12-10-2020

Gene Expression of Candidate Chemoreceptor Protein Families in Transcriptomes of Two Major Chemosensory Organs and Brain in Decapod Crustaceans

Mihika T. Kozma
Georgia State University

GENE EXPRESSION OF CANDIDATE CHEMORECEPTOR PROTEIN FAMILIES IN TRANSCRIPTOMES OF TWO MAJOR CHEMOSENSORY ORGANS AND BRAIN IN DECAPOD CRUSTACEANS.

by

MIHIKA TOTTEMPUDI KOZMA

Under the Direction of Charles D. Derby, PhD

ABSTRACT

Chemoreceptor proteins are necessary for animals to detect chemical signals and cues in their environment in a process known as chemical sensing. The diversity and number of chemoreceptor proteins have been characterized in many groups of animals, but few have studied the repertoire of chemoreceptor proteins expressed by decapod crustaceans. Crustaceans express at least three classes of putative chemoreceptor proteins. These are: Variant Ionotropic Receptors (IRs), derived from the ionotropic glutamate receptors (iGluRs); Transient Receptor Potential (TRP) channels, a diverse set of sensor-channels; and Gustatory Receptor Like receptors (GRLs), a family of ionotropic

receptor proteins that are ancestral to Gustatory Receptors (GRs) of insects. IRs are typically the most numerically dominant of these receptor proteins in crustaceans.

In order to identify families of candidate chemoreceptor proteins that are expressed by decapod crustaceans, I examined and compared transcriptomes from four decapod crustaceans that are established models of chemoreception: the Caribbean spiny lobster *Panulirus argus*, the clawed lobster *Homarus americanus*, the red swamp crayfish *Procambarus clarkii*, and the blue crab *Callinectes sapidus*. Transcriptomes were generated from: a) two major chemosensory organs, the lateral flagella of the antennules (LF) and dactyls of the walking legs (dactyl), of all four decapod crustaceans; and b) the supraesophageal ganglion (brain) of only three decapod crustaceans, *P. argus*, *H. americanus*, and *P. clarkii*. Each species expressed genes for at least ca. 100 to 250 IRs, ca. 15 TRP channels including those shown to be chemoreceptors in other species, and 1 to 4 GRLs. The IRs show different degrees of phylogenetic conservation: protostome-conserved, arthropod-conserved, pancrustacean-conserved, crustacean-conserved, and species-specific. Many IRs appear to be more highly expressed in the LF than dactyl. In the brain transcriptomes, few IRs, almost all TRP channels, and GRLs (in the case of *H. americanus*) were also detected. Immunocytochemistry in LF and dactyl of *P. argus* and *H. americanus*, revealed protein expression of co-receptor IR, IR25a, in olfactory sensory neurons and chemosensory neurons. This research lays the foundation for future functional studies by showing that decapod crustaceans have an abundance of gene expression for chemoreceptor proteins of different types, phylogenetic conservation, and expression patterns.

INDEX WORDS: Chemical sensing, Chemosensory neuron, Chemoreceptor proteins, Crab, Crayfish, Crustacean, Gustatory receptor, Ionotropic receptor, Lobster, Olfaction, Olfactory sensory neuron, TRP channel, Transcriptome

GENE EXPRESSION OF CANDIDATE CHEMORECEPTOR PROTEIN FAMILIES IN TRANSCRIPTOMES OF TWO MAJOR CHEMOSENSORY ORGANS AND BRAIN IN DECAPOD CRUSTACEANS.

by

MIHIKA TOTTEMPUDI KOZMA

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2020

GENE EXPRESSION OF CANDIDATE CHEMORECEPTOR PROTEIN FAMILIES IN TRANSCRIPTOMES OF TWO MAJOR CHEMOSENSORY ORGANS AND BRAIN IN DECAPOD CRUSTACEANS.

by

MIHIKA TOTTEMPUDI KOZMA

Committee Chair: Charles D. Derby

Committee: Daniel N. Cox

Paul S. Katz

Manfred Schmidt

Adriano Senatore

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

December 2020

For Ken, the love of my life and my truest friend.

For my parents and sister, who always believe in me.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Charles Derby, for being an excellent mentor and friend. Chuck, thank you for teaching me how to be a good scientist and for introducing me to the wonderful world of neuroethology and comparative biology. You are a constant source of inspiration to me and it has been a great pleasure to be your student. Your unwavering support even through my lowest points has meant a lot to me. Thank you for letting me hop (literally) into your office anytime for a conversation. I could always be my oddball self around you and express my ideas freely. I will always cherish our long and many conversations on matters related to research, science, life, great music, terrible movies, baseball, or college football. And thanks for not judging me too harshly for THE college football team I support. Derby Lab has been and always will be my family. Thank you for being my advisor!

Thank you to the members of my dissertation committee. To Dan Cox, thank you for giving me constant encouragement in my work and life. Thank you for your constructive comments on my research and for our many discussions on sensory systems of arthropods. Your work and our conversations inspired me to search for TRP channels in decapod crustaceans. To Paul Katz, thank you for being my teacher, being on my qualifying committee, and for encouraging me when I needed it. Thank you for bringing transcriptomics to the NI and for the 2CI Neurogenomics Fellowship, it helped my research tremendously. To Manfred Schmidt, thank you for being a wonderful mentor and friend. Learning from you and working with you has been a great pleasure. Our many conversations on science and life made me look at things from a different perspective. I will always be one of the biggest fans of your microscopy work. I am also lucky to have

seen your dance moves, they are quite groovy! To Adriano Senatore, thank you for being an amazing mentor and collaborator. None of my dissertation research would have been possible without your support. I am extremely grateful for the time and effort you took in training me. You are a role-model to me on how to mentor students. Thank you for always keeping your door open for me.

Thank you to the past and present members of Derby Lab, your friendship made work life so much better. To Vivian Ngo-Vu, thank you for being such a great friend and sister to me. Thank you for training me in ICC and so many more skills. I will never forget the countless hours we spent together to extract and purify RNA from pesky crustaceans. To my former mentees: Lanna Wolfe, Neal Shukla, and Matt Rump, thank you for your hard work and your friendship. It has been my privilege to publish with you as co-authors. To Shea Sparks, thank you for your beautiful ICC work and your friendship.

Thank you to my collaborators, Yan Wong and Shrikant Pawar for your help in analyzing transcriptome data. Thank you to Jonathan Boykin, Suranga Edirisinghe, David Sinkiewicz, Chris Deeb, and Randall Williford for your help in figuring out the bioinformatics world. To Emily Hardy, thank you making grad school manageable.

Thank you to my peers and friends for sharing my ups and downs. Grad school was so much more enjoyable because of you. Thank you for your love and friendship, it means a lot to me. I am lucky to have so many friends who care deeply for me. To my D&D crew, thank you for the good times and all the times we lost our heads. Adventure on!

I would like to express my deep gratitude to my whole family. Thank you for your constant love and support. To the Deebs, I am so lucky to have you not only as my family, but as my dearest friends. Thank you for always having my back and loving me no matter what. To Andrea and Randall, you are my home away from home. How lucky we are to have each other! Thank you for your friendship, your love, and for being my family. I am so proud to be Godmother (of Science!) to your children. Always.

I owe much in my life and career to my grandmother, Smt. Kolli Shakuntala. She gave me direction when I needed it the most. Her life and legacy are an inspiration to me, and I am honoured to have been her granddaughter. Rest in peace, Ammamma.

I am deeply indebted to my parents, Rama Devi and Harischandra Prasad, and my sister, Bhavya. Amma, Nanna, and Akka, thank you for your indomitable spirit, your strength, your sense of humour, and your unwavering love. Over the years, I have given you all many grey hairs with my antics, but somehow you still believed that I would find my way. Thank you for that belief, I would not be here without you.

Most of all, I would like to thank my husband, Ken. Kenny, you and our little family of critters have been the light of my life. Without you, Tuk, Simon, Krenshaw, and Argus, I would have never made it through. “I know you can do it,” these little words from you gave me the strength I needed to fight the great big bags of lemons that life kept throwing at me. Thank you for loving me so well. Wu Tang!

– Mihika

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ATD, Amino-terminal domain

BLAST, Basic local alignment search tool

BUSCO, Benchmarking universal single-copy orthologs

CNS, Central nervous system

Csap, *Callinectes sapidus*

CSN/CRN Chemosensory neuron/chemoreceptor neuron

CTD, carboxyl-terminal domain

Dmel, *Drosophila melanogaster*

Dpul, *Daphnia pulex*

ENaC, Epithelial sodium channel

EVG pipeline, Evidentialgene pipeline

GPCR, G protein coupled receptor

GR, Gustatory receptor

GRL, Gustatory receptor-like

Hame, *Homarus americanus*

ICC, Immunocytochemistry

ICD, Ion channel domain

iGluR, Ionotropic glutamate receptor

IR, Variant ionotropic receptor

LBD, Ligand binding domain

LF, Lateral flagella of the antennules

MAFFT, Multiple alignment using fast Fourier transform

MSN/MRN, Mechanosensory neuron/mechanoreceptor neuron

NMDAr, N-methyl-D-aspartate receptor

OL, Olfactory lobe

OSN/ORN, olfactory sensory neuron/olfactory receptor neuron

OR, Odorant receptor

Parg, *Panulirus argus*

Pcla, *Procambarus clarkii*

PCR, Polymerase chain reaction

RNA-Seq, RNA sequencing

RSEM, RNA-Seq by expectation-maximization

SCT, Single cell transcriptome

TRP channel, Transient receptor potential channel

UFBoot, Ultrafast bootstrap

1 GENERAL INTRODUCTION

All forms of life on Earth from bacteria and plants to animals with complex behaviors evolved various sensory systems to monitor available resources and risks in their environment. Resources can include means of sustenance, a host, shelter, proximity and availability of mates, and companionship. Risks can include predators, competitive conspecifics, and natural elements. For an organism to survive and be successful in its environment it needs to find resources and evaluate risks in its environment with competence. Sensory systems form a critical first step in gathering such information. This need to sense the presence and location of light, chemicals, sound, pressure, temperature, and other environmental stimuli is one of the driving forces of evolution. Among the various sensory systems that evolved to detect environmental stimuli, chemical sensing is a critical ability that drives many behaviors of most organisms. These behaviors range from plants evolving flowering strategies to manipulate visitation order of pollinators (Tsuji et al. 2020) to food tracking behavior in animals, and beyond. Chemical stimuli in an environment broadly range from single compounds to complex mixtures of compounds released either accidentally or on purpose by an organism, or as a product of a natural occurrence (Derby and Sorensen 2008, Kamio et al. 2014). To detect these chemical stimuli, organisms have evolved and express several classes of receptor proteins within their sensory cells known as chemoreceptor proteins. My dissertation is a study of gene expression of the different classes of chemoreceptor proteins that are expressed by four decapod crustaceans, Caribbean spiny lobster *Panulirus argus* (Latreille, 1804), American clawed lobster *Homarus americanus* (Milne-Edwards, 1837), red swamp crayfish *Procambarus clarkii* (Girard, 1852), and blue crab *Callinectes sapidus* (Rathbun,

1896), within their chemosensory organs, the identity and evolutionary history of these proteins, and their expression in sensory cells. Studying chemoreceptor protein expression in decapod crustaceans will bring insights to the first steps of chemical stimuli detection, organization of the peripheral and central nervous system, and consequently behavior mediated by chemical sensing.

