organs, the labellar palps are covered with sensilla of different morphological types; short (s), intermediate (i), and long bristles (l), housing between 2-4 gustatory sensory neurons (GSNs) and one mechanosensory cell. The pharynx has three separate bilaterally distributed cluster of taste cells, the labral sense organ (LSO), dorsal and ventral cibarial sense organs (DCSO and VCSO). Taste is neuronally represented in the suboesophageal ganglion (SOG) located ventrally to the antennal lobes. The SOG is the relay station of axonal projections from pharynx, labellum and tarsal segment of the prothoracic legs (37, 44). The chemosensory sensilla of the legs and wings express in a stereotypical fashion and are found interspersed with mechanosensory sensilla. Axonal projections from the legs and wings synapse on to the thoracic ganglia (TG) (14). All sensory sensilla house chemosensory neurons, surrounded by three support cells. These cells express the biochemical network of proteins that functionally characterize the sensory cell: Odorant binding protein (OBP), Olfactory, Ionotropic or Gustatory Receptors (ORs, IRs, GRs), and SNMPs.
SNMPs are chemosensory proteins, which are highly conserved throughout the holometabola (138). Unlike the OBP s and ORs, SNMP does not seem to have arisen through convergent evolution, belonging to a gene lineage of Scavenger Receptors Class B type 1 (SR-B1), which are broadly expressed family of fatty acid transporters in vertebrates (86, 138). The insect SNMPs were first identified in Lepidoptera where SNMP1 and its orthologue SNMP2 were identified in antennal pheromone sensilla, expressing in neuronal and support cell respectfully (88). Enhancer trap characterization of the Drosophila melanogaster SNMPs reported peripheral expression of SNMP1 in olfactory (antennae and palps) and gustatory tissue (proboscis, legs, and wings). The neuronal projections from SNMP1-Gal4 revealed possible co-expression with ORs in the majority of trichoid OSNs. However it was seen in other sensilla types, trichoid expression was vastly superior in glomeruli expression. Gustatory and mechanosensory neuronal projections were minimal and the majority was characterized as support cell specific. This study proposes that the projection patterns exhibited by the Snmp1-Gal4 enhancer/trap analysis are under cis-regulatory control and co-expression with non-orthologous
genes is a result of a conserved network of regulatory elements.

Likely functional CRES were identified using a comparative approach. The identified CRES were used as guides for dissection of the SNMP1 upstream region. A broadly characterized 400 bp core promoter (S1.4-GFP) was identified within the vicinity of the transcriptional start site (tss). Reporter GFP expression was increased in the neuronal projections from gustatory and mechanosensory sensilla, suggesting an upstream repressor element possibly controlling cell type expression. The S1.4-GFP core promoter driver additionally abolished neuronal expression in a number of trichoid projections. Closer inspection of the core promoter element using motif predictive algorithms identified a highly conserved element amongst Drosophila species orthologous and non-orthologous co-expressed ORs. Site-directed mutagenesis of a 7-bp motif was responsible for further loss of trichoid OSN expression.

This study suggests broad snmp1 expression is under control of a complex, possibly diverse, cis - regulatory landscape consisting of specific enhancers for subsets of trichoid OSNs and a repressor element involved in sensory neuron versus support cell instructions. This work also identified a motif similarly found to be overrepresented
amongst trichoid specific ORs (137), suggesting regulatory elements used to direct functionally related chemosensory proteins to the appropriate sensory cell.

3.2 RESULTS

Many factors contribute to the spatial and temporal expression of genes. Environmentally regulated morphogenic changes lead to methylated and acetylated enzymatic reactions unwinding DNA and revealing an embedded nucleic acid code in the non-coding regions of DNA flanking a gene. Broadly expressed chemosensory SNMP1 in cell type (neuronal / support cell) and functionally distinct sensory cell modalities (olfactory / gustatory) hint at an intricate regulatory landscape flanking the coding region of the snmp gene. A reporter/enhancer methodology of qualifying expression using an upstream enhancer driven Gal4 indirectly expressing mCD8::GFP in regions spatially and temporally coded by the enhancer code is implemented in the characterization of regulatory elements. Based on the assumption that functional regulatory non-coding regions hold an analogous level of conservation compared to coding regions of DNA we compared ~2 kb upstream of the start ATG of the D. melanogaster snmp1 sequence and aligned it to the
D. pseudoobscura snmp1 orthologous sequences, identifying conserved ~ 20 bp non-coding regions (139, 140).

Comparative analysis of orthologous 5' upstream SNMP1 in D. melanogaster and D. pseudoobscura

Non-coding nucleotide sequences in the vicinity of the coded gene are involved in recruiting the appropriate machinery in spatial and temporal gene transcription. CREs are primarily identified by a comparative approach, assuming that most CREs are under selective pressures, hence demonstrating a high level of sequence conservation between two orthologous sequences. D. pseudoobscura (D.pse) is one of the closest but highly diverged suborders from D. melanogaster (D.mel), showing an evolutionary split from the melanogaster and obscura group at ~ 65 – 43 Mya (141)(Fig3.1A). An arbitrarily chosen 2 kb upstream sequence from annotated D.mel snmp1 including the transcriptional start site (Fig3.1B) was compared to the D.pse 5’ upstream non-coded orthologous snmp1 sequence. The two gene comparative Family Relations software (142) algorithm compares a window of fixed size of one sequence against an orthologous sequence. These windows were analyzed at an 80-95 % threshold level. The threshold level allows a certain number of mismatches per 20-bp. Pairwise comparative analysis identified discrete regions
of 21 – 56 bp at 90 % threshold of conservation in snmp1 (Fig3.1B). These regions hint at possible CREs and were used as guides for primer design for 2.0 Kb snmp1 driver dissection (Fig3.1C). The fractionated upstream sequences were separately cloned into a Gal4::VP16 vector for qualified examination of mCD8::GFP reporter expression. Out of the eight constructs four (S1-1, S1-2, S1-3, and S1-4) transgenic lines showed characteristic peripheral expression previously characterized in Chapter 2. A 400 bp core promoter, S1.4-Gal4, with relatively conserved peripheral expression was identified. Other dissected driver sequences, S1.5-Gal4 to S1.8-Gal4 did not show any GFP reporter expression. This is most likely a result of missing the basal transcriptional sequences from the core promoter.

**Analysis of neuronal projections from dissected Snmp1::Gal4.**

Dissection of S1-Gal4 upstream elements report possible regulatory control through repressor and enhancer mechanisms, controlling the characteristic expression of Snmp1-Gal4 in glomeruli expression of the AL. GSN projections to the primary gustatory center (PGC) in SOG (63) and support cell expression of mCD8::GFP positive sensilla in chemosensory and mechanosensory cells of the
legs and wings (Fig2.2B). Neuronal projections to the fly CNS (Brain and TG) from S1.1-Gal4 (1.95 Kb), S1.2-Gal4 (1.56 Kb), and S1.3-Gal4 (0.86 Kb) was similar to the S1.0-Gal4 (3.10 Kb) driver previously characterized in Chapter 2 (Fig 2.5). The core promoter minimal driver, S1.4-Gal4 (0.40 Kb) did exhibit visually characterized significant differences in neuronal projection pattern compared to S1.1-Gal4 driver. (Fig3.2D). The AL of S1.4-Gal4 exhibited expression of trichoid DA3, DA4m, DA4l, DL3, VA1v and VA1d. Expression of the Or67d specific DA1 and VA6 were noticeably missing, including the coeloconic VL2a. Expression to the coeloconic VM4 is reported in the S1.4-Gal4 driver, this is one of the glomeruli targets of SNMP2-Gal4 expression. The SOG projections are slightly more pronounced, reporting expression to the posterior maxillary sensory (PMS) 2 and 3 zones in the PGC representative of Gr47a GSNs from the labellum (Fig2.2D) (Miyazaki and Ito, 2010). Uncharacteristic of chemosensory neuronal projections in the TG seem to represent mechanosensory neural expression. The expression pattern of intersegmental processes running from prothoracic to metathoracic in the TG most resembles a previously characterized expression of femoral campaniform sensilla in the Dipteran Drosophila relative Phormia (Merritt and
The reported expression in the minimal core promoter element hints to possible CRE enhancers between S1.3–Gal4 and S1.4-Gal4 controlling spatial expression of DA1 and VA6 trichoid OSNs and coeloconic VL2a and a repressor for gustatory and mechanosensory projections from labellum and leg respectfully.

**Identification of highly conserved and represented motif in S1-4::Gal4 core promoter driver.**

A difficulty in CRE motif discovery is the increasing signal to noise ratio as sequences move away from the core promoter (Stark et al., 2007; Down et al., 2007; Berendzen et al., 2006). Therefore, the highly conserved core promoter region (Ohler et al., 2002) was implemented in discovery of motifs without a priori knowledge of a known transcription factor motif involved for the S1.4-Gal4 expression. The D. melanogaster minimal core promoter element was compared to 2.00 Kb upstream of the Drosophila orthologous snmp1 coding region and the presumably co-expressed OR upstream sequences with the purpose of identifying conserved and over-represented elements. Using the oligo-analysis program from the RSAT suite of repetitive elemenent discovery algorithms (Thomas-Chollier et al., 2012), under a background Markov model (MM = 3) calculated from the input sequence for optimal
identification of hexanucleotides, returned motif hits for four different conserved motifs in S1.4-Gal4 and related sequences. The GCAATTA motif was chosen based on location within comparatively deduced conserved elements, identified previously (Fig3.1B), local overrepresentation, and the conserved nature between orthologous and co-expressed genes (Fig2.3A,B).

**Mutagenesis of GCAATTA motif suggests it is a likely enhancer for a number of AL glomeruli.**

Using site-directed mutagenesis targeting the GCAATTA motif allowed the insertion of an AvrII restriction endonuclease site. Dmel transgenic lines reporting mCD8::GFP expression driven by S1.4-Gal4 with mutated motif element, (S1.4AvrII-Gal4) was characterized. Robust difference in projection pattern was noticed in the AL glomeruli. S1.4AvrII-Gal4 exhibited expression in the dorsal glomeruli: DA2, DA3, DA4l, DA4m, and DL3. Expression to the basiconic DA2 was “rescued” in the S1.4AvrII-Gal4 compared to the S1.4-Gal4 (Fig3.5B,C). The expression pattern correlated with the mapping of the GCAATTA element to the comparatively analyzed OR sequences expressing in OSNs projecting to DA2, DA3, DA4l, suggesting that most of these ORs are under control of a different enhancer.
3.3 DISCUSSION

Projections from chemosensory neurons previously characterized from the S1.0-Gal4 (3.10Kb) driver showed GFP expression from OSNs tracking to pheromone pathways in the antennal lobe, and GSNS to the labellar PMS region of the SOG. Support cell expression was suggested to explain the absence of neuronal projections to the SOG and to the TG. The expression profile from the majority of dissected driver constructs showed a similar expression, with the exception of the minimal core promoter driver S1.4-Gal4.

Driver constructs S1.1-Gal4 to S1.4-Gal4 were positive for reporter mCD8::GFP expression in the chemosensory periphery. Even the S1.4-Gal4 expression in chemosensory sensilla from the head, legs, and wings was conserved, suggesting that this broad expression pattern is instructed by regulatory elements housed in an ~350 bp region adjacent to the tss. Motif discovery algorithms were able to show four hexanucleotide elements within the core promoter element sequence (not shown). Further dissection of overlapping sequences would be necessary to find the possibility of sensory epithelia specific elements.

Projections from chemosensory neurons from the S1.0-Gal4 (3.5Kb) driver showed GFP showed expression in the glomeruli in a number of trichoid sensilla, one basiconic,
one intermediate, and one coeloconic specific glomerulus. The minimal core promoter driver S1.4 – Gal4 (0.40kb) reports a loss of projections to a number of regions in the AL. Antennal basiconic expression seen in the full driver is abolished in the core promoter expression. Neuronal expression to the Or67d glomeruli and coeloconic VL2a was abolished. It would be of interest to analyze if the early wiring of these OSNs to the AL is under a similar TF developmental pathway. However, previous study detailing the role of certain TFs in OR regulatory gene expression deduced that some TFs may act as both enhancer and repressors and this may be a result of a combinatorial effect of repeated elements for the same TF, demonstrating a CRE positional factor in CRE function (Miller and Carlson, 2010; Jafari et al., 2012). The increased neuronal expression to the SOG and mechanosensory projections to the TG suggest an upstream repressor element involved in differentiating SNMP1 localization within neuronal or support cells. The loss of expression in certain olfactory glomeruli could also be an issue of either an enhancer directing neuronal expression or a repressor involved in choosing between support cell or sensory neuron expression. Detailed in-situ analysis of the SNMP1 expression pattern in the antennae should be
attempted in the future to confirm this theory. Motif discovery algorithms accurately identified an enhancer element controlling the expression of certain trichoid ORs. A previous study found the same GCAATTA element functioning as both an enhancer and repressor element (137). This study confirms that SNMP1 expression in certain trichoid OSNs are under a similar regulatory control possibly through transposition or crossing-over of non-orthologous gene network of regulatory elements.

3.4 EXPERIMENTAL PROCEDURE

Plasmid Construction

Construction of enhancer / driver insert sequences. SNMP1 upstream enhancer fragments were amplified by PCR from BAC library, BACR10M16. Primers were designed based on short segments of high conservation identified through evolutionary comparative analysis (Fig 3.1B). Figure 3.1C, nucleotide length scale details primer edges for the construction of enhancer fragments. Primers were designed inserting a 5’ SpeI (ACTAGT) and 3’ BamHI (GGATCC) restriction site for subsequent cloning. Table 3.1 list designed primers for each SNMP1 upstream enhancer construct, including restriction site ends.
VP16 – SNMP1-Gal4 enhancer driver cloning. SNMP1 enhancer fragments were isolated through PCR (Table 3.1) and blunt-end ligated and cloned into a pcrII TOPO Vector provided by TOPO TA Cloning Kit (Invitrogen). TOPO – SNMP1 upstream fragment plasmid was transformed through heat-shock treatment of chemically competent Top10 E. coli cells (Invitrogen) and grown in LB Agar plates with a final concentration of ampicillin of 100 µg / ml. Culture was grown overnight at 37 ° C and selected based on disruption of LacZ gene (white versus blue colonies). Genomic DNA was extracted using Qiagen mini-prep spin coloumn protocol and sent for sequencing using standard M13 Primers (Invitrogen). Insertion of correct upstream SNMP1 fragment was further qualified by restriction digest (SpeI and BamHI) and visualized on a 4 % TAE DNA electrophoresis gel. Digested SNMP1 upstream fragment with restriction enzyme sticky ends were dephosphorylated (Antartic phosphatase: NEB) and ligated in to SpeI and BamHI digested VP16-Gal4 P-element vector (128). VP16 vector was constructed by Makoto Makishima and David Mangledorf and was a generous gift from Dean Smith (Southwestern Medical Center. Dallas, TX). VP16 is derived from a pCASPER 4 vector with a GAL4 region and adjacent cloning site suitable for inserting driver sequence of interest (128). The resulting ligated
VP16 – SNMP1-Gal4 vector plasmid was sequenced using Forward 5’ – GCGTATGCGTGATATTGTTA – 3’ and Reverse 5’ – AAGCTTCTTGATGGCGGATA – 3’ VP16 primers. Upon sequence verification plasmids were transformed into Top10 E-coli competent cells and grown in 150 ml LB broth (100 µg / ml Ampicillan) culture for subsequent genomic isolation using Qiagen midi-prep columns. Concentrated genomic pellet was diluted in 50 µl of dH2O to a final concentration of > 1 µg/µl of dissolved DNA, measured in a Thermo Scientific NanoDrop 2000. Concentrated plasmid preps were sent for P-element transformation to Model System Genomics of Duke University for Drosophila (w1118) embryo injections of concentrated plasmid.

Motif mutagenesis. Selected hexanucleotide cis-regulatory motif was mutated using QuikChange II (Agilent Technologies) for high-fidelity replication of complete plasmid and insert sequences. Mutations were introduced via annealing of forward and reverse mutagenic primers flanking the site of interest.

Molecular determination of transgenic progeny. Two to three days post-eclosion of F1 crosses, single fly was subjected to whole body genomic isolation (based on standard protocol) for PCR. PCR was carried out with primers flanking the driver and gal4 insert region of VP16
(For: 5’ – CGAAAGAACCTGGTACA – 3’; Rev: 5’-
 CCCAGTTGTTTCTTCAGAC – 3’). PCR products were cloned and sequenced using previously stated protocol.