1.1 Chemosensory systems of decapod crustaceans

Decapod crustaceans have two broad chemosensory systems: 'olfaction' and 'distributed chemoreception.' Olfaction is the process of sensing distant chemical signals and cues that mediate behavior. Distributed chemoreception encompasses taste and other chemoreceptor events that are not mediated through olfaction. In decapod crustaceans, olfaction mediates courtship behavior in response to distant sex pheromones, agonistic interactions between conspecifics and across species, aggregation behavior, avoidance behavior which includes response to conspecific alarm cues, and individual recognition of conspecifics (Schmidt and Mellon 2011, Derby and Weissburg 2014). Antennular grooming and detection of contact sex-pheromones are mediated by distributed chemoreception (Schmidt and Derby 2005, Bauer 2010). While olfaction and distributed chemoreception have distinct chemical sensing functions, they also have several overlapping functions such as food search, associative odor learning, initiation of search behavior, and orientation towards distant food-related chemicals.

Decapod crustaceans are covered in chemosensory sensilla that belong to these two chemosensory systems. Organs that include chemosensory sensilla are antennules (1st pair of antennae), 2nd antennae, mouthpart appendages, walking legs, tail fan, gill chamber, and pleopods (Schmidt and Mellon 2011, Derby et al. 2016). Chemosensory

sensilla are also broadly categorized into two groups: aesthetascs and bimodal sensilla. While aesthetascs are the only members of their group, there are many types of bimodal sensilla such as guard hairs, asymmetric setae, hooded sensilla, funnel-canal organs, etc. (Schmidt and Mellon 2011, Derby and Weissburg 2014).

Aesthetascs have only one morphological type and are only found on the distal ends of the lateral flagella (LF) of the antennules. In the example of *P. argus*, aesthetascs are organized into parallel rows of ~ 20 aesthetascs per annulus of the LF. Aesthetascs are thin, long, and unimodal sensilla made of porous cuticle that allows exchange of molecules between the environment and lumen of the aesthetascs but filters out molecules larger than 8.5 kDa (Schmidt and Gnatzy 1984, Derby et al. 1997). Aesthetascs are innervated by finely branched dendrites of bipolar olfactory sensory neurons (OSN) which mediate 'olfaction.' Depending on the species of decapod crustacean, each aesthetasc is innervated by 50 – 400 bipolar OSNs (Schmidt and Mellon 2011). The somata of the OSNs form clusters and are housed within the annuli of the LF underneath the aesthetascs. Axons from OSNs project through the length of the LF and directly terminate into the glomeruli of the olfactory lobe (OL) (Sandeman et al. 1992), which are paired neuropils in the deutocerebrum of a decapod crustaceans brain. OLs are comprised of small and dense wedge-shaped clumps of neuropil known as glomeruli, and a fiber core. The size of OLs and the number of glomeruli in each OL varies depending on the decapod species; Spiny lobster *P. argus* has very large OLs with ~1300 glomeruli, while crayfish *P. clarkii* has very small OLs and ~500 glomeruli (Schmidt 2016).

The other group of chemoreceptive sensilla is the bimodal sensilla and include all chemoreceptive sensilla in decapod crustaceans except aesthetascs. Bimodal chemoreceptive sensilla are comprised of multiple morphological types and mediate

'distributed chemoreception' (Schmidt and Mellon 2011). In the case of *P. argus*, on antennules alone, there are nine morphological types of bimodal sensilla such as guard hairs, companion setae, asymmetric setae, hooded sensilla, plumose sensilla, short setuled sensilla, short simple sensilla, medium simple sensilla, and long simple sensilla (Cate and Derby 2001, Derby and Weissburg 2014). These sensilla typically are shafts with a terminal pore at or near the tip of the shaft and a scolopale body located below the base of the setae in the annulus. These sensilla are called bimodal as they are innervated by chemosensory neurons (CSN) and mechanosensory neurons (MSN). Bimodal sensilla in decapod crustaceans are typically innervated by 1 – 22 CSNs and 1 – 3 MSNs. Dendrites from these sensory neurons are unbranched. Axons from bimodal sensilla on antennules terminate into the striated second antenna neuropils, bilaterally paired lateral antennular neuropils (LAN) and unpaired median antennular neuropil (MAN). Axons from bimodal sensilla on legs terminate into leg neuromeres of the central nervous system (CNS) and axons from bimodal sensilla on all other body parts terminate into other neuropils in the brain and CNS. These neuropils are also local motor centers for the appendages from which the axons originate (Sandeman et al. 1992, Schmidt and Mellon 2011, Schmidt 2016).

1.2 Chemoreceptor proteins of decapod crustaceans

A critical part of chemical sensing is the expression and function of receptor proteins in OSNs and CSNs. Chemoreceptor proteins are necessary for making 'first contact' with chemicals and form one of the initial stages of a chemosensory event. Chemicals in an environment can come into contact with chemoreceptor proteins expressed in sensory neurons either by active sensing, such as voluntary antennular

flicking by crustaceans to probe their environment, or passive sensing by an animal. The classes of chemoreceptor proteins that are expressed in sensory neurons determine downstream effects in signaling cascades and are responsible for mediating behavior of an animal in response to a particular chemical stimulus in its environment. Therefore, knowing the identity of chemoreceptor proteins that are expressed in OSNs and CSNs is essential for the study of chemical sensing in animals. Research on discerning the identity of chemoreceptor proteins that are expressed in OSNs and CSNs of decapod crustaceans has lagged compared to similar research in other arthropods such as insects and other protostomes such as nematodes and molluscs. Based on expression in other arthropods and protostomes, several classes of chemoreceptor proteins emerge as candidates. Chief among them are G-protein coupled receptors (GPCR), gustatory receptor-like(GRL)/gustatory receptors (GR)/odorant receptors (OR) family, variant ionotropic receptors (IR), and transient receptor potential (TRP) channels (Derby et al. 2016).

GPCRs are one of the most widely studied and well described class of metabotropic chemoreceptor proteins in animals. They were first discovered and classified in rats (Buck and Axel 1991). GPCRs are the predominant class of chemoreceptor proteins in chordates. Some protostomes such as nematodes and molluscs have independently evolved GPCR chemoreceptor proteins (Thomas and Robertson 2008, Cummins et al. 2009a, Cummins et al. 2009b, Albertin et al. 2015). Among decapod crustaceans, there is evidence for second messengers of GPCR signaling cascades in OSNs (Boekhoff et al. 1994, Hatt and Ache 1994, McClintock et al. 2006), but little support for GPCRs as primary chemoreceptor proteins. GPCRs as chemoreceptor proteins are largely understudied in arthropods with a notable and recent exceptions, such as non-visual opsins playing a role in mediating bitter sensing in the fruit fly, *Drosophila melanogaster* (Leung et al. 2020).

GRL/GR/OR family is a class of 7-transmembrane chemoreceptor proteins that are presumed to be ionotropic receptors (Robertson et al. 2003, Silbering and Benton 2010, Benton 2015, Freeman and Dahanukar 2015, Saina et al. 2015). GRL are the most ancient of this family as they were detected in Placozoa, Cnidaria, in deuterostome phyla Hemichordata and Echinodermata, and across protostomes (Saina et al. 2015). So far, GRLs are not known to be involved in chemical sensing. GRs are detected across arthropods including in chelicerates and myriapods (Chipman et al. 2014, Egekwu et al. 2014), and are well described chemoreceptor proteins especially in insects. The most recently evolved of this family of chemoreceptor proteins are ORs, which are predominantly found in OSNs of insects and mediate olfaction (Peñalva-Arana et al. 2009, Groh-Lunow et al. 2014, Benton 2015, Robertson 2015, Eyun et al. 2017, Kozma et al. 2020).

An ancient class of receptor proteins called ionotropic glutamate receptors (iGluR) is found in nearly all organisms from plants to animals. iGluRs form homo- and heterotetrameric ionotropic receptor channels (Mayer and Armstrong 2004, Mayer 2011). Subclasses of iGluRs include *N*-methyl-D-aspartate NMDA, α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA), and kainate receptors that are activated by glutamate (co-activated by glycine in the case of NMDA). Another subclass of iGluRs are the variant ionotropic receptors (IR) that were first described by Benton et al. (2009). Variant IRs function as chemoreceptor proteins in insects (Benton et al. 2009, Silbering et al. 2011) and are considered as candidate chemoreceptor proteins in other protostomes as well (Croset et al. 2010, Corey et al. 2013, Groh et al. 2013, Eyun et al. 2017, Vizueta et al. 2018, Kozma et al. 2020). Variant IRs are structurally similar to iGluRs and share highly conserved domain regions (Benton et al. 2009, Rytz et al. 2013). IRs also form

heterotetrameric channels but require at least one sub-unit of the protein to be a co-receptor IR to be a functional receptor channel (Abuin et al. 2011, Rytz et al. 2013, Abuin et al. 2019). Most co-receptor IRs are highly conserved across arthropods with one co-receptor IR (IR25a) being conserved across protostomes (Croset et al. 2010, Rytz et al. 2013).

TRP channels are 6-transmembrane cation channels that function as multimodal sensors in most animals. A single TRP channel can be activated by several different modes of sensory stimuli such as temperature, chemicals, light, and pressure. In chemical sensing, TRP channels are mostly involved in directly or indirectly mediating avoidance of aversive compounds (Venkatachalam and Montell 2007, Venkatachalam et al. 2014). The exact mechanism of activation of TRP channels by the different sensory modalities remains under study, but an emerging theory is that reactive oxygen species released from cell damage caused by aversive stimuli activate these channels (Guntur et al. 2015, Arenas et al. 2017). While TRP channels have been detected in transcriptomes of the central nervous system of decapod crustaceans, the function of these channels in decapods remains unknown (McGrath et al. 2016, Northcutt et al. 2016).

Learning the gene expression patterns of these chemoreceptor protein families in OSNs and CSNs of decapod crustaceans will give us insights into the molecular machinery of these sensory neurons and facilitate our understanding of how chemosensory neurons in decapod crustaceans respond to different chemical signals and cues. Therefore, identification and characterization of chemoreceptor molecules will lead to better understanding of the functional organization of chemosensory systems in decapod crustaceans. Using RNA-sequencing and *de novo* transcriptome assembly, we generated transcriptomes of two chemosensory organs, lateral flagella (LF) and dactyls of walking

legs (dactyls), from four decapod crustaceans: *P. argus*, *H. americanus*, *P. clarkii*, and *C. sapidus*. We also generated transcriptomes from the brains of *P. argus*, *H. americanus*, and *P. clarkii*. Chapter 2 (Kozma et al. 2018) focuses on identifying and characterizing genes that are expressed in these *P. argus* transcriptomes. By building maximum likelihood phylogenetic trees with chemoreceptor protein sequences from other animals, candidate chemoreceptor proteins belonging to variant IRs, TRP channels, and GRs that are expressed by *P. argus* were identified and classified into subtypes of a protein family based on sequence homology. Using immunocytochemistry experiments we detected protein expression of IRs in OSNs and CSNs of LF and dactyls in *P. argus*. We also detected protein expression of IRs in large cells located in the axon sorting zone in the brain of *P. argus*. Chapter 3 (Kozma et al. 2020) extends this work to the transcriptomes of *H. americanus*, *P. clarkii*, and *C. sapidus*, using annotated genes of candidate chemoreceptor proteins from P. argus transcriptomes as references. In this chapter we generated de novo transcriptomes using the evidential gene pipeline and produced high quality transcriptome assemblies for all four decapod species. Despite coming from diverse environmental conditions and belonging to different families within Decapoda, the four decapod crustaceans had high conservation in the classes of candidate chemoreceptor proteins expressed. Variant IRs were the most predominant and highly expressed genes by all four decapods, followed by TRP channels. While we detected GRs/GRL in these transcriptomes, they typically had very low expression and fragmented sequences. We also identified several GPCRs in these transcriptomes, but most were categorized into neuromodulator GPCRs. By building maximum likelihood phylogenetic trees, we defined levels of conservation of variant IRs from the species-specific level to inter-phyla. Although our data only used transcriptomes, we attempted to determine the

number of genes expressed by each class of candidate chemoreceptor proteins. Variant IRs numbered in the hundreds, ~ 15 TRP channels, and less than 5 GRs were detected for all four decapod crustaceans. This body of work presents a long and diverse list of genes to be examined as potential chemoreceptors in decapod crustaceans.

2 CHEMORECEPTOR PROTEINS IN THE CARIBBEAN SPINY LOBSTER, *PANULIRUS ARGUS*: EXPRESSION OF IONOTROPIC RECEPTORS, GUSTATORY RECEPTORS, AND TRP CHANNELS IN TWO CHEMOSENSORY ORGANS AND BRAIN

Previously published as: Kozma, M. T., M. Schmidt, H. Ngo-Vu, S. D. Sparks, A. Senatore and C. D. Derby (2018). PLoS One **13**(9): e0203935.

Author contribution: Kozma, M.T. designed experiments, collected and analyzed data, and wrote the paper. Schmidt, M. designed experiments, collected and analyzed immunocytochemistry data, and wrote the paper. Ngo-Vu, H. and Sparks, S.D. collected data. Senatore, A. contributed to data analysis and editing of manuscript. Derby, C.D. designed experiments and wrote the paper.

Supplemental Information available at: https://doi.org/10.1371/journal.pone.0203935

Acknowledgements: We thank Dr. Timothy McClintock for the antibody to IR25a, Dr. Donald Behringer for providing lobsters, Dr. Edirisinghe for technical guidance on HPC, generating R scripts for transcript abundance estimation and heatmaps, and guidance and discussions on using Trinity transcriptome assembly software, Neal Shukla for assistance with analysis of TRP channels, and Drs. Dan Cox and Paul Katz for many helpful discussions. Funding was provided by a seed grant from the Georgia State University's Brains & Behavior program and fellowships from Georgia State University's 2CI Neurogenomics program, Center for Neuromics, and Neuroscience Institute.

2.1 Introduction

Acquiring environmental cues is key to the survival of animals, since it informs them about the location and quality of food, mates, predators, shelter, and other resources and risks. The first steps in detecting and discriminating environmental chemicals are performed by chemoreceptor cells, which possess the receptor proteins that bind the chemicals. Characterizing these receptor proteins is fundamental to understanding chemosensory transduction and, more broadly, mechanisms of chemical sensing.

Although crustaceans are one of the largest and most diverse animal taxa with nearly 70,000 extant species living in diverse environments (Ahyong et al. 2011, Schram 2012), relatively little is known about their chemoreceptor proteins. Given that crustaceans are well-established models of chemoreception (Ache 2002, Stensmyr et al. 2005, Schmidt and Mellon 2011, Derby and Weissburg 2014), the lack of data on their receptor proteins has limited our understanding of the organization of their chemical senses at the cellular and molecular levels. Crustaceans, in particular decapod crustaceans such as Caribbean spiny lobster, *Panulirus argus,* have two major chemosensory systems – olfaction and distributed chemoreception – which differ in their peripheral and central organization (Schmidt and Ache 1996, Schmidt and Ache 1996, Schmidt and Mellon 2011, Derby and Weissburg 2014, Derby et al. 2016) (Figure 2.1). Olfaction is mediated by aesthetasc sensilla on the distal end of the lateral flagella of the antennules (first antennae). Aesthetascs are unimodal, being innervated only by olfactory receptor neurons (ORNs). More than 300 ORNs can innervate a single aesthetasc. Distributed chemoreception includes gustation plus other chemical senses except for olfaction. The sensilla in the distributed chemoreception pathway are diverse in form and are found not only on the antennules but also on the second

antennae, mouthparts, legs, and other parts of the body. Despite structural diversity, the distributed chemosensilla have a fundamental characteristic: they are bimodal, being innervated by chemoreceptor neurons (CRNs) and mechanoreceptor neurons (MRNs). Chemoreceptor cells in olfaction and distributed chemoreception also differ in their central projections. ORNs project to the brain's olfactory lobes, which have a glomerular organization (Schmidt and Mellon 2011). CRNs project to neuropils in the brain, subesophageal ganglion, and thoracic and abdominal ganglia, which have a somatotopic organization and also serve as local motor centers for the appendages providing sensory input to the respective neuropils (Schmidt and Ache 1996, Schmidt and Ache 1996, Schmidt and Mellon 2011).

Olfaction and distributed chemoreception have some overlapping chemosensitivity and function, though they also detect different chemicals and mediate different behaviors. For example, both olfactory and distributed chemoreceptors on the antennules detect some of the same food-related chemicals, including representatives of amino acids, amines, nucleotides, organic acids, and other molecules. The detection of these chemicals drives the orientation of animals toward the source of the chemicals (reviewed in (Schmidt and Mellon 2011, Derby and Weissburg 2014)). Distributed chemoreceptors on the legs and mouthparts detect these same food-related molecules, yet stimulation of them drives different behaviors such as grabbing and ingestion of food or rejection of food, and stimulation of distributed chemoreceptor neurons in asymmetric sensilla on the antennules mediates grooming behavior (Schmidt and Derby 2005). However, olfaction and distributed chemoreception differ in their sensitivity to intraspecific chemical cues. Olfaction is largely responsible for detecting waterborne conspecific chemicals such as sex pheromones, social cues, and alarm cues, and then mediating

various behaviors driven by these cues (Breithaupt and Thiel 2010, Kamio et al. 2014). Despite these known differences in olfaction and distributed chemoreception in crustaceans, the molecular basis for differences at the receptor level is unknown.

A phylogenetic and evolutionary approach to the study of chemoreceptor proteins has revealed that animals have adopted many different types of proteins as chemoreceptors, including a diversity of ligand-gated ion channels and metabotropic receptors (reviews: (Benton 2015, Joseph and Carlson 2015, Robertson 2015, Derby et al. 2016, 2016, Wicher and Groβe-Wilde 2017)). The chemoreceptor proteins used by Protostomia and Deuterostomia are largely different from each other (Derby et al. 2016). Protostomes predominately use ionotropic receptors as chemoreceptors, though GPCRs have been shown to function in a few protostomes, most notably in the nematode *Caenorhabditis elegans* (Thomas and Robertson 2008, Nordstrom et al. 2011, Krishnan et al. 2014), the sea hare *Aplysia californica* (Cummins et al. 2009a, Cummins et al. 2009b), and the crown-of-thorns starfish (Roberts et al. 2017). In this study, we investigate the major classes of chemoreceptor proteins in the Protostomia, including Ionotropic Receptors (IRs), gustatory receptors (GRs), and transient receptor potential (TRP) channels.

An ancestral type of ionotropic receptor present in all Protostomia, and not present in the Deuterostomia, evolved from ionotropic glutamate receptors (iGluRs) and are thus called Variant Ionotropic Receptors, or simply Ionotropic Receptors (IRs) (Benton et al. 2009, Croset et al. 2010, Abuin et al. 2011, Rytz et al. 2013). The iGluRs form homotetrameric ion channels, with each monomer consisting of four main domains (Mayer and Armstrong 2004, Mayer et al. 2006, Mayer 2011): an extracellular amino-terminal domain (ATD) involved in assembly of the heteromeric channel; an extracellular

ligand binding domain (LBD) consisting of two half-domains (S1 and S2) to which L-glutamate, glycine, or serine agonists bind; an ion channel domain (ICD) that forms the ion channel, consisting of three transmembrane domains (M1, M2, M3) and a pore loop (P); and an intracellular carboxyl-terminal domain (CTD) (Mayer et al. 2006) (Figure 2.1).

There are several families of IRs, including co-receptor IRs, conserved IRs, and divergent IRs. Co-receptor IRs – IR8a, IR25a, IR76b, and IR93a – are co-expressed with other IRs in cells and are necessary for the function of those receptor-channels. Out of these, IR25a and IR8a are the IRs most closely related to iGluRs in that they possess all four domains of the iGluRs. With the exception of IR25a and IR8a, IRs contain only three of four domains of iGluRs, either lacking or having a truncated ATD region (Figure 2.1). The other families of IRs are the conserved IRs and divergent IRs. Conserved IRs are IRs present across insects, crustaceans, and in the case of the co-receptor IRs IR25a and IR8a, all protostomes examined so far (Benton et al. 2009, Croset et al. 2010, Rytz et al. 2013, Eyun et al. 2017). Conserved IRs include IR21a, IR31a, IR40a, IR41a, IR60a, IR64a, IR68a, IR75, IR76a, IR84a, and IR92a. Due to the presence of their homologues in other species, all four co-receptor IRs are also considered to be conserved IRs. However, divergent IRs are species-specific IRs with no known homologues. When co-expressed with co-receptor IRs, the conserved and divergent IRs form functional heteromeric channels with ligand-specific binding properties that depend on the specific divergent IR that is expressed (Silbering et al. 2011, Croset et al. 2016, Hussain et al. 2016, Pitts et al. 2017).

The number of IRs varies across species ranging from three in *C. elegans*, around 60 IRs in drosophilids, to over 120 in the termite *Zootermopsis nevadensis* (Benton et al.

2009, Croset et al. 2010), and other protostomes have more or less than these. Though IRs are best studied in Insecta, they have also been identified in other arthropods, including myriapods (centipedes: (Egekwu et al. 2014); millipedes: (Kenny et al. 2015)), chelicerates (ticks: (Rytz et al. 2013, Chipman et al. 2014)), and crustaceans (Croset et al. 2010, Eyun et al. 2017) (reviewed in (Derby et al. 2016), and see below). The genome of the water flea, *Daphnia pulex*, has 85 IRs, though the anatomical location of these IRs is unknown (Croset et al. 2010, Rytz et al. 2013, Eyun et al. 2017). Some co-receptor IRs have been identified in a number of crustaceans including clawed lobsters, *Homarus americanus* (Hollins et al. 2003), spiny lobsters (Tadesse et al. 2011, Corey et al. 2013), hermit crabs (Groh et al. 2013, Groh-Lunow et al. 2014), seven species of shrimp (Zbinden et al. 2017), and copepods (Lenz et al. 2014, Núñez-Acuña et al. 2014, Núñez-Acuña et al. 2016, Eyun et al. 2017). The co-receptor IR25a is expressed in ORNs of the antennular lateral flagellum of *H. americanus* (Hollins et al. 2003, Stepanyan et al. 2004), spiny lobster *P. argus* (Tadesse et al. 2011, Corey et al. 2013), and hermit crab *Coenobita clypeatus* (Groh-Lunow et al. 2014), and it also appears to be expressed in other chemosensory organs (Corey et al. 2013, Zbinden et al. 2017). However, very little is known about IRs, especially the divergent IRs, in most crustaceans. Complete sequences of two divergent IRs (and additional partial sequences) were identified in *P. argus* (Corey et al. 2013), 16 divergent IRs were found in the hermit crabs *Pagurus bernhardus* (Groh et al. 2013), and 22 divergent IRs were identified in *C. clypeatus* several of which were shown to be expressed in ORNs (Groh-Lunow et al. 2014).

A major type of ionotropic chemoreceptor protein in Protostomia, apparently first appearing in insects and constituting a major class of receptors in them, is the Odorant Receptors (ORs) (Robertson et al. 2003, Benton et al. 2006, Joseph and Carlson 2015,

Robertson 2015, Saina et al. 2015). These ionotropic ORs are not homologous with the GPCR ORs of deuterostomes and are evolutionarily related to arthropod Gustatory Receptors (GRs) (Figure 2.1) (Scott et al. 2001, Robertson et al. 2003, Chipman et al. 2014, Ling et al. 2014, Benton 2015, Freeman and Dahanukar 2015) which are considered a more early-diverging class of ionotropic receptors. GR homologues ancestral to Arthropoda, for which chemosensory function has not been identified, are called Gustatory Receptor-Like proteins (GRLs). GRLs appeared early in metazoan evolution, at least in cnidarians and placozoans (Benton 2015, Robertson 2015, Saina et al. 2015, Eyun et al. 2017).