**Drosophila strains**

Positive p-element insertion was assessed by a resulting red-eyed P1 generation, resulting from back-crossing each individual adult with w1118 stock flies. Drosophila stocks were maintained in a standard corn flour, yeast and molasses medium under a 12 h light and dark cycle at 25 C. The following transgenic strains were used for SNMP1 upstream regulatory analysis: S1.1(1.95 kb) – Gal4, S1.2(1.56 kb) – Gal4, S1.3(0.86 kb) – Gal4, S1.4(0.4 kb) – Gal4. UAS-mcd8:GFP virgin flies were crossed to red-eyed males to produce a P1 generation exhibiting enhancer mediated cd8::GFP expression. Enhancer - Gal4 insertions in transgenic lines were mapped using yw; Cyo; Sb,TM3 / T(2;3)ap[Xa], ap[Xa] (Bloomington Stock #2475).

**Histology**

*Visualization of chemosensory appendages.* P1 enhancer driven cd8::GFP expressing flies were screened under CO2 in a 10x Nikon dissecting scope fitted with an epifluorescence lamp and screened for bright fluorescence in chemosensory appendages. Flies were further anesthetized in cold and
washed 3x in PBST. Antennae, palps and labellum, legs, wings and halters were dissected and fixed in 4 % paraformaldehyde solution for 15 minutes. Tissue was washed 3x for 5 min in PBST. Tissue was treated with Vectashield (Vector Laboratories) and mounted on a slide with a coverslip with 2 1 μm coverslips used as spacers. Legs and wings were visualized under 40x objective epifluorescence microscope (Nikon SMZ800). Images were captured under fluorescence (λ: Em = 488 nm; Ex = 509 nm) and brightfield; .tif images were merged and aligned using Adobe Photoshop CS5 extended software. Antennae, palps, and labellum were visualized under 60x objective confocal microscope (Nikon Eclipse TE300 fitted with a BioRad MRC 1024ES laser scanning microscope). GFP (λ: Em = 488 nm; Ex = 509 nm) and cuticle background fluorescence (λ: Em = 578 nm; Ex = 603 nm) stacks were scanned at 1 – 2 μm increments, images were adjusted and merged using Image J and Adobe Photoshop CS5 software.

**Immunohistochemistry of CNS tissue.** Brains and thoracic ganglion (TG) were prepared as described in Wu and Luo, 2006. Flies were anesthetized in cold and pinned under 1xPBST (0.3% Triton-X), cuticle surrounding brain and TG was broken and fixed at RT for 20 minutes in 4% paraformaldehyde solution. Fly bodies were washed in 1x
PBST 3x for 5 minutes, brains and TG were dissected and immersed in 5 % Normal goat serum (NGS) and placed on a nutation device for 1 hour at RT. Brains and TG were washed once in 1xPBST and treated with primary antibody at 4 °C for 48 hrs (Rabbit anti-GFP [Invitrogen], 1:2000; mouse monoclonal anti-nc82 [DSHB], 1:40). Tissue was washed 3x PBST for 20 minutes each and treated with secondary antibody for 24 hours at 4°C overnight under gentle nutation (Alexa 488 conjugated Goat anti-rabbit, 1:2000; Alexa 568 conjugated Goat anti-mouse [Invitrogen], 1:2000). Tissue was washed 3x in PBST for 20 min each and immersed in a 20 μl of Slowfade Gold (Invitrogen). Tissue was mounted on a handmade staged slide; 1 μm glass coverslips with a 1 μm coverslip spacer. All confocal images were collected within 24 hours of mounting.

Confocal imaging and processing. Whole brain and TG images were acquired under 20x and 40x objectives. AL with SOG and partial TG were scanned under 40x objectives, whole AL and SOG images were captured under 100x. Tissue was scanned at 1 – 2 μm increments and corrected for noise using a Gaussian spatial filter. Image stacks were compiled, adjusted, and merged using ImageJ software.
Bioinformatics

Comparative sequence analysis. Possible cis-regulatory regions were identified by comparing ~2.0 Kb upstream of the SNMP1 (CG7000) transcriptional start site from both *D. melanogaster* (FlyBaseID: FBgn0260004) and *D.pseudoobscura* (FlyBaseID: FBgn0080014). Family Relations II is a web server application with an algorithm used for a pairwise alignment of ~10 – 35 bp sequences with a percentage of mismatches. SNMP1 analysis was set at a 85 – 90 % stringency per alignment window within the ~2.0 kb sequence. Primers were designed bordering regions of high percentage of homology.

Motif analysis. The regulatory sequence analysis tools (RSAT, http://rsat.ulb.ac.be/rsat/) web based software suite contains many tools used for the detection of cis-regulatory elements (CREs) in genome sequences. 5’ upstream sequences for orthologous SNMP1 in *Drosophila* species and *D. melanogaster* SNMP1 co-expressed chemosensory genes were acquired through the retrieve ENSEMBL seq tool. The ENSEMBL seg tool permitted the simultaneous acquisition of ~2.0 kb upstream sequences. The sequences were entered into the oligo-analysis tool with a background correction Markov model (MM = 3). Motif was selected based on the
graphical overrepresentation and position / sequence similarities.
Figure 3.1: Comparative approach between *D. melanogaster* and *D. pseudoobscura* identifying conserved elements in 5′ upstream region.

A. Phylogeny of holometaboulous insects showing 65 – 43 Mya separation between Dmel and Dpse. B. Comparative analysis of Dmel (top sequence) and Dpse (bottom sequence), red blocks indicate elements of 90% similarity. C. Diagram of dissected SNMP1 upstream region, S1-00 and S1-0 are drivers used in previous study. Green blocks indicate the coded Gal4 gene. Green indicates positive mCD8::GFP expression, blue indicates no expression was identified.
Figure 3.2: Neuronal expression of SNMP1 dissected regulatory region showing difference in expression.

A. Map of dissected elements. Red blocks indicate comparatively identified conserved elements. B. Expression pattern in brain and TG of S1.0-Gal4; S1.1-Gal4; S1.4-Gal4. S1.4-Gal4 brain shows change in expression of projections to glomeruli and SOG. Bottom panel S1.4-Gal4 TG shows expression of possible mechanosensory projections.
Figure 3.3: RSAT oligo-analysis output of S1.4 upstream sequence compared to orthologous and co-expressed genes.

A. Output of motif analysis identified a highly conserved element (GCAATTA). B. Map featuring S1.4 upstream sequence compared to orthologous SNMP1 sequences from Dsim, Dere, and Dpse. Possibly co-expressed genes based on expression profile in the AL glomeruli. Blue blocks indicate position of conserved GCAATTA element in other upstream sequences.
Figure 3.4: Mutagenesis of GCAATTA element in S1.4-Gal4 driver sequence.

A. Site-directed mutagenesis of GCAATTA element by base-pair change inserting AvrII restriction site (GCCTAGG). B. 100x confocal stack images of AL from S1.1-Gal4 driver and S1.4(AvrII)-Gal4 driver with mutated element.
Figure 3.5: Expression of SNMP1 driven mCD8::GFP by different SNMP1 enhancer drivers reveals change in glomeruli expression.