To date, despite efforts, GRs have rarely been identified in crustaceans and ORs not at all. *D. pulex* has 58 GRs but no ORs (Peñalva-Arana et al. 2009, Benton 2015, Robertson 2015, Saina et al. 2015), and their anatomical location or involvement in chemical sensing has not been demonstrated. Eyun et al. (2017) found a few GRs in some species of Copepoda and in a barnacle (Cirripedia). From their analysis, they concluded that GRs appeared early in metazoan evolution but expanded only in some groups: in the Arthropoda, the expansion was in Insecta and some Chelicerata but not most Crustacea.

Another class of receptors that are less explored in crustaceans is the transient receptor potential (TRP) channel superfamily. All TRP channels have six transmembrane regions and are cation channels. Based on sequence homology, TRP channels are classified into seven subfamilies belonging to two groups: group 1 includes subfamilies TRPA, TRPC, TRPM, TRPN, and TRPV; and group 2 includes subfamilies TRPML and TRPP (Figure 2.1) (Venkatachalam and Montell 2007, Venkatachalam et al. 2014). TRP channels can be activated by several mechanisms including sensory stimuli such as temperature, light, chemicals, sound, and touch. Members of group 1 have been shown

to be a part of chemosensory systems across animals including insects and nematodes. OSM-9 is a TRPV channel in *C. elegans* that mediates avoidance to bitter chemicals (Colbert et al. 1997, Upadhyay et al. 2016). TRPA1 is expressed in the labellum and mouthparts of *D. melanogaster* and detects aversive chemicals. TRPA1 also indirectly and directly mediates avoidance of citronella in flies and mosquitoes, respectively (Kang et al. 2010, Kwon et al. 2010). *painless*, a member of the TRPA sub-family, is a chemoreceptor that prevents male-male courtship and also mediates avoidance of deterrent compounds (Wang et al. 2011). Members of TRPC, TRP, and TRPL subfamilies are expressed in CO_2-detecting ORNs in *D. melanogaster* and contribute to CO_2 avoidance (Venkatachalam and Montell 2007, Venkatachalam et al. 2014).

Although their function has not been described in any crustacean, TRP channels are found in the few crustaceans examined. For example, *D. pulex* has 14 TRP channels representing all the subfamilies (Peng et al. 2015). Among the decapods, *H. americanus* has a combined eight TRP channels in the transcriptomes of its central nervous system (CNS) and heart (McGrath et al. 2016, Northcutt et al. 2016), the Jonah crab, *Cancer borealis*, has six TRP channels in its CNS transcriptome (Northcutt et al. 2016), and the hermit crab, *C. clypeatus*, has one potential TRP channel, a homologue of the *D. melanogaster* TRPN channel, NompC, in its antennule transcriptome. However, TRP channels were not detected in the antennule transcriptome of the marine hermit crab, *P. bernhardus* (Groh et al. 2013).

The goal of our study was to use transcriptomics to identify chemoreceptor proteins in the olfactory and distributed chemoreception systems of a major crustacean model of chemoreception, the Caribbean spiny lobster *Panulirus argus*, and to compare them to homologous chemoreceptor proteins in *Drosophila*, *Daphnia*, and other species

in major animal groups. We aimed to determine the number and diversity of receptor proteins of different types in olfactory sensilla bearing tissue (antennular lateral flagellum) and distributed chemoreception (leg dactyl) including whether these chemosensory organs express the same or overlapping sets of receptor proteins. We also examined the brain of *P. argus* due to the identification of chemoreceptor proteins, IR68a and IR75u, in the brain of the honey bee, *Apis mellifera* (Croset et al. 2010). To accomplish these goals, we performed next-generation sequencing of mRNA preparations of the antennular lateral flagellum, leg dactyl, and brain, assembled their transcriptome, and performed bioinformatics searches for molecules of interest. In a very conservative estimate, we identified 108 IRs, with more of these expressed in the antennular lateral flagellum than in the dactyl. Using immunocytochemistry, we showed that IR25a is expressed in all ORNs and most CRNs but not in MRNs. We found IR25a expressed in specific cells of the ORN axon sorting zone near the olfactory lobe in the brain. We also identified one GR and homologues from seven subfamilies of TRP channels. Our results show a diversity of putative chemoreceptor proteins in *P. argus*

2.2 Materials and Methods

2.2.1 Animals

Experiments were performed on male and female Caribbean spiny lobsters, *Panulirus argus*. Most animals were obtained from the Florida Keys Marine Laboratory, and some animals were kindly provided by Dr. Don Behringer (University of Florida). Animals were held at Georgia State University in communal 800-L aquaria or in

individual 10-L aquaria containing aerated, recirculated, filtered artificial seawater (Instant Ocean, Aquarium Systems, Mentor, OH) in a 12-hr:12-hr light:dark cycle. They were fed shrimp or squid three times per week.

2.2.2 Tissue collection and RNA isolation for generating transcriptomes

Tissue was collected from four animals: female 1 (F1) with carapace length 66 mm, weight 255 g; female 2 (F2) with carapace length 80 mm, weight 397 g; female 3 (F3) with carapace length 63 mm and weight 232 g; and male 1 (M1) with carapace length 77 mm and weight 364 g. Tissues were dissected from animals that were anesthetized in ice. Three tissues were collected, shown in Figure 2.1. The aesthetasc-bearing region of both antennular lateral flagella (LF) of F1 and M1 was collected and pooled. The dactyl of the right second walking leg (dactyl) leg was collected from F2. The supraesophageal ganglion, or brain, was collected from F3. For collecting LF and dactyl, the soft tissue within the cuticle was dissected out. For brain, the head was separated from the rest of the body and the brain was removed from the posterior aspect, with care being taken to prevent hepatopancreas from entering the head space during dissection. Immediately following dissection, collected tissues were instantly frozen in liquid nitrogen and stored at -80^{o}C.

Total RNA was extracted from these tissues using Tri-Reagent (Sigma-Aldrich, St. Louis, Missouri). Frozen tissues were weighed and then homogenized in Tri-Reagent using sterile disposable pestles. Chloroform was used for separation through centrifugation, and RNA was precipitated with ethanol. To ensure the quality of RNA and to remove any contamination with DNA, protein, or carbohydrate, the precipitated RNA

was reconstituted in diethylpyrocarbonate (DEPC)-treated water and again precipitated using lithium chloride. Next, the RNA was reconstituted in DEPC-treated water, and other contaminants such as sodium dodecyl sulfate (SDS) and SDS-bound proteins were precipitated out of the RNA using potassium acetate. Total RNA was then precipitated out of solution with ethanol and reconstituted in DEPC-treated water. The total RNA extracted for each tissue was quantified and checked for purity using NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts). Running agarose gel electrophoresis and staining the gels with ethidium bromide visualized RNA integrity. Aliquots of the total RNA extracted for each tissue were frozen over liquid nitrogen and stored at -80^{o}C.

2.2.3 Illumina sequencing, de novo *assembly, and transcript abundance estimation*

Total RNA extract for each tissue was diluted to 100 ng/µL in TE buffer (10 mM Tris, 1 mM EDTA, pH 7.5). Diluted RNA samples were instantly frozen in liquid nitrogen and shipped on dry ice to Beckman Coulter Genomics for quality assessment on Agilent Bioanalyzer2000 and TapeStation, mRNA specific cDNA synthesis, and paired-end sequencing of cDNA on Illumina HiSeq 2500 high-throughput sequencer. The read length was 2x100 (base pair reads) for LF and dactyl and 2x125 (base pair reads) for brain, and the number of reads per sample was > 120 million. Prior to data delivery, adapter sequences incorporated for tracking Illumina reads from multiplexed samples were removed.

All high performance computing (HPC) was performed on ORION and ACoRE HPC systems at Georgia State University (Sarajlic et al. 2016). A transcriptome was

generated by combining raw reads from the three tissues by following the *de novo* assembly protocol (Haas et al. 2013) for the program Trinity v2.6.5 (Grabherr et al. 2011) (Supplemental S1 Text: https://doi.org/10.1371/journal.pone.0203935.s004). Prior to assembly, reads were quality trimmed with default settings using the Trimmomatic software (Bolger et al. 2014) bundled into Trinity. The TransDecoder program (http://transdecoder.github.io/) was used to predict proteins with open-reading frames (ORFs) in each transcriptome. The predicted proteins for the transcriptome are referred to as 'Parg protein database.'

CD-Hit (Li et al. 2001) was performed on the transcriptome to remove redundancy. Benchmarking Universal Single-Copy Orthologs (BUSCO) v3 (Simao et al. 2015, Waterhouse et al. 2017) was run on the transcriptome before and after running CD-Hit to analyze the completeness of the assembly (Supplemental S2 Text: https://doi.org/10.1371/journal.pone.0203935.s005). Following removal of redundancy, the abundance of transcripts was estimated using RSEM, an alignment-based quantification software (Li and Dewey 2011). Custom 'R' scripts were used to extract the counts for each tissue for the transcripts of interest from the gene counts matrices that were generated using RSEM perl scripts bundled in Trinity, and were plotted using the 'heatmap.2' function in R (Supplementals: Figure 2.11 and Figure 2.12; Supplemental S2 and S3 Tables: https://doi.org/10.1371/journal.pone.0203935).

2.2.4 IR identification, sequence alignment, and phylogenetic analysis

iGluR and IR protein sequences from the common fruit fly (*Drosophila melanogaster*, Dmel), mosquito (*Anopheles gambiae*, Agam), water flea (*Daphnia pulex*,

Dpul), and Caribbean hermit crab (*Coenobita clypeatus,* Ccly) were collected from published data (Benton et al. 2009, Croset et al. 2010, Corey et al. 2013, Groh-Lunow et al. 2014) and NCBI databases. These were used as query sequences to screen for iGluRs and IRs in our Parg transcriptome and protein databases. NCBI-BLAST+ versions 2.3+ and 2.4.0 on ORION and ACoRE were used for tBLASTn and BLASTp searches on the Parg transcriptome and protein databases. Sequence hits with e-value < e^{-4} from the tBLASTn and BLASTp searches, along with Dmel and Dpul query sequences, were used as queries for PSI-BLAST searches against the Parg protein databases. These hits were then further screened for having both the ICD and LBD (Figure 2.1), which characterize iGluRs and IRs (Croset et al. 2010). This screening was performed with TMHMM v2.0 (Krogh et al. 2001) for transmembrane domain prediction and signature-screened against InterPro (Finn et al. 2017) with InterProScan 5 (v5.28-67.0) (Jones et al. 2014) for conserved Pfam domains. The ICD domain and S2 of the LBD were predicted by the presence of the Pfam domain, PF00060 (which contains M1, P, M2, S2, and M3, see Figure 2.1). S1 of the LBD was predicted by the presence of the Pfam domain, PF10613 (Croset et al. 2010, Finn et al. 2016).

To perform phylogenetic analysis, we used predicted protein sequences for putative IRs and iGluRs from the Parg protein databases containing both PF00060 and PF10613 domains (Supplemental S3 and S4 Text: https://doi.org/10.1371/journal.pone.0203935) and reference sequences of IRs and iGluRs from Dmel and Dpul. These sequences were aligned with MAFFT (Katoh et al. 2002, Katoh and Standley 2013) using default settings and visualized on Jalview (Waterhouse et al. 2009). Alignments were manually trimmed to remove gaps (Supplemental S5-S15 Text: https://doi.org/10.1371/journal.pone.0203935) guided by

amino acid conservation as annotated on Jalview. Sequences with < 150 amino acid residues that were missing large regions of the S1 or S2 domains of the LBD and ICD were removed from further analysis, with some exceptions. The best model of substitution was identified using ModelFinder (Kalyaanamoorthy et al. 2017), and maximum likelihood phylogenetic trees were constructed using IQ-Tree (Nguyen et al. 2015, Trifinopoulos et al. 2016). Confidence values for the trees were generated using ultrafast bootstrap (UFBoot) (Minh et al. 2013) integrated into IQ-Tree. The phylogenetic trees were visualized using FigTree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/), and color schemes were edited on Adobe Illustrator CS6, San Jose, CA.

Phylogenetic analysis of conserved IRs also included conserved IR sequences from several insect species: *Aedes aegypti* (Aaeg), *Culex quinquefasciatus* (Cqui), *Anopheles gambiae* (Agam), *Bombyx mori* (Bmor), *Tribolium castaneum* (Tcas), *Apis mellifera* (Amel), *Nasonia vitripennis* (Nvit), *Acrythosiphon pisum* (Apis), and *Pediculus humanus humanus* (Phum); and two gastropod molluscs: *Aplysia californica* (Acal) and *Lottia gigantea* (Lgig).

2.2.5 IR and iGluR nomenclature

IRs were named with a prefix of the species 'Parg-' and a random number assignment starting from 1000 and chronologically increasing. For example, the IRs were named PargIR1000, PargIR1001, and so on. Homologous sequences are named according to their IR homologues, e.g., PargIR25a, PargIR8a, and so on. NMDA-like iGluRs were named according to their homologues from Dmel and Dpul. Other iGluRs were named with random numbers.

2.2.6 GR identification and sequence alignment

GR sequences from Dmel, Dpul, and a copepod (*Eurytemora affinis*, Eaff) were used as queries in BLASTp searches of the Parg transcriptome. All hits against the query sequences were selected for further analysis even though their e-values were > e^{-4}. This was expected based on the high sequence divergence among GRs of other species (Peñalva-Arana et al. 2009, Robertson 2015, Saina et al. 2015). The transcriptome was also screened using InterProScan for the Pfam domain family, 7tm_7 (PF08395), since this family includes GRs and ORs from insects.

2.2.7 TRP channel identification and sequence alignment

TRP channel sequences from group 1 and group 2 subfamilies of Dmel, Dpul, Jonah crab (*Cancer borealis*, Cbor), American clawed lobster (*Homarus americanus,* Hame), nematode (*Caenorhabditis elegans,* Cele), mouse (*Mus musculus*, Mmus), and rat (*Rattus norvegicus,* Rnor) were used as queries in BLAST searches against the Parg transcriptome and protein database. TRP channel sequences in Dmel were extracted from www.flybase.org. TRP channel sequences from other insects (Bmor, Tcas, Amel, Nvit, Phum) were collected from Matsuura et al. (2009). Cele and mammalian sequences were extracted from publicly available databases on NCBI. TRP channel sequences of Dpul (Peng et al. 2015) were extracted from the JGI genome portal (https://genome.jgi.doe.gov/Dappu1/Dappu1.home.html). The rest were collected from NCBI databases. Only those sequences from the Parg transcriptome and protein database with e-values < e^{-4} to the queries were selected for further analysis. In parallel, we also screened the Parg transcriptome with InterProScan for TRP channel domain regions of different subfamilies. These included Pfam groups PF06011, PF08344, PF00520,

PF12796, PF00023, PF16519, and PF08016, which are TRP, TRP_2, ion transport, ankyrin 2, ankyrin, TRPM_tetra, and PKD channel domains, respectively. For maximum likelihood phylogenetic analyses, along with the query sequences, the target sequences were aligned with MAFFT (Supplemental S13-S15 Text: https://doi.org/10.1371/journal.pone.0203935), and maximum likelihood phylogenetic trees were constructed using IQ-Tree after ModelFinder assessed the best model of substitution as described above.

2.2.8 PCR

For PCR, total RNA was extracted from the following tissues of a male *P. argus* (549 g, 86.4 mm carapace length) after anesthetizing the animal on ice for about 20 min: (1) the distal, aesthetasc-bearing part of the LF (LFD); (2) the proximal, non-aesthetasc bearing part of the LF (LFP) (note that LFP lacks the 20 proximal-most annuli in the non-aesthetasc region, which contains the proximal proliferation zone and its developing aesthetascs (Steullet et al. 2000)); (3) flagella of second antennae (A2); (4) dactyls of second pereiopods (walking legs) (dactyl); (5) central brain (brain); and (6) green glands (GG). After anesthetizing the animal on ice for about 20 min, appendages and head were cut off. All appendages were wiped with 100% ethanol to remove attached epibionts before they and the remaining body parts were dissected further under autoclaved *P. argus* saline. From LF, A2, and dactyls, the internal tissue including the epithelium was removed from the surrounding cuticle. All tissues were instantly frozen in liquid nitrogen and stored at -80° C.

For extraction of total RNA, the appropriate amount of TRIzol (Thermo Fisher Scientific, Waltham, Massachusetts) was added to tissue samples (10 µl of TRIzol per mg of sample) and these were homogenized using disposable tissue grinders (Biomasher II, Kimble Chase, Vineland, New Jersey) or sterilized glass tissue homogenizers. Extracts were processed according to the TRIzol protocol until phase separation, and then RNA was purified using the Direct-zol RNA Mini Prep (Zymo Research, Irvine, California) including treatment with DNAse I to remove genomic DNA. About 1 µg of total RNA was used for cDNA synthesis with SuperScript III reverse transcriptase (Thermo Fisher Scientific - Invitrogen, Waltham, Massachusetts) and random hexamers following the manufacturer's instructions.

Primers for PCR (Supplemental Table 2.3) targeting the housekeeping gene GAPDH and the genes of interest from the Parg transcriptome (NMDA-R1, IR25a, IR8a, IR93a, IR1028, IR1074) were designed using Primer Blast (NCBI) and purchased from Integrated DNA Technologies (Coralville, Iowa). PCR amplification was performed in 50 µl reactions containing about 250 ng of cDNA with the DNA polymerase Phusion according to the manufacturer's instructions (Thermo Fisher Scientific, Waltham, Massachusetts) using the thermal cyclers Mastercycler SXI (Eppendorf, AG, Hamburg, Germany) and C1000 Touch (Bio-Rad Laboratories, Hercules, California). Annealing temperatures of primer pairs were calculated with an online Tm calculator (New England Biolabs, Ipswich, Massachusetts). PCR products were subjected to gel electrophoresis (1.5% agarose gel) together with a 100 bp DNA ladder (G-Biosciences, St. Louis, Missouri) for size determination.

To determine if PCR products matched the target sequences, bands of the expected size were cut from the agarose gel and cDNA was extracted from them using the PureLink

kit (Thermo Fisher Scientific - Invitrogen, Waltham, Massachusetts). Extracted cDNA was sequenced in the Georgia State University (GSU) molecular core facility.

2.2.9 Immunocytochemistry

For immunocytochemistry, LF, dactyls of pereiopods, second antennae, and brains of male and female *P. argus* (34–65 mm carapace length and 45–250 g in weight, n = 6) were dissected after anesthetizing the animals on ice for about 20 min and were fixed for 6–24 hr at room temperature by immersion in 4% paraformaldehyde in 0.1 M Sörensen phosphate buffer (SPB) containing 15% sucrose. Prior to incubation in fixative, LF were cut into 8-annuli long pieces (covering the LFD and LFP) as described previously (Schmidt et al. 2006), dactyls were cut into 2 or 3 pieces, and second antennae were cut into 5-8 annuli long pieces with coarse scissors. After fixation, LF, dactyls, and second antennae were decalcified by incubation in 10% EDTA (in SPB) for about one week (with several changes of the medium). All fixed tissues were stored in 0.02 M SPB with 0.02% sodium azide at 4°C until sectioning. For sectioning, tissues were embedded in gelatin (100 bloom for brain, 300 bloom for LF, dactyls, and second antennae) and cut on a vibrating microtome (VT 1000 S; Leica, Wetzlar, Germany) into 80–100 μm thick sections as described in detail previously for brain (Schmidt 2001) and LF (Schmidt et al. 2006). LF, dactyls, and second antennae were cut sagittally and brains were cut horizontally or sagitally (after splitting them into hemibrains).

Free-floating sections were incubated overnight at room temperature with an affinity-purified polyclonal rabbit antiserum against IR25a of the American lobster, *Homarus americanus* (anti-HaIR25a - courtesy of Dr. Tim McClintock, University of

Kentucky) diluted 1:750 in SPB containing 0.3% Triton-X-100 (TSPB). Anti-HaIR25a (previously annotated as anti-GluR1) was generated using two non-overlapping peptides (P1: TGEGFDIAPVANPW; P2: REYPTNDVDKTNFN) from the C-terminus of *H. americanus* IR25a (originally annotated as OET-07; Genbank accession #AY098942) (Hollins et al. 2003, Stepanyan et al. 2004). Sequence alignments of P1 and P2 with the deduced amino acid sequences of all IRs and iGluRs of *P. argus* identified in our transcriptome sequencing project showed matches for both peptides only in *P. argus* IR25a (P1: 79% identity; P2: 86% identity). This high degree of antigen identity between *H. americanus* and *P. argus* strongly suggests that anti-HaIR25a specifically labels IR25a in *P. argus*, as has been demonstrated by Western blots for *H. americanus* (Stepanyan et al. 2004).

For sections from LF, dactyls, second antennae, and some brains, anti-HaIR25a was combined with a mixture of two mouse monoclonal antibodies against modified α-tubulin isoforms that are enriched in neurons (Fukushima et al. 2009) to achieve labeling of all sensory neurons. These tubulin antibodies were anti-tyrosine tubulin (T9028, clone TUB-1A2, Sigma-Aldrich, St. Louis, Missouri) diluted 1:2000 and anti-acetylated tubulin (sc-23950, Santa Cruz Biotechnology, Dallas, Texas) diluted 1:200. After incubation in primary antibodies, sections were rinsed 4 x 30 min in TSPB and then incubated in a mixture of two secondary antibodies, goat anti-rabbit CY3 (111-165-003, Jackson Immunoresearch, West Grove, Pennsylvania) diluted 1:400 and goat anti-mouse DyLight-488 (35502, Thermo Fisher Scientific, Waltham, Massachusetts) diluted 1:100 in TSPB. After rinsing 3 x 30 min in TSPB, sections were incubated for 20 min in Hoechst 33258 diluted 1:150 in TSPB from a stock solution of 1 mg/ml to stain nuclei. After a final rinse in SPB, sections were mounted on slides in 1:1 glycerol:SPB containing 5% DABCO

(diazabicyclol[2.2.2]octane) to prevent photobleaching. Coverslips were secured with nail polish, and slides were stored at 4° C or at -20° C (for extended storage time). In some brains, labeling with anti-HaIR25a was combined with labeling by the lectin wheat germ agglutinin (WGA) that we previously identified as selective neuronal marker for the brain of *P. argus* (Schmidt and Derby 2011). In these brain sections, WGA-AlexaFluor-488 (Thermo Fisher Scientific - Invitrogen) diluted 1:1000 replaced goat anti-mouse DyLight-488 in the secondary antibody incubation medium.