100x merged confocal stacks of mCD8::GFP expression in AL and anterior (Ant), medial (Med), and posterior (Post) illustration of GFP positive glomeruli between S1.0-Gal4 (A); S1.4-Gal4 (B); S1.4(AvrII)-Gal4 (C).
<table>
<thead>
<tr>
<th>Target Name</th>
<th>Primers</th>
<th>Reporter Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1.00 (7.00 Kb)</td>
<td>FOR. 5’ - ACTAGTCACCTTGGCTTTCCATCGACTACG - 3’&lt;br&gt;REV. 5’ - GGATCCGAGCTGAAAGACGAGGCTTC - 3’</td>
<td>Yes</td>
</tr>
<tr>
<td>S1.0 (3.10 Kb)</td>
<td>FOR. 5’ - ACTAGTAGCCATATCGCTGGGAGAAC - 3’&lt;br&gt;REV. 5’ - GGATCCGAGCTGAAAGACGAGGCTTC - 3’</td>
<td>Yes</td>
</tr>
<tr>
<td>S1.1 (1.95 Kb)</td>
<td>FOR. 5’ - ACTAGTCACCTTGGCTTTCCATCGACTACG - 3’&lt;br&gt;REV. 5’ - GGATCCCTGCTCCTCGAGCTGAAGACC - 3’</td>
<td>Yes</td>
</tr>
<tr>
<td>S1.2 (1.56 Kb)</td>
<td>FOR. 5’ - ACTAGTAACCTCGAGTCTGGGAGGCGA - 3’&lt;br&gt;REV. 5’ - GGATCCCTGCTCCTCGAGCTGAAGACC - 3’</td>
<td>Yes</td>
</tr>
<tr>
<td>S1.3 (0.86 Kb)</td>
<td>FOR. 5’ - ACTAGTACAAACCTCGAGTCTGGCT - 3’&lt;br&gt;REV. 5’ - GGATCCCTGCTCCTCGAGCTGAAGACC - 3’</td>
<td>Yes</td>
</tr>
<tr>
<td>S1.4 (0.40 Kb)</td>
<td>FOR. 5’ - ACTAGTGGCCCATTGAGCTGAGCA - 3’&lt;br&gt;REV. 5’ - GGATCCCTGCTCCTCGAGCTGAAGACC - 3’</td>
<td>Yes (core promoter)</td>
</tr>
<tr>
<td>S1.5 (1.36 Kb)</td>
<td>FOR. 5’ - ACTAGTGCCCCCTTGAGACCTAGGACATTAGCTACC - 3’&lt;br&gt;REV. 5’ - GGATCCACCTTAGTCCAGCTGGATTAGCTACC - 3’</td>
<td>No</td>
</tr>
<tr>
<td>S1.6 (0.75 Kb)</td>
<td>FOR. 5’ - ACTAGTGCCCCCTTGAGACCTAGGACATTAGCTACC - 3’&lt;br&gt;REV. 5’ - GGATCCACCTTAGTCCAGCTGGATTAGCTACC - 3’</td>
<td>No</td>
</tr>
<tr>
<td>S1.7 (0.31 Kb)</td>
<td>FOR. 5’ - ACTAGTGCCCCCTTGAGACCTAGGACATTAGCTACC - 3’&lt;br&gt;REV. 5’ - GGATCCACCTTAGTCCAGCTGGATTAGCTACC - 3’</td>
<td>No</td>
</tr>
<tr>
<td>S1.8 (0.98 Kb)</td>
<td>FOR. 5’ - ACTAGTAACCTCGAGTCTGGGAGGCGA - 3’&lt;br&gt;REV. 5’ - GGATCCACCTTAGTCCAGCTGGATTAGCTACC - 3’</td>
<td>No</td>
</tr>
</tbody>
</table>
CHAPTER 4

CHARACTERIZATION OF SNMP2 GENE REGULATION BY ENHANCER DRIVER DISSECTION

4.1 INTRODUCTION

Broad chemosensory expression of SNMPs between cell type (neuronal/support cell) and chemosensory modality (olfactory/gustatory) are evidence for an intricate regulatory landscape flanking the transcriptionally coded gene. Based on previous methodology using an upstream enhancer driver indirectly reporting the expression of GFP facilitates the discovery of important cis-regulatory elements (CRE). Targeted dissection of the SNMP2 enhancer driver was attempted in order to qualify the regulatory control of broadly represented Snmp2-Gal4 expression pattern.

Evolutionary pressure on non-coded DNA flanking a gene might be a principal driving force behind endless forms of morphological diversity (143-145). The insect chemosensory system offers a robust evolutionary model demonstrating diverse chemotactic behavior as a function of a highly
expanded and divergent protein network with the functional role of selectively recognizing environmental chemical signals (5, 22, 66). Networks of chemosensory proteins working at the periphery (OBPs, ORs, GRs, IRs ODEs, SNMPs) are selectively expressed at an individual sensory neuron or adjacent support cell conferring a functional phenotype. The expression patterns of chemosensory proteins are essential to an insect’s fitness by conferring the ability to detect and integrate both attractive and aversive stimuli (3, 110, 146, 147). Recent interest has been building on understanding the role of cis-regulation on highly organized and orchestrated expression of proteins within a complex network of functionally related cells (81, 82, 137, 148). It is well accepted that the non-transcribed nucleotides flanking the transcriptionally coded gene themselves code for the spatial and temporal instructions during gene transcription. Alternate biological molecular functions of regulatory control cannot be overlooked i.e. chromatin modification and small RNA silencing. However, the cis-regulatory control elements present a method of regulation under similar evolutionary pressure as a coding sequence without confounding negative pleotropic effects (144, 149).
SNMPs were first discovered in Lepidoptera and characterized as pheromone specific proteins expressing in neuronal (SNMP1) and support (SNMP2) cells (85, 88). In Drosophila melanogaster SNMP1 has been attributed in assisting receptor activity within the social aggregate and aggression pheromone cis-vaccenyl acetate (cVA) pathway (21, 25). However, the characterized expression pattern suggest a broader functional role for SNMPs, which coincides with its sequence homology to the broadly expressed and functioning family of type B1 scavenger receptor, SR-B1 / CD36 (86).

The D. melanogaster SNMP2 has been characterized in both olfactory and gustatory tissue expressing in the antenna, palps, legs, and wings. Axonal projections from the antennae to the antennal lobe (AL) glomeruli report OSN expression in a number of palp and antennal basiconic and coeloconic sensilla. SNMP2 expression largely differentiates from SNMP1 in its representation in the gustatory sensory neurons (GSNs). GSN projections exhibited by Snmp2-Gal4 (named S2.0-Gal4 in this study) to the suboesophageal ganglion (SOG) reveal broadly expressed pattern from the proboscis and frontal tarsi. Broad gustatory projections from the legs and wings are represented in the pro-, meso-, and meta-thoracic
neuromeres of the thoracic ganglia (TG), with pro-thoracic projections demonstrating sexual dimorphism.

SNMP2 represents a novel chemosensory phenotype in insects, expressing between two sensory modalities and cell types (neuronal / support cell). Possibly co-expressing with many poorly characterized GSNs. Considering how largely represented SNMP2 is in the gustatory network of cells in different sensory epithelia delivers a fascinating model for understanding regulatory control of a sensory gene represented in different developmental networks. Implementing the powerful tool of characterizing native gene expression by indirect upstream enhancer driven expression of GFP, the upstream regulatory control of SNMP2 is characterized through enhancer / driver dissection.

Uncovering a cis-regulatory network requires a highly integrated approach of sequence analysis and molecular dissection. Under the assumption that a 5′-UTR is under the same evolutionary constraints as the coding region small islands of conservation may be detected and assumed to be of regulatory importance. The binary enhancer trap system facilitates the visual inspection of different regulatory elements associating changes in temporal and spatial expression to elements of upstream control. This analysis provides a broad characterization of regulatory
function, however it facilitates a top-down approach. Cis-regulatory elements are small conserved nucleotide sequences varying in size, sequence, and position. Bioinformatic web-based tool-kits implement regulatory motif databases and algorithms based on over-representation and conserved position to identify motifs of possible regulatory importance. In this study we identified and mutated a likely functional motif and found it to alter the spatial expression of the reporter gene mCD8::GFP. Suggesting a motif responsible for broad expression of SNMP2 in gustatory neuronal cell expression as evidenced by axonal projections to the SOG. This study provides evidence of a diversely regulated chemosensory protein with a complex regulatory landscape of enhancers for SNMP2 GSN expression within a core-promoter element of transcriptional regulation.

4.2 Results

Comparative analysis of orthologous 5’ upstream SNMP2 in D. melanogaster and D. pseudoobscura

Comparative analysis between two SNMP2 orthologous sequences between D.melanogaster and D. pseudoobscura was employed in order to discover cis-regulatory elements
(CREs) within 2000bp of the start ATG of Dmel SNMP2. The precise tss of D. melanogaster SNMP2 was later identified and annotated using 5'RACE. In the snmp2 gene the tss delineates the start of a transcribed non-coded portion of the first exon with the start ATG located 397 bp upstream of the tss. Based on comparative analysis the exon1 non-coded transcribed sequence is not highly conserved (Fig4.1B). Presumed CREs were identified from a sequence comparison held at 95 and 85% stringency. Possible CREs were used as guide marks for primer design in the dissection of the previously characterized S2.0-Gal4 (5.0Kb) driver. Eight drivers were characterized through transgenic fly mCD8::GFP expression of which only four drivers showed positive GFP expression in previously characterized SNMP2 peripheral expression profile. Drivers proximal to the tss and from the transcribed non-coded portion of exon 1 did not show to have core-promoter function. The upstream sequence between -500 and -150 bp of tss S2.8-Gal4 (0.37Kb) shows evidence of core promoter function based on GFP reporter expression at the periphery. (Fig4.1C).
Analysis of neuronal projections from dissected Snmp2 drivers shows conserved expression within the gustatory centers of the fly CNS.

Dissection of Snmp2 upstream driver exhibited a broad change in expression of neuronal projections from the antennae and palps indicating far upstream regulatory enhancers between the previously characterized driver S2.0- Gal4 (5.00Kb) and S2.1-Gal4 (1.43Kb). Antennal neuronal projections in S2.0-Gal4 reported possible SNMP2 expression in a number of antennal coeloconic, basiconic and palp basiconic glomeruli. The first dissected driver S2.1-Gal4 showed a completely abolished neuronal projection pattern from olfactory centers, suggesting a far upstream control of either complete sensory specific regulatory control or olfactory cell type (neuronal or support cell) regulation. SNMP2 expression in GSNs is broad and characterized from neuronal projections from labellar, pharyngeal, and tarsal axons to the SOG. Additionally, gustatory sensilla from legs and wings express the ipsilateral GSN projections in the TG neuromeres. The GSN projections to the SOG and TG were vastly unchanged from the S2.1-Gal4 (1.40Kb) driver to the minimal S2.8-Gal4 (0.37Kb) core promoter (Fig4.2B).

This analysis demonstrates highly partitioned regions of regulatory control based on sensory modality. Further
characterization of the far upstream region spliced to the core promoter would be necessary to test the validity of this assumption.

**Identification of highly conserved and over-represented motif between Snmp2 and co-expressed gene sequences.**

Motif sequence analysis tool (MEME) (150) was used to determine functional motifs in S2.2-Gal4 (0.90Kb) driver. Motif predicting algorithms are designed for identifying and characterizing shared motifs in a set of orthologous or co-transcribed genes. Most algorithms work under three assumptions: 1) Functional motifs are highly conserved in sequence and position, 2) Conserved motifs have a high occurrence amongst a given dataset, 3) within a range of ~5 to 20 nucleotides (151). S2.2-Gal4 was compared to orthologous sequences from D. simulans, D. erecta, and D.pseudoobscura. Additionally, the enhancer + core promoter S2.2-Gal4 driver sequence was compared with co-expressed genes from GSNs in labellum and tarsi (Fig4.3A). MEME analysis calculated a degenerate position weight matrix of 20 nucleotides. Cross analysis with the open access Drosophila TF binding motif database, JASPAR (152), identified a previously classified motif that was found overrepresented amongst the majority of Snmp2 compared sequences (Fig4.3B).
Mutagenesis of TTAATCG motif reduces expression in the SOG.

Site directed mutagenesis of four base pairs, introducing the NheI endonuclease restriction site, converted strictly conserved AT nucleotides to GC. This mutation was sufficient to partially abolish expression of mCD8::GFP by S2.2 (NheI)-Gal4 driver in SOG. Snmp2 representative expression using the S2.2-Gal4 (0.90Kb) promoter driver faithfully represented the previously characterized expression pattern in the SOG and TG. Transgenic lines with the S2.2(NheI)-Gal4 mutated sequence exhibited a minimalized SOG expression, Miyazaki and Ito, 2010 identified this projection as the PMS4 projection from the labellum. Peripheral expression was noted in all the chemosensory organs, however expression to the labellum, legs and wings was greatly reduced, however change to the TG expression pattern was not noticed.

3.3 DISCUSSION

Snmp2-Gal4 expression was characterized as broadly associated with GSNs, characterized expression in the SOG and chemosensory projections to the TG. The antennal and palp OSNs associated with SNMP2 were immediately lost when reducing the driver from 5.0 Kb to a 1.5 Kb element, suggesting the previously characterized expression in the
AL glomeruli is driven by enhancer elements far upstream from the core-promoter. Expression from the GSNs driven by the core promoter was still maintained. It is possible that the SNMP2 upstream control requires two groupings of enhancers based on sensory modality, spatial regulation of gustatory sensilla are seemingly held close to the core promoter, while olfactory inputs are controlled by distal enhancers. Motif discovery and subsequent mutagenesis of a possible TF binding element, demonstrated a localized change in neuronal expression to the SOG.

Computational estimation of a regulatory binding motif is often met with difficulty. Consensus sequences bound by transcription factors are often very short, exhibiting variability in TF binding affinities, at times highly permissive to sequence degeneracy. Many programs are designed to implement algorithms that take into account many of these features. The RSAT suite of pattern discovery and recognition is a powerful collection of tools and algorithms with a high success rate of discovering functional motifs. However, the previously successful motif discovery for the SNMP1 core-promoter using oligo-analysis program proved insufficient for the S2.2-Gal4 sequence. This was possibly a result of incorrect parameters, unlikely co-expressed sequences, or a highly
degenerate yet functional motif between orthologous sequences. The MEME motif discovery algorithm uses a similar maximum likelihood based approach, however the analysis is based on a relaxed statistical interpretation of longer nucleotide strings of interspersed repeated elements. This was permissive for the identification of highly degenerate twenty-nucleotide sequence, which was further matched to a possible candidate TF binding motif.

The S2.2(NheI)-Gal4 mutated motif abolished the majority of neuronal projections to the SOG. However, no change to the TG was noted, signifying a highly conserved element with specified function for either complete expression of a group of GSNs or control of cell type localization. Previous study characterized the labellar sensilla into seven classes of neurons sending axonal afferents to the SOG through the labial nerve using horseradish peroxidase injections to the lateral sensilla (39). This study classified projections based on groups of labellar sensilla, grouping sensilla into type I, II, and IV-VII neurons. Types I and II seem to project to the majority of broad projections patterns in the SOG, covering the PMS, AMS, and LS branches. While type VII neurons were found to correspond to the PMS4 projections patterns demonstrated by Miyazaki and Ito, 2010. The expression
pattern seems to most resemble those tracing from the trehalose receptor Gr5a-Gal4 in various studies (153-156). It is likely that the discovered motif is responsible for the cellular identity of sugar sensitive projections.

Further characterization of the discovered motif using the TOMTOM query web based software (157) providing a p-value based match of the query sequences and previously identified position-weight matrices for binding motifs from multiple databases, returned a significant match to a binding motif associated with the homeobox Bar-H1 transcription factor (p = 0.0015). This association proves to be most interesting because of the role of Bar-H1 in lineage based cellular development in determining sensory or support cell identity, in addition to maintaining a high expression level in thecogen support cell and sensory neuron. (158-162). Future studies, would attempt to characterize the olfactory inputs from the Snpmp2 by splicing far upstream with the minimal core promoter sequence. Additionally, in situ hybridization technique is needed to establish if the enhancer mutation affected overall expression or mediated cell-type expression. Binding assays would be needed to confirm a role of BarH1 in organizing the biochemical network of sensory proteins within the labellar GSNs.
4.4 **EXPERIMENTAL PROCEDURE**

**Plasmid Construction**

*Construction of enhancer / driver insert sequences.* SNMP2 upstream enhancer fragments were amplified by PCR from BAC library, BACR10M16. Primers were designed based on short segments of high conservation identified through evolutionary comparative analysis (Fig 4.1B). Figure 4.1C, nucleotide length scale details primer edges for the construction of enhancer fragments. Primers were designed inserting a 5’ SpeI (ACTAGT) and 3’ BamHI (GGATCC) restriction site for subsequent cloning. Table 4.1 list designed primers for each SNMP1 upstream enhancer construct, including restriction site ends.

*VP16 – SNMP2-Ga4 enhancer driver cloning.* SNMP1 enhancer fragments were isolated through PCR (Table 4.1) and blunt-end ligated and cloned into a pcrII TOPO Vector provided by TOPO TA Cloning Kit (Invitrogen). TOPO – SNMP1 upstream fragment plasmid was transformed through heat-schock treatment of chemically competent Top10 E. coli cells (Invitrogen) and grown in LB Agar plates with a final concentration of ampicillin of 100 µg / ml. Culture was grown overnight at 37 °C and selected based on disruption of LacZ gene (white versus blue colonies). Genomic DNA was
extracted using Qiagen mini-prep spin column protocol and sent for sequencing using standard M13 Primers (Invitrogen). Insertion of correct upstream SNMP1 fragment was further qualified by restriction digest (SpeI and BamHI) and visualized on a 4 % TAE DNA electrophoresis gel. Digested SNMP2 upstream fragment with restriction enzyme sticky ends were dephosphorylated (Antarctic phosphatase: NEB) and ligated in to SpeI and BamHI digested VP16-Gal4 P-element vector (128, 129). VP16 vector was constructed by Makoto Makishima and David Mangledorf and was a generous gift from Dean Smith (Southwestern Medical Center. Dallas, TX). VP16 is derived from a pCASPER 4 vector with a GAL4 region and adjacent cloning site suitable for inserting driver sequence of interest (128). The resulting ligated VP16 – SNMP2-Gal4 vector plasmid was sequenced using Forward 5′ – GCGTATGCGTGATATTTTTGTTA – 3′ and Reverse 5′ – AAGCTTCTTGTATG GCCGATA – 3′ VP16 primers. Upon sequence verification plasmids were transformed into Top10 E-coli competent cells and grown in 150 ml LB broth (100 µg / ml Ampicillan) culture for subsequent genomic isolation using Qiagen midi-prep columns. Concentrated genomic pellet was diluted in 50 µl of dH2O to a final concentration of > 1 µg/µl of dissolved DNA, measured in a Thermo Scientific NanoDrop 2000. Concentrated plasmid preps were sent for P-
element transformation to Model System Genomics of Duke University for Drosophila (w1118) embryo injections of concentrated plasmid.