Labeled sections were viewed and imaged at low magnification in an epifluorescence microscope equipped with color CCD camera (AxioScope FL LED with Axiocam 503, Carl Zeiss Microscopy, Thornwood, New York) and imaged at higher magnification in a confocal microscope (LSM 700, Carl Zeiss Microscopy) using the associated software package ZEN. Stacks of 0.3- to 1.0-µm thick optical sections covering the entire section thickness of 60–80 µm were collected. A different software package (LSM Image Browser Version 4.2.0.121, Carl Zeiss MicroImaging GmbH, Jena, Germany) was used to select sub-stacks of optical sections and collapse them to two-dimensional images using maximum-intensity projection.

2.2.10 Histological staining with ethyl gallate and methylene blue

Brains with attached antennular nerve roots were perfusion-fixed with 5% glutaraldehyde in 0.1 M SPB + 15% sucrose for 4 hr and post-fixed for 2-3 hr with OsO_4 according to a TEM-fixation protocol (Gnatzy et al. 1984). After dehydration in an ascending ethanol series and incubation in propylene oxide (2 x 30 min), brains were either embedded in hard Epon for semi-thin (1–2 µm) sectioning and subsequent staining

with methylene blue (1% in aqueous borax) or were incubated in ethyl gallate solution, embedded in Spurr's resin, and cut into 10–20 μm thick sections (according to methods described in detail in Schmidt et al. (1992) (Schmidt et al. 1992)). Sections were viewed and imaged in a bright field microscope equipped with color CCD camera (AxioScope FL LED with Axiocam 503, Carl Zeiss Microscopy, Thornwood, New York).

2.2.11 Scanning electron microscopy

Walking leg dactyls were cut off, cleaned by sonication for about 10 min (VWR Model 50T, VWR International, West Chester, Pennsylvania), cut into 2 or 3 pieces, and were fixed by immersion in 5% glutaraldehyde in 0.1 M SPB + 15% sucrose for 4 hr and post-fixed for 2-3 hr with OsO_4 according to a TEM-fixation protocol (Gnatzy et al. 1984). Fixed pieces were dehydrated, air-dried from hexamethyldisilazane, sputtered with palladium and viewed in a scanning electron microscope equipped with digital image capturing (Stereoscan 420 with LEO-32, Leica , Wetzlar, Germany) as described in detail by Schmidt and Derby (2005).

2.2.12 Processing of digital images

The digital images were processed by filtering out high frequency noise and by adjusting brightness and contrast with a free image and photo editing software (Paint.net 4.0.16, dotPDN LLC) before they were arranged into the final figures with Adobe Illustrator CS6.

2.3 Results

2.3.1 Identification of iGluRs and IRs in P. argus

To identify iGluRs and IRs in *P. argus*, we probed the predicted protein sequences (Parg protein database) from the transcriptome generated from three tissues – distal lateral flagellum of first antennule (LF), walking leg dactyls, and brain. We conducted BLAST searches with iGluR and IR sequences from *Drosophila melanogaster* (Dmel) and *Daphnia pulex* (Dpul) as query files and also performed an InterProScan screen of the transcriptomes for conserved the Pfam domains, PF00060 (consisting of M1, P, M2, S2, and M3 regions of iGluRs and IRs) (Supplemental S3 Text: https://doi.org/10.1371/journal.pone.0203935.s006) and PF10613 (consisting of S1 region of iGluRs and IRs) (Supplemental S4 Text: https://doi.org/10.1371/journal.pone.0203935.s007). The results of the InterProScan for the number of "Trinity genes" representing unique Trinity ID numbers and therefore protein sequences for iGluRs and IRs (Grabherr et al. 2011) yielded 342 sequences with PF00060 domain, 286 sequences with the PF10613 domain, and 132 with both of these domains.

To refine the identification of Parg homologues of iGluRs and IRs in the Parg transcriptome, we performed a phylogenetic analysis of all selected sequences (Supplemental S5 Text: https://doi.org/10.1371/journal.pone.0203935.s008) together with iGluRs and IRs from Dmel and Dpul. In this analysis, we used only the protein sequences having two domains of iGluRs and IRs (i.e., both PF00060 and PF10613 domains). Furthermore, sequences from this set that were short and had large gaps in these two domains, and sequences that did not have any transmembrane regions as predicted by TMHMM 2.0, were not included in the following analysis, with a few

exceptions that included sequence hits to conserved IRs from the BLAST searches. Thus, this set of 114 selected sequences from the transcriptome represents a conservative estimate of IRs and iGluRs. With over 340 unique sequences in the Parg transcriptome having at least the PF00060 domain, we expect the total number of IRs to be greater in Parg than our current estimate. Based on our phylogenetic analyses, the selected sequences were categorized into four groups: iGluRs, co-receptor IRs, conserved IRs, and divergent IRs, and the results are shown in Figure 2.2, Table 2.1, and Supplemental Figure 2.11 and Supplemental S2 Table: https://doi.org/10.1371/journal.pone.0203935.s002.

2.3.1.1 iGluRs

Six iGluRs were identified in the Parg transcriptome. They phylogenetically cluster with iGluRs of Dmel and Dpul, and as expected, are distinguished from IRs by lacking the highly divergent LBD of the IRs (Benton et al. 2009). Parg iGluRs include homologues to the Dmel NMDA receptors NMDAr1 and NMDAr2, non-NMDA (kainate) receptors GluRII and DmelCG11155, and AMPA receptors DmelGlu-R1 and DmelGlu-R1b. One Parg sequence is homologous to Dmel NMDAr1 sequence. Parg also has two sequences that are homologous to the Dmel NMDAr2 sequence, PargNMDAr2a and PargNMDAr2b. Among the non-NMDA iGluRs, there are two homologues to kainate receptors, Parg86731 and Parg73816. There is one homologue to Dmel AMPA receptors, Parg83058.

2.3.1.2 IRs

Our analysis demonstrated 108 sequences to be IRs in the Parg transcriptome. These included four co-receptor IRs, nine conserved IRs, and 95 divergent IRs (Table 2.1).

Co-receptor IRs. Four co-receptor IRs were identified in the Parg transcriptome: IR8a, IR25a, IR76b, and IR93a. These sequences and their tissue specific expression are shown in Figure 2.2 and Supplemental Figure 2.11. IR25a and IR8a cluster close to the iGluRs because only they have a defined ATD like the iGluRs. We did not identify IR8a in the Dpul database, as others have also reported (Croset et al. 2010, Eyun et al. 2017).

Divergent IRs. The Parg transcriptome had 95 divergent IRs. Out of these 95 divergent IRs, four are putative IR75-like sequences (PargIR1091, PargIR1092, PargIR1093, and PargIR1094. See below in **Conserved IRs**) and one is a putative IR68a homologue (PargIR68a-put. See below in **Conserved IRs**). The distribution of the remaining divergent IRs between the tissues was highly skewed and mostly non-overlapping (Supplemental Figure 2.11 and Supplemental S2 Table: https://doi.org/10.1371/journal.pone.0203935.s002): 51 divergent IRs were only in LF, two divergent IRs only in dactyl, and over 14 divergent IRs were found in both LF and dactyl. Two additional divergent IRs that are different from our 95 sequences were identified in the LF by Corey et al. (2013), who named them IR4 and IR7.

Multiple sequence alignments of iGluRs and IRs of Parg and Dmel illustrate two major similarities and one dissimilarity between these two species (Figure 2.3; Supplemental S6 Text: https://doi.org/10.1371/journal.pone.0203935.s009). The first similarity is a high variability of amino acids in S1 and S2 regions of the IRs of Parg and Dmel. This is probably related to the specificity of ligand binding of these receptors. The second similarity is that the S2 region of most IRs of both Parg and Dmel lack both the amino acid residues present in iGluRs that bind the glutamate ligand. The dissimilarity in the IRs of Parg and Dmel is that only 31% of the IRs in Dmel have the glutamate binding

arginine residue (R) in S1 that is conserved in iGluRs (Mayer et al. 2006, Benton et al. 2009), but 99% (107 of 108) of the Parg IRs have this residue. One Parg divergent IR, PargIR1059, lacks this arginine residue (Figure 2.2, sequence name in black). Instead of the conserved 'R' residue of S1, PargIR1059 has a non-polar tryptophan (W) residue. PargIR93a does not have a substitution is this conserved site unlike Dmel IR93a (Figure 2.3; (Benton et al. 2009)).

Conserved IRs. Based on BLAST search results and phylogenetic analyses (Figure 2.2 and Figure 2.4), we identified nine homologues to conserved IRs in the Parg transcriptome. Out of these nine conserved IRs, one conserved IR is PargIR21a, while the other eight are an expanded family of IR40a homologues. These sequences are denoted as Parg IR40a-1 to Parg IR40a-8 (Figure 2.4).

To represent the Parg co-receptor IRs and conserved IRs in a broader phylogenetic context, we constructed another tree using only conserved IRs and with more species included. These include Parg, Dmel, and Dpul as in Figure 2.2, another decapod crustacean (Ccly), eight additional insect species (Aaeg, Cqui, Agam, Bmor, Tcas, Amel, Nvit, Apis, and Phum), and two gastropods (Acal and Lgig). The results from this analysis (Figure 2.4) highlight three major points. First, this analysis confirms the results of Figure 2.2 in showing four co-receptor IRs (IR8a, IR25a, IR76b, IR93a) and conserved IRs (IR21a and the expanded IR40a family) in the Parg transcriptomes in a broader phylogenetic context. Second, it also suggests that Parg may have additional conserved IRs, including members of the IR75 clade that are present in the transcriptome. There are four sequences that we consider putative conserved IRs due to their proximity to the insect IR75 clade of conserved IRs. The four Parg sequences (PargIR1091, PargIR1092, PargIR1093, PargIR1094) may be homologous to the sequences from other species in the

IR75 clade (Figure 2.2 and Figure 2.4). In this study, we classified these putative IR75 sequences as divergent IRs and the numbers are appropriately reflected in Table 2.1. Third, one Dpul IR, DpulIR304, which was previously classified as divergent IR by Croset et al. (2010), clustered with the homologues of the co-receptor IR76b. Another Dpul divergent IR, DpulIR298, may also be a conserved IR. Based on BLAST searches against non-redundant NCBI databases and our phylogenetic analyses, we classify it as homologue of IR68a. A Ccly hermit crab IR (CEF34375.1) clustered with insect IR68a (Groh-Lunow et al. 2014), and we identified a putative IR68a in Parg, PargIR68a-put. We consider this sequence as a putative homologue due to the sequence being incomplete and reciprocal BLAST searches against non-redundant NCBI databases returning hits for other conserved insect IRs. We also did not find other insect conserved IRs in Parg or Dpul.

We point out that these analyses did not include over 200 sequences of putative Parg IRs due to their small sequence length and missing domain regions (Supplemental S3 and S4 Text: https://doi.org/10.1371/journal.pone.0203935), and therefore our estimates of the type and number of IRs are conservative.

2.3.2 Immunolocalization of the P. argus *IR25a receptor*

Given the importance of the co-receptor IR25a in forming functional receptors in sensory neurons in other arthropods, we sought to localize the expression of this protein in LF, dactyl, second antennae, and brain of *P. argus*. In all these tissues, anti-HaIR25a intensely and selectively labeled particular types of cells (Figure 2.5; Figure 2.6; Figure 2.7). Negative controls, by labeling tissue sections using the same ICC protocol but

omitting anti-HaIR25a, showed no labeling of these cell types or any other cells. Use of anti-tubulin to label other tissue components and the nuclear marker Hoechst 33258 to reveal nuclei of all cells allowed a robust interpretation of the labeling pattern obtained by anti-HaIR25a. In LF, dactyl, and the flagellum of the second antennae (all of which contain no musculature), anti-tubulin primarily labeled bipolar sensory neurons, epithelial cells, and the walls of hemolymph vessels. Analysis of the overall distribution of cell nuclei strongly suggests that anti-tubulin indeed labeled all bipolar sensory neurons. In the brain, anti-tubulin primarily labeled neuronal elements, especially axonal tracts and neuropils.