Motif mutagenesis. Selected hexanucleotide cis-regulatory motif was mutated using QuikChange II (Agilent Technologies) for high-fidelity replication of complete plasmid and insert sequences. Mutations were introduced via annealing of forward and reverse mutagenic primers flanking the site of interest.

Molecular determination of transgenic progeny. Two to three days post-eclosion of F1 crosses, single fly was subjected to whole body genomic isolation (based on standard protocol) for PCR. PCR was carried out with primers flanking the driver and gal4 insert region of VP16 (For: 5' – CGAAAGAACCTGGTACA – 3'; Rev: 5’-CCCAGTTGTTCTTCAGAC – 3’). PCR products were cloned and sequenced using previously stated protocol.

Drosophila strains

Positive p-element insertion was assessed by a resulting red-eyed P1 generation, resulting from back-crossing each individual adult with w1118 stock flies. Drosophila stocks were maintained in a standard corn flour, yeast and molasses medium under a 12 h light and dark cycle at 25°C.
The following transgenic strains were used for SNMP1 upstream regulatory analysis: S2.1(1.43 kb) – Gal4, S2.2(0.90 kb) – Gal4, S2.7(0.90 kb) – Gal4, S2.8(0.37 kb) – Gal4. UAS-mcd8:GFP virgin flies were crossed to red-eyed males to produce a P1 generation exhibiting enhancer mediated cd8::GFP expression. Enhancer - Gal4 insertions in transgenic lines were mapped using yw; Cyo; Sb,TM3 / T(2;3)ap[Xa], ap[Xa] (Bloomington Stock #2475).

**Histology**

*Visualization of chemosensory appendages.* P1 enhancer driven cd8::GFP expressing flies were screened under CO2 in a 10x Nikon dissecting scope fitted with an epifluorescence lamp and screened for bright fluorescence in chemosensory appendages. Flies were further anesthetized in cold and washed 3x in PBST. Antennae, palps and labellum, legs, wings and halter were dissected and fixed in 4 % paraformaldehyde solution for 15 minutes. Tissue was washed 3x for 5 min in PBST. Tissue was treated with Vectashield (Vector Laboratories) and mounted on a slide with a coverslip with 2.1 μm coverslips used as spacers. Legs and wings were visualized under 40x objective epifluorescence microscope (Nikon SMZ800). Images were captured under fluorescence (λ: Em = 488 nm; Ex = 509 nm) and brightfield; .tif images were merged and aligned using
Adobe Photoshop CS5 extended software. Antennae, palps, and labellum were visualized under 60x objective confocal microscope (Nikon Eclipse TE300 fitted with a BioRad MRC 1024ES laser scanning microscope). GFP ($\lambda$: Em = 488 nm; Ex = 509 nm) and cuticle background fluorescence ($\lambda$: Em = 578 nm; Ex = 603 nm) stacks were scanned at 1 – 2 μm increments, images were adjusted and merged using Image J and Adobe Photoshop CS5 software.

**Immunohistochemistry of CNS tissue.** Brains and thoracic ganglion (TG) were prepared as described in Wu and Luo, 2006. Flies were anesthetized in cold and pinned under 1xPBST (0.3% Triton-X), cuticle surrounding brain and TG was broken and fixed at RT for 20 minutes in 4% paraformaldehyde solution. Fly bodies were washed in 1x PBST 3x for 5 minutes, brains and TG were dissected and immersed in 5 % Normal goat serum (NGS) and placed on a nutation device for 1 hour at RT. Brains and TG were washed once in 1xPBST and treated with primary antibody at 4 °C for 48 hrs (Rabbit anti-GFP [Invitrogen], 1:2000; mouse monoclonal anti-nc82 [DSHB], 1:40). Tissue was washed 3x PBST for 20 minutes each and treated with secondary antibody for 24 hours at 4°C overnight under gentle nutation (Alexa 488 conjugated Goat anti-rabbit, 1:2000; Alexa 568 conjugated Goat anti-mouse [Invitrogen],
1:2000). Tissue was washed 3x in PBST for 20 min each and immersed in a 20 μl of Slowfade Gold (Invitrogen). Tissue was mounted on a handmade staged slide; 1 μm glass coverslips with a 1 μm coverslip spacer. All confocal images were collected within 24 hours of mounting.

Confocal imaging and processing. Whole brain and TG images were acquired under 20x and 40x objectives. AL with SOG and partial TG were scanned under 40x objectives, whole AL and SOG images were captured under 100x. Tissue was scanned at 1 – 2 μm increments and corrected for noise using a Gaussian spatial filter. Image stacks were compiled, adjusted, and merged using ImageJ software.

Bioinformatics

Comparative sequence analysis. Possible cis-regulatory regions were identified by comparing ~ 2.0 Kb upstream of the SNMP2 (CG7422) transcriptional start site from both D. melanogaster (FlyBaseID: FBgn0035815) and D. pseudoobscura (FlyBaseID: FBgn0080333). Family Relations II is a web server application with an algorithm used for a pairwise alignment of ~ 10 – 35 bp sequences with a percentage of mismatches. SNMP2 analysis was set at a 80 – 95 % stringency per alignment window within the ~ 2.0 kb
sequence. Primers were designed bordering regions of both 85% and 90% high percentage of homology.

Motif analysis. The regulatory sequence analysis tools (RSAT, http://rsat.ulb.ac.be/rsat/) web based software suite contains many tools used for the detection of cis-regulatory elements (CREs) in genome sequences. 5′ upstream sequences for orthologous SNMP1 in Drosophila species and D. melanogaster SNMP1 co-expressed chemosensory genes were acquired through the retrieve ENSEMBL seq tool. The ENSEMBL seq tool permitted the simultaneous acquisition of ~2.0 kb upstream sequences. The sequences were entered into the oligo-analysis tool with a background correction Markov model (MM = 3). Motif was selected based on the graphical overrepresentation and position / sequence similarities. An alternate method of motif discovery was used. The MEME (http://meme.nbcr.net) web application tool is based on expectation maximization, providing a relaxed interpretation of possible cis-regulatory nucleotides, increasing false–positives, but also revealing possible previous false-negatives. Results were cross-analyzed using TOMTOM tool for alignment of identified statistically significant conserved region of up to 20 nucleotides with short string hexanucleotides previously identified in Drosophila transcription factor motif databases.
Figure 4.1: Comparative approach between *D. melanogaster* and *D. pseudoobscura* identifying conserved elements in 5’ upstream region of SNMP2.

A. Phylogeny of holometaboulous insects showing 65 – 43 Mya separation between Dmel and Dpse. B. Comparative analysis of Dmel (top sequence) and Dpse (bottom sequence), red bars indicate elements of 90 % similarity, blue bars indicate area of 85 % similarity. C. Diagram of dissected SNMP2 upstream region, S2.0 driver was characterized in previous study. Green blocks indicate the coded Gal4 gene. Green indicates positive mCD8::GFP expression, gray indicates no expression.
Figure 4.2: Neuronal expression of SNMP2 dissected regulatory region show minimal differences in expression.

A. Map of dissected elements. Red blocks indicate comparatively identified conserved elements. B. Expression pattern in brain and TG of S2.0-Gal4; S2.1-Gal4; S2.8-Gal4. S2.1-Gal4 and core promoter S2.8-Gal4 show loss of expression in the AL. Expression in the TG was conserved down to the core promoter.
Figure 4.3: RSAT motif-finder analysis of S2.2 upstream sequence motif compared to orthologous and co-expressed genes.

A. Output of motif analysis; map featuring (TAA(T/A)(T/C)G) motif sequence compared to upstream SNMP2 orthologue in Dsim, Dere, and Dpse. Sequence was also compared to possibly co-expressing GR upstream sequences and Obp49a, Obp57d, and Obp57e. B. MEME analysis broadly characterized 20-nucleotide conserved sequence. Output was queried on JASPAR database and matched (TAA(T/A)(T/C)G) motif sequence. C. Mutagenesis of TTAAATCG motif sequence by insertion of NheI restriction site. Star indicates the position of point mutations.
Figure 4.4: Mutagenesis of TTTAATCG element in S2.2-Gal4 driver sequence.