2.3.2.1 Lateral flagella of antennules

In the aesthetasc-bearing tuft region of LF, anti-HaIR25a intensely labeled the clusters of ORN somata associated with the aesthetascs (Figure 2.5e-q) as shown previously (Tadesse et al. 2011). Double-labeling with anti-tubulin revealed that most or all ORN somata of a cluster are HaIR25a-positive (Figure 2.5l-o). Glia-like auxiliary cells of aesthetascs whose somata are located at the apical pole of ORN clusters surrounding the emerging inner dendritic segments (Figure 2.5d) were not labeled by anti-HaIR25a or anti-tubulin (Figure 2.5l-o). In the ORN somata, HaIR25a-like immunoreactivity was not restricted to the cell membrane, but rather it extended throughout the entire cytosol. Labeling intensity within the cytosol was not uniform but more intense in spherical inclusions that may represent stacks of smooth endoplasmic reticulum (as described in ORN somata of another spiny lobster species (Spencer and Linberg 1986)). The ORN somata in a cluster varied very little in size (n=82, Feret diameter: 9.2-12.9 µm

(minimum-maximum); mean ± SD: 11.0 ± 0.9 μm) and in the intensity of HaIR25a-like immunoreactivity. From the ORN somata, HaIR25a-like immunoreactivity and also tubulin-like immunoreactivity extended into the inner dendritic segments, which arise from the apical pole of the ORN somata, and into the axons, which arise from the basal pole of the ORN somata. For anti-HaIR25a, the labeling intensity was higher in the ORN somata than in the axons and inner dendritic segments (Figure 2.5m), whereas for anti-tubulin, the labeling intensity was higher in the axons and inner dendritic segments than in the somata (Figure 2.5n). Proximal to the ORN clusters, the axon fascicles arising from them fuse with the common lateral flagellum nerve which was intensely labeled by anti-tubulin and less intensely labeled by anti-HaIR25a (Figure 2.5p, q). The inner dendritic segments arising from the apical pole of the ORN somata of a cluster form a common fascicle, which traverses the cuticle in a wide canal leading to the base of the associated aesthetasc seta (Figure 2.5c, d). Within the aesthetasc seta, at about 25% of the total setal length, the inner dendritic segments give rise to the outer dendritic segments representing modified cilia. The outer dendritic segments are highly branched and together the thin branches of all ORN outer dendritic segments fill the entire lumen of the aesthetasc seta distal to the transition area (Figure 2.5d). Cross sections through aesthetasc setae revealed that ORN outer dendritic segments were strongly labeled by anti-HaIR25a (Figure 2.5g-j). In fact, labeling intensity was higher in the outer dendritic segments than in the inner dendritic segments or the ORN somata. Outer dendritic segments were also strongly labeled by anti-tubulin (Figure 2.5g, i, k), but in this case labeling intensity was lower in the outer dendritic segments compared to the inner dendritic segments.

In the tuft region of LF, aesthetascs are accompanied by three types of smooth sensilla likely innervated by distributed CRNs and MRNs (Cate and Derby 2001, Cate and Derby 2002, Schmidt and Mellon 2011). These presumptive bimodal sensilla are guard setae, asymmetric setae, and companion setae (Figure 2.1d and Figure 2.5a, c). They are located lateral to the two rows of aesthetascs traversing each annulus. Sections through the lateral region of the LF revealed the presence of clusters of bipolar sensory neurons that were distinctly labeled by anti-tubulin and anti-HaIR25a but clearly differed from clusters of ORNs in several parameters (Figure 2.5p-s). First, the number of sensory neurons in a cluster was substantially lower (between 10 and 15) than in a typical cluster of ORNs (up to 320 (Grünert and Ache 1988)). Second, the overall intensity of HaIR25a-like immunoreactivity in these clusters was noticeably lower than in neighboring clusters of ORNs. Third, the somata of the sensory neurons composing these clusters were considerably larger on average than the somata of ORNs and they varied more in size (n = 45, Feret diameter: 8.8–30.1 μm (minimum-maximum); mean ± SD: 18.3 ± 4.8 μm. Fourth, not all sensory neurons delineated by distinct tubulin-like immunoreactivity were double labeled by anti-HaIR25a. Typically, two or three sensory neurons in the apical region of the cluster and occasionally single sensory neurons located further basally were completely devoid of HaIR25a-like immunoreactivity. Most likely each of these clusters of sensory neurons innervates one of the bimodal sensilla accompanying the aesthetascs, but we were unable to unequivocally link particular clusters to the setae they innervate. This was mainly because the inner dendritic segments arising apically from the somata were only weakly labeled by either antibody and therefore could not be traced to the base of setae. The axons arising basally from the somata were more intensely labeled by both antibodies (Figure 2.5p, q). We failed to unequivocally detect anti-HaIR25a-like or

tubulin-like immunoreactivity in the outer dendritic segments running through a narrow canal within the thick-walled setae. This was because the cuticle forming these setae is intensely autofluorescent in the channels used to visualize anti-HaIR25a-like or tubulin-like immunoreactivity.

Proximal to the aesthetasc-bearing tuft region, the LF of the antennule bears fewer and less regularly arranged presumptive bimodal sensilla of different types (Cate and Derby 2001, Cate and Derby 2002). Sections through the proximal region of the LF revealed the presence of widely dispersed clusters of bipolar sensory neurons that were distinctly labeled by anti-tubulin and anti-HaIR25a. These clusters of sensory neurons were similar to the ones presumably innervating bimodal non-aesthetasc sensilla in the tuft region of the lateral flagellum in number (between 8 and 15) and size (n = 8, Feret diameter: 10.4-20.4 μm (minimum-maximum); mean ± SD: 14.5 ± 3.4 μm) of sensory neurons and in the presence of one or two sensory neurons in the apical region of the cluster that were completely devoid of HaIR25a-like immunoreactivity. The inner dendritic segments arising apically from the somata were only weakly labeled by either antibody, and we were therefore unable to link the clusters to particular setae on the surface of the cuticle.

2.3.2.2 Walking leg dactyl

The dactyl of walking legs of *P. argus* is organized into a smooth epicuticular cap at the tip and a much longer proximal region that bears numerous dense bundles of robust smooth setae that are organized in six longitudinal rows (Figure 2.6a-c). In addition, single smooth and very robust spines are interspersed between the rows of bundles of setae. In having a thin central canal that reaches all the way to their tip, the smooth setae

and smooth spines show one morphological characteristic defining bimodal chemo- and mechanosensory sensilla (Schmidt and Mellon 2011). The epicuticular tip bears numerous small depressions at the end of thin canals traversing the cuticle (Figure 2.6c). This organization is typical of bimodal sensilla called funnel-canal organs that lack an outer seta, and whose ultrastructure and chemosensory physiology was described in shore crabs, *Carcinus maenas* (Gnatzy et al. 1984, Schmidt and Gnatzy 1984, Schmidt and Gnatzy 1989). None of the presumptive sensilla on the dactyls of *P. argus* have been studied with electron microscopy, and chemoreception mediated by sensilla on the dactyls of *P. argus* has been demonstrated in only two studies (Gleeson et al. 1989, Kem and Soti 2001).

Sections through the distal region of the dactyl (excluding the epicuticular cap, which could not be sectioned well with the vibrating microtome even after decalcification) revealed the presence of numerous large clusters of bipolar sensory neurons that were distinctly labeled by anti-tubulin and anti-HaIR25a (Figure 2.6d-f). These clusters of sensory neurons were similar to the clusters of sensory neurons that presumably innervate bimodal chemo- and mechanosensory sensilla in the LF (see above) in having a relatively low number of sensory neurons (between 15 and 20) that were larger than the somata of ORNs and varied substantially in size (n = 77, Feret diameter: 9.1-22.9 μm (minimum-maximum); mean ± SD: 13.9 ± 2.9 μm) (Figure 2.6g-l). Most strikingly, also the clusters of sensory neurons in the dactyl typically had two or three sensory neurons (identified by their intense tubulin-like immunoreactivity) completely devoid of HaIR25a-like immunoreactivity in their apical region. In addition, the clusters contained one or two sensory neurons sub-apically that were only lightly labeled by anti-HaIR25a in contrast to the remaining neurons in the basal part of the clusters that expressed

intense anti-HaIR25a-like immunoreactivity. Axon fascicles emerging basally from the clusters of sensory neurons and fascicles of inner dendritic segments emerging apically were intensely labeled by both antibodies. Fascicles of inner dendritic segments projected to the bases of bundled smooth setae but could not be traced further into the setae because of the intense autofluorescence of the setal cuticle. Interspersed between clusters of sensory neurons were single, spindle-shaped bipolar sensory neurons that expressed intense tubulin-like but no anti-HaIR25a-like immunoreactivity (Figure 2.6d, f).

2.3.2.3 Second antennae

The second antennae of *P. argus* are organized into three basal segments and an extraordinarily long and sturdy flagellum that is organized into numerous annuli. The anatomical organization of the annuli including their setation has not been studied in *P. argus* or any other spiny lobster species. Superficial examination of antennal flagella with light microscopy revealed that each annulus carries a stereotypical arrangement of setae concentrated at the distal edge of the annulus. Most of these setae are smooth and have a thin central canal that reaches all the way to their tip and hence show morphological characteristics of bimodal chemo- and mechanosensory sensilla (Schmidt and Derby 2011). Sections through the antennal flagellum revealed the presence of some clusters of bipolar sensory neurons that were distinctly labeled by anti-tubulin and anti-HaIR25a (Figure 2.6m-o). These clusters of sensory neurons were located directly under the thick cuticle, and they were similar in number (between 6 and 12) and size of sensory neurons to the clusters of sensory neurons innervating bimodal non-aesthetasc sensilla in the

distal and proximal lateral flagellum. One or two sensory neurons in the apical region of each cluster were completely devoid of HaIR25a-like immunoreactivity.

2.3.2.4 Brain

The neuroanatomy of the brain of *P. argus* is known in substantial detail especially with regard to the deutocerebrum that receives sensory input from sensilla on the antennules (Schmidt and Ache 1992, Schmidt et al. 1992, Schmidt and Ache 1996, Schmidt and Ache 1996, Schachtner et al. 2005). Axons of sensory neurons innervating sensilla on the antennules form the common antennular nerve that is subdivided into three main divisions. The lateral division contains axon fascicles originating from the LF including numerous extremely thin axons of ORNs, the medial division contains axon fascicles originating from the medial flagellum, and the dorsal division contains axon fascicles originating from the basal segments of the antennule (Schmidt et al. 1992). Upon entering the brain, axons in the lateral division of the antennular nerve undergo a massive rearrangement in which the extremely thin ORNs axons get sorted out from all other axons in the lateral division. Proximal to this axon sorting zone, the ORN axons form a large fascicle projecting towards the olfactory lobe whereas all other axons form a fascicle projecting to the lateral lobe of the lateral antennular neuropil (Schmidt and Ache 1992, Schmidt et al. 1992). The lateral division of the antennular nerve at its entry into the brain and the axon sorting zone contain conspicuous, very large cells (diameter 30–50 μm) that are intensely stained by ethyl gallate and methylene blue (Figure 2.7b, d, e; (Schmidt and Ache 1992)) and were originally described by Herbst (1916) in the antennular nerve of the spiny lobster, *Palinurus vulgaris*.