A. Site-directed mutagenesis of TTTAATCG element by base-pair change inserting NheI restriction site (GCTAGCCG). B. 100x confocal stack images of SOG from S1.2-Gal4 driver and S2.2(NheI)-Gal4 driver with mutated element.
Table 4.1: SNMP2 upstream dissection primers

<table>
<thead>
<tr>
<th>Target Name</th>
<th>Primers</th>
<th>Reporter Expression</th>
</tr>
</thead>
</table>
| S2.0 (5.00 Kb) | FOR. 5' - ACTAGTGCGATGCGATATTTTTTGTATAGTTTA - 3'  
REV. 5' - GGATCCAGATACGGATGGATG - 3' | Yes |
| S2.1 (1.43 Kb) | FOR. 5' - ACTAGTAGTTATATAGTTTATAGCCGATGATG - 3'  
REV. 5' - GGATCCAGATACGGATGGATG - 3' | Yes |
| S2.2 (0.90 Kb) | FOR. 5' - ACTAGTGCTACGTGACGTGGGCA - 3'  
REV. 5' - GGATCCAGATACGGATGGATG - 3' | Yes |
| S2.3 (0.55 Kb) | FOR. 5' - ACTAGTGCTCAGACCGCAAACGAA - 3'  
REV. 5' - GGATCCAGATACGGATGGATG - 3' | No |
| S2.4 (0.35 Kb) | FOR. 5' - ACTAGTTATAGTTAAAGCAGGATGGATG - 3'  
REV. 5' - GGATCCAGATACGGATGGATG - 3' | No |
| S2.5 (0.21 Kb) | FOR. 5' - ACTAGTGCTCAGACCGCAAACGAA - 3'  
REV. 5' - GGATCCAGATACGGATGGATG - 3' | No |
| S2.6 (0.90 Kb) | FOR. 5' - ACTAGTTATAGTTAAAGCAGGATGGATG - 3'  
REV. 5' - GGATCCAGATACGGATGGATG - 3' | Yes |
| S2.7 (0.16 Kb) | FOR. 5' - ACTAGTGCTCAGACCGCAAACGAA - 3'  
REV. 5' - GGATCCAGATACGGATGGATG - 3' | No |
| S2.8 (0.37 Kb) | FOR. 5' - ACTAGTGCTCAGACCGCAAACGAA - 3'  
REV. 5' - GGATCCAGATACGGATGGATG - 3' | Yes (Core Promoter) |
CHAPTER 5
EVOLUTION THROUGH CIS-REGULATORY CHANGE

5.1 INTRODUCTION

Charles Darwin stated in the Origin of Species:

“There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved.”

Evo-Devo proponent and biologist Sean Carroll often uses this quote to exemplify a vast diversity of morphology that exist in the natural world, that which cannot be explained solely by studying the coded genome (163). There are approximately 250,000 genes in Arabidopsis, 22,000 in humans, 14,000 in fruit flies, and 6,000 in yeast. These narrow differences in the coded genome do not surmount to the level of organismal complexity observed in the natural
world. Rather, the orchestration of a spatial and temporal gene regulatory code is able to connect different regulatory networks providing biological complexity from a highly conserved developmental tool-kit. In the last two decades a number of studies have observed that complex phenotypic traits have evolved from a re-wiring of old genes. How are these pre-existing genes re-wired? Either their wiring is modulated de novo bringing about subtle changes or they are re-wired through co-opted regulatory instructions. While both of these are possible scenarios the time-scale at which the appearance of novel traits is dependent on the mechanism of regulatory evolution.

5.2 TRANSCRIPTIONAL REGULATION: MECHANISM

Transcriptional regulation is a multi-faceted process mediated by chromatin, histone modification, microRNAs, and transcription factors. While these are all important factors in the orchestration of gene regulation this analysis will focus on cis-regulation, regulation mediated by transcription factors (TF) upstream nucleotide interactions (Fig5.1A,B). In the vicinity of a gene the often expansive landscape of non-coded DNA, extending to over thousands of base pairs in eukaryotes, houses discrete TF binding sites between 4-30 bp long. These binding
sites, cis-regulatory elements, are responsible for connecting expressed TF from one gene to another, creating a network of gene-to-gene communication. They are categorized as either enhancer or repressor sequences that work in concert with the transcriptional initiator promoter sequences to generate levels of transcription. The regulatory interaction is a binding process under biophysical constraints. Transcription factor binding is a 2-fold interaction involving non-specific electrostatic interactions between the negatively charged DNA backbone and a specific interaction involving the appropriate energy signature binding to a specific locus (Fig 5.2). The exact nature of this process is not known, but it can be theoretically assumed through physical chemical interactions that TF interact in a one-dimensional diffusion along the DNA backbone and three-dimensional diffusion characterized by a stochastic bind / release until it reaches a state equilibrium on a specific binding site (164) (Fig5.2). The regulatory binding machinery is probabilistic at a given organismal and cellular time scale dependent on quantity and type of TF present in the nuclear environment and type and position of binding sites. Cis-regulation is a sensitive mechanism; in prokaryotes the binding of one factor is enough to drive transcription, in
Eukaryotes it involves multiple factors (Fig5.1A). Eukaryotic transcription requires the equilibrium between TF availability and binding elements, permitting fine tuning through broad adaptations.

5.3 Functional role of Cis-regulation

Cis-regulatory sequences are found within the 5'- / 3'-untranslated region flanking a gene, in some cases within intron sequences. Approximately 100bp from the transcriptional start site (tss) is the core promoter housing the necessary binding sites for the transcriptional machinery including tissue specific enhancers, resulting in basal levels of transcription (165). Further upstream of the core promoter is made up of enhancer, repressor, and insulator elements, diverse in both sequence and position. Enhancer elements have been found as far as 10Mb from the tss in some cases and are involved in altering the transcriptional rate of the basal core promoter function (Fig5.1A). Enhancer elements may exist as a series of repeated elements or a series of diverse enhancing elements. Cases where the same enhancer element is found repeated throughout the regulatory upstream region either demonstrate the ability of repeated enhancers to significantly increase the rate of transcription or provide
a fail-proof mechanism of phenotypic maintenance (166). In the later case repeated enhancers, termed shadow enhancers, are practically non-functional elements that are activated once another enhancer or series of enhancers fail to specify TF binding. Increasing transcriptional regulation by multiple enhancers is a common feature in cis-regulation, while shadow enhancers are thought to be more common in early developmental genes, which hold critical functional roles (167). More complex roles in cis-regulation involve the function of controlling the space and time of expression. Temporal and spatial expression involves the integrated effort of many regulatory factors (168). Cis-regulation enhancers are accompanied by repressor and insulator elements offering a diverse instructional signal. Enhancer elements come in many different flavors permissive to selectivity of the signal given a specific TF environment. At the level of cis-regulation time and space is a function of the TFs in the nuclear environment and the selectivity of enhancer sequences. In some cases this selectivity may further be altered by a silencer or insulator elements, resulting in modulating constitutive TF binding and regulation mediated by enhancer sequences. This property promotes selectivity filters resulting in highly orchestrated and patterned
expression of genes. Considering the importance in the maintenance of genetic identity the non-coding regulatory regions does not seem to be under tight evolutionary constraints.

5.4 CIS-REGULATORY EVOLUTION

Mutations on regulatory binding sites can often result in deleterious changes, on occasions these mutations may lead to adaptation of a complex trait. However, much debate surrounds the importance of cis-regulatory divergence mechanisms on evolution (169, 170) (Fig 5.3). The importance of cis-regulatory evolution becomes one of practicality. Mutations on developmentally important proteins will often lead to deleterious pleotropic effects, while mutations to the cis-regulatory code would likely produce a more gradual route of adaptation (145, 171, 172). Evolutionary changes in the development of morphological features in Drosophila have provided important data on the mechanism of cis-regulatory evolution. Within populations of African Drosophila melanogaster, abdomen pigmentation seems to vary on distribution and width of yellow and dark bands, which is correlated to geographic distributions (Rebeiz et al., 2009). Using an enhancer / reporter method on the melanin inducing developmental gene ebony, the
researchers found that the distribution of pigmentation was a result of a series of enhancer elements and that single base-pair mutations in certain populations accounted for the differences in the distribution of pigmentation (Fig5.3A). This discovery coincides with a study on sexual dimorphic pigmentation mediated by directed spatial expression of the bric-a-brac TF controlled by two separate CREs. One CRE was found responsible for expression in the anterior abdomen while another was directing female specific expression. The researchers found that this sexual dimorphic pattern has evolved from a monomorphic ancestor through point mutations to the enhancer elements (173). Cis-regulatory evolution can also take place through large-scale changes to the regulatory code. Within the Drosophila melanogaster sub-group, larvae are decorated along its dorsal epithelia with small ‘hairs’, however Drosophila sechellia larvae do not show such a characteristic. Such a divergence in phenotype has been demonstrated to be a result of broad deletions and mutations within enhancer elements (174). Similar studies in wing pigmentation in Drosophila has demonstrated that divergence of the wing pigmentation pattern has resulted by a novel acquisition of enhancer sites for the already existing engrailed TF, directing the location of spot
pigmentation in the Drosophila wing (Fig5.3B). This study presents a method of cis-regulatory evolution through co-option of a regulatory network not present in the ancestral species (144, 171). The authors suggest this co-opted element to be a result of random mutations occurring in an enhancer site resulting in novel TF binding affinity, however it should be addressed that these elements may result from similar genetic processes that result in gene duplication, such as transposition or inter-chromosomal crossing over events (Fig5.3B). Studies in Drosophila have efficiently aided in the understanding of cis-regulatory evolution by comparing closely related versus ancestral species. However, the debate between the roles of cis-regulatory evolution on adaptive morphological changes is contradicted by changes to protein function at the level of comparative analysis of ancestral species (170). One of the best examples of this is the evolution of homeotic HOX proteins in the development of limb body plan in crustacean versus hexapod species (175). Crustacea and arthropods share two HOX protein class (Ubx and AbdA) and the evolution of these led to broad changes in limb body plan. Where mutations to the C-terminal Ser/Thr phosphorylated residues are implicated in changes of body plan between ancestral species (176) (Fig5.3B). A comparative study on
the homeotic protein *Antp* between *Daphnia magna* and *Drosophila melanogaster* indicates that mutations to the *Antp* coding sequence is involved in post-transcriptional modifications in leg morphology and body plan. *Antp* in *Drosophila* interacts with *Dll* and active *Dll* is involved in limb formation amongst other morphological features (172). Unlike *Drosophila*, *Daphnia* *Antp* holds a mutation involved in post-transcriptional repression of *Dll* function, indicating a protein evolutionary modification change leading to wide morphological adaptations (177). Studies of this nature question the validity of the importance of cis-regulatory evolution of morphological changes between ancestral species. In this study a highly conserved protein is implicated in functional modulation of diversified chemoreceptors and it is found that the expression of this protein is dependent on conserved regulatory elements between non-related co-expressed genes. Both branches of this debate have a common question in mind: How does evolution generate organismal novelty? Future evolutionary studies on insect chemosensation may provide an adequate model exemplifying both evolution of gene coding and cis-regulation in a marriage of two mechanisms leading to both wide ancestral changes and narrow modifications in closely related species.
5.5 Summary: Insect Chemosensory Evolution Revisited

Cis-regulation occurs through the stochastic, electrostatic binding of developmental TFs to either enhancers or repressors elements upstream of the coded gene. The pattern of TF binding directs the rate of transcription at specific times and tissues. The nature of these elements allow for a level of degeneracy, controlling the level of binding affinity, also indicating a higher rate of neutral evolution in some species leading to alteration of expression patterns and possible adaptive changes.

Evolutionary changes with cis-regulation occur through point mutations, changing the percentage of affinity or complete specificity of the regulatory sequence, another mechanism proposes co-opted regulatory sequences through transposition or crossing over distributing common elements in number and position. Evo-devo proposes that the non-pleiotropic changes in cis-regulation would be sufficient to explain the divergence of morphological features within species. However, this does not account for changes within the TF gene coding sequence, resulting in changes of binding specificity or post-transcriptional functionality.

Insect chemosensory evolution exemplifies broad changes of chemosensory make-up and repertoire from ancestral crustacean species to terrestrialization of the
hexapoda lineages, expansion of convergently evolved ORs in flying insects and eusocial insects. At the same time, comparative analysis of closely related Drosophila species demonstrates a higher level on conservation, however demonstrating diverse behavioral adaptations. A recent study using a population genomic approach on the evolution of the Drosophila genome has found that de-novo genes from ancestral intergenic unexpressed open reading frames, are rapidly occurring within the Drosophila genome (178). In the insect chemosensory model ORs were believed to arise de-novo with an apterygote hexapod species expressing three OR co-receptor ORCO genes amongst the diverged ancient IR chemoreceptors. If ORs appeared through de-novo gene evolution two evolutionary paths may occur: 1) Directional selection may occur in the coded sequence resulting in a novel functionality. 2) Neutral evolutionary forces may result in loss of the gene through pseudogenization. A newly acquired functional gene may expand through further gene duplication events. Zhao et al, 2014 further demonstrated that evolving 5’-UTR regulatory sequences influenced de novo gene functionality. Within the de novo or gene duplication birth-and-death model of evolution of the ORs adaptive functional changes may occur through the change in temporal or spatial expression of that gene,
resulting in either increased fitness or decreased fitness within the organism. In an increased fitness hypothesis the cis-regulatory region would direct gene expression within a novel cell. In the case of chemosensory proteins this may result in an expansion of the olfactory or gustatory code permitting the exploitation of a novel environment. However, novel expression patterns may lead to redundant functionality or a detrimental loss of a gene. Within closely related Drosophila species the OR repertoire is relatively unchanged compared to more distantly evolved species. While gene expansion events might be constrained, evolution at the level of cis-regulation might influence the re-organization of chemosensory protein networks. The highly conserved and broadly expressing nature of the SNMPs might act by stabilizing a rapidly evolving network of chemosensory proteins permitting flexibility in chemosensory evolution. Future studies would require a comparative approach of SNMP expression profiles within closely and distantly related species. This study in conjunction with mapping of identified cis-regulatory motifs would be helpful in determining an evolutionary role for ancient proteins in stabilizing physiology.
Figure 5.1: Transcriptional Regulation.

A. The 5’ upstream region from a coding gene is made up of the core promoter and upstream elements. Elements are made up of groups of repressors, enhancers, and insulators. The core promoter is made up of highly conserved elements for the binding of the RNA Pol II holoenzyme (Transcriptional machinery). The core promoter is also associated with highly conserved regulatory motifs. Upstream elements may be found far upstream from the core promoter, promoting folding of the nucleotide backbone. Upstream repressors might affect the folding by enhancer bound transcription factors (TFs).

B. Transcriptional regulation occurs through either cis- or trans- mechanisms. Cis-regulation is any transcriptional control mediated by elements bound to proximal non-coded DNA. Trans- factors are many, this example demonstrates a scenario of post-translation modification of transcription.
Proposed stochastic mechanism of TF binding takes place in three-dimensional diffusion or one-dimensional diffusion along the DNA backbone. A TF may interact multiple times with the DNA backbone being repelled or attracted depending on the specific affinity of the TF binding locus and nucleotide binding sequence. Low-affinity binding may occur in non-regulatory regions of DNA, eventually reaching a point of equilibrium on the binding element.

Figure 5.2: Proposed mechanism of TF binding.
Direct or indirect mutational modification of a series of elements may cause changes in the specificity of transcriptional regulation. A. Direct modifications occur in one or a series of point mutations changing the binding affinity for TF to that particular element. Indirect modifications may occur through transposition of regulatory elements, resulting in broad modification to the expression of a gene. The acquired regulatory elements may further go through direct modifications resulting in a complete novel pattern of expression for that gene. Point mutations to a gene, in this case a TF, may result in a change of binding affinity to a regulatory element, causing regulatory change through a trans-acting factor.
REFERENCES


of Physiology - Regulatory, Integrative and Comparative Physiology.


111. Shahnaz Rahman L & Vijay Kumar S (2012) Or47b receptor neurons mediate sociosexual interactions in the fruit


129. Elina I & Michael GR (2011) Animal communication: flies’ ears are tuned in. *Current biology : CB.*


136. Vogt R (Biochemical Diversity of Odor Detection: OBPs, ODEs and SNMPs. *Biochemical Diversity of Odor Detection: OBPs, ODEs and SNMPs.*


transcription factor binding profiles. *Nucleic Acids Research*.


159. Dong P, Dicks J, & Panganiban G (2002) Distal-less and homothorax regulate multiple targets to pattern the