In the brain, anti-HaIR25a intensely and selectively labeled these large cells in the proximal part of the lateral division of the antennular nerve and the adjacent axon sorting zone (Figure 2.7f-h). No other cellular structure in the brain including the ORN afferent axons that were intensely labeled by anti-tubulin (Figure 2.7f, i, k, m) expressed any anti-HaIR25a-like immunoreactivity. Labeling with anti-HaIR25a revealed that about 100 of the large cells are present in each hemibrain and that most of these cells (≥ 95%) are unipolar, giving off one process that typically projects towards the olfactory lobe (Figure 2.7i, j). A smaller fraction of large cells labeled by anti-HaIR25a (about 5%) were distinctly bipolar with a spindle-shaped soma and two processes arising from the opposing cell poles (Figure 2.7k, l). Very few large cells were pseudo-unipolar with two processes arising from one pole of the cell and projecting in opposite directions. Because the intensity of anti-HaIR25a declined rapidly in the processes, they could not be traced to their terminals. To elucidate if the large cells have neuronal identity, we double-labeled brains with anti-HaIR25a and WGA, which labels the vast majority of neurons residing in the CNS of *P. argus* (Schmidt and Derby 2011). In the large cells, WGA did not label the cytosol or the cell membrane, only the nuclear envelope was weakly WGA-positive. In contrast, neuronal somata in the same sections imaged using the identical setting of the confocal microscope showed intense labeling of the cytosol and very intense labeling of the cell membrane with WGA.

2.3.3 PCR

PCR was used to verify and extend the transcriptomics results by examining expression of IRs in an array of tissues that included those used to generate the Parg transcriptome but also extended to other tissues. We tested two regions of the LF, as

described in the previous section: the distal region of the lateral flagellum of the antennule (LFD), which bears the olfactory (aesthetasc) sensilla; and the proximal, non-aesthetasc bearing region of the lateral flagellum of the antennule (LFP). Besides the dactyl of the pereiopods (Da) and the brain (Br), from which we developed transcriptomes, we also tested two other tissues: the second antenna (A2) and green gland (GG), an excretory organ. We examined three co-receptor IRs (IR25a, IR8a, IR93a) and two divergent IRs, the divergent PargIR1074 (expressed only in dactyl; Supplemental Figure 2.11 and Supplemental S2 Table: https://doi.org/10.1371/journal.pone.0203935.s002) and the divergent PargIR1028, found in the Parg transcriptome. We also tested the iGluR, NMDAr1, and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as a positive control. Figure 2.8 shows representative results from our PCR runs, and Table 2.2 summarizes all findings. Our results show complete agreement between our transcriptomics and PCR results: genes expressed in a transcriptome were also found there through PCR. Thus, IR25a, IR8a, IR93a, and NMDAr1, identified in the Parg transcriptome, were also identified in all three tissues via PCR (Figure 2.8). IR1074, expressed only in dactyl and not LF or brain (Supplemental Figure 2.11 and Supplemental S2 Table: https://doi.org/10.1371/journal.pone.0203935.s002), were also only found in dactyl via PCR. IR1028, expressed only in LF (Supplemental Figure 2.11 and Supplemental S2 Table: https://doi.org/10.1371/journal.pone.0203935.s002), was also only found in LF via PCR (Figure 2.8). IRs expressed in the LF were expressed in both the LFD and LFP, although expression levels typically appeared to be qualitatively much higher in the LFD than LFP. PCR results showed that IR25a, IR8a, and IR93a are also found in the green gland, which is an excretory organ (Figure 2.8). We sequenced PCR products and verified that they have > 90% similarity to the expected sequence (Table 2.2).

2.3.4 Identification of a GR

We searched for insect-like ORs and GRs in the Parg transcriptomes. We did not find any ORs, but we did find one GR, PargGR1 (Figure 2.9, Supplemental S11 Text: https://doi.org/10.1371/journal.pone.0203935.s014 and Supplemental S12 Text: https://doi.org/10.1371/journal.pone.0203935.s015), according to the following lines of evidence as suggested by Robertson (2015). First, BLAST searches of our Parg transcriptomes using arthropod GRs as query files identified PargGR1 as a GR (see **Methods** for e-values). Second, the top hits from BLAST searches of non-redundant NCBI databases with PargGR1 as the query were arthropod GRs. Third, an InterProScan screen of our Parg database for the GR specific 7tm_7 domain, PF08395, identified PargGR1 (Fox et al. 2001, Robertson 2015). The screen using InterProScan did not produce other candidates with OR specific 7tm_6 domain, PF02949, or the 7tm_7 domain in the Parg transcriptome. Although the PargGR1 sequence is a fragment (134 amino acids), Figure 2.9 shows the conserved region of the 7tm_7 domain (Robertson 2015) for PargGR1 along with several GRs from Dmel, Dpul, and the copepod *Eurytemora affinis* (Eaff) acquired from Eyun et al. (2017), demonstrating the presence of 7tm_7 in the PargGR1 fragment. Consistent with GRs and GRLs (GR-Like) from other species, PargGR1 also has the highly conserved motif 'TYxxxxxQF' found in the TM7 region of all GRs (Robertson 2015, Saina et al. 2015). Using TMHMM, we were only able to predict one TM region in PargGR1 due to the small size of the fragment.

2.3.5 Identification of TRP channels

We searched our Parg protein database for Transient Receptor Potential (TRP) channels by using BLASTp and TRP channel sequences as queries from insects (Dmel,

Bmor, Tcas, Amel, Nvit, and Phum), nematodes (Cele), and mammals (Mmus and Rnor). Using the BLAST results and an InterProScan screen, we selected Parg sequences with TRP channel domain regions (see **Methods** for list of Pfam domains) from the transcriptome, and then used these sequences along with the TRP channel query sequences and Dpul TRP channels (Peng et al. 2015) to construct a phylogenetic tree.

Homologues to all seven TRP subfamilies (Venkatachalam and Montell 2007, Venkatachalam et al. 2014) were found in the Parg transcriptome (Figure 2.10; Supplemental S13-S15 Text: https://doi.org/10.1371/journal.pone.0203935). An expanded TRPA subfamily of the Group 1 TRP channels of insects was identified that was similar to insects (Matsuura et al. 2009). Overall, our analysis determined nine types of TRPA channels in the Parg transcriptome. One homologue to TRPA1 was identified and found to cluster with Cele TRPA1. Two more homologues to TRPA1, PargTRPA1-like1 and PargTRPA1-like2, were also detected. Two homologues to painless, an insect TRPA channel, were identified. Two homologues to hymenopteran TRPA5 channels, PargTRPA5-1 and PargTRPA5-2, were also detected in the transcriptome (Supplemental S15 Text: https://doi.org/10.1371/journal.pone.0203935.s018). Another TRPA subfamily member, PargTRPApw, clustered with the insect TRPA channel clades pyrexia, waterwitch, and painless. The remaining member of the TRPA subfamily, PargTRPA-like1, clustered with the TRPA subfamily clade but did not directly cluster with any of the TRPA channel types in insects or Dpul.

Two homologues to the TRPM subfamily were identified. The first, Parg TRPM, is a homologue to the insect TRPM subfamily (Matsuura et al. 2009) (Figure 2.10; Supplemental S15 Text: https://doi.org/10.1371/journal.pone.0203935.s018). The second, Parg TRPMm, is a homologue to the mammalian TRPM channel subfamily and

clustered with the Cele TRP channel, CED-11. PargTRPMm had higher abundance estimated in the LF (Supplemental Figure 2.12 and Supplemental S3 Table: https://doi.org/10.1371/journal.pone.0203935.s003).

Two homologues to the TRPV subfamily were found. A homologue to the arthropod TRPV channel 'Nanchung,' PargNan, and the homologue, PargIav, to the TRPV channel 'inactive' in arthropods and OSM-9 in Cele were found in the Parg transcriptome. Two homologues to the classical TRP channels, TRPC subfamily, were identified. The first, PargTRPgamma, is homologous to arthropod TRPgamma channels. The second, PargTRPC (Figure 2.10; Supplemental S15 Text: https://doi.org/10.1371/journal.pone.0203935.s018) clusters with the Cele TRP-1a channel. A single homologue, PargNompC, to the insect TRPN channel, NompC, was identified.

Homologues to the group 2 TRP subfamilies, TRPP (Pkd2) and TRPML, were also found in Parg (Figure 2.10). A single homologue, Parg TRPML, to insect TRPML sequences and the homologue to insect Pkd2 channel, PargPkd2, were found.

2.4 Discussion

The overarching aim of this study is to gain a molecular understanding of the first step in chemoreception – the receptor proteins on olfactory and distributed chemoreceptor neurons – in the Caribbean spiny lobster, *Panulirus argus*, from a functional and phylogenetic perspective. We chose *P. argus* because it is a major model decapod crustacean for chemoreception and thus represents a good starting point for comparison with other decapod crustaceans and other arthropods, in particular insects.

Our principal objective was to identify and compare a major set of chemoreceptor proteins of insects – the ionotropic receptors (IRs) – in the two major chemosensory organs – the antennular lateral flagellum representing olfaction and distributed chemoreception, and the leg dactyls representing only distributed chemoreception – and brain of *P. argus*. A secondary objective was to search for other types of chemoreceptor proteins besides IRs, including Olfactory Receptors (ORs), Gustatory Receptors (GRs), and Transient Receptor Potential (TRP) channels. We conservatively found 108 IRs that include co-receptor IRs, conserved IRs, and divergent IRs. However, most (95) are divergent IRs and the vast majority (51) of these divergent IRs are found only in the antennular lateral flagellum and thus likely involved in olfaction. We show that one co-receptor, IR25a, is expressed in olfactory receptor neurons and other chemosensory neurons. We also identified one GR and homologues to all subfamilies of TRP channels. Some of these molecules are expressed not only in chemosensory tissue but also in the brain, suggesting that they may mediate diverse functions.

2.4.1 Diversity and distribution of IRs in two chemosensory organs of P. argus

We identified 108 IRs in the antennules, leg dactyls, and brain of *P. argus*: four co-receptor IRs, nine conserved IRs, and 95 divergent IRs. This is a conservative estimate of the diversity of IRs, because we only included sequences with both the ligand binding domain and the ion channel domain that did not have large gaps in these domain regions.

The four IR co-receptors of *P. argus* – IR8a, IR25a, IR76b, and IR93a – were found in both the LF and dactyl of *P. argus*. IR25a has been found in all protostomes examined

to date with few exceptions, suggesting it is the ancestral IR and has conserved functions (Croset et al. 2010, Eyun et al. 2017). IR8a, IR76b, and IR93a evolved more recently, and their expression appears to be limited to the arthropods. IR25a and IR8a have been found in most crustaceans examined to date, including *Homarus americanus* (Hollins et al. 2003), spiny lobsters (Tadesse et al. 2011, Corey et al. 2013), seven species of shrimp (Zbinden et al. 2017), and copepods (Lenz et al. 2014, Núñez-Acuña et al. 2014, Núñez-Acuña et al. 2016, Eyun et al. 2017). Some crustaceans have been found to have IR25a but not IR8a, including the branchiopod *Daphnia* and two species of hermit crabs (Figure 2.4) (Croset et al. 2010, Groh et al. 2013, Groh-Lunow et al. 2014). Eyun et al. (2017) propose that IR8a appeared early in the pancrustaceans and may have been secondarily lost in the branchiopods. IR76b has been found not only here in *P. argus* but also in copepods (Eyun et al. 2017) and *Daphnia*. The relatively infrequent reporting of IR76b in crustaceans might be because it has only recently been identified in *Daphnia* ((Eyun et al. 2017), Figure 2.4) and could not be identified in phylogenetic analyses without homologues from other crustaceans.

Orthologues of some IRs have been found across insects. These are called conserved IRs and include IR21a, IR31a, IR40a, IR41a, IR60a, IR64b, IR68a, the IR75 family, IR76a, IR84a, and IR92a (Figure 2.4) (Croset et al. 2010, Rytz et al. 2013). Some of these insect conserved IRs – IR21a and IR40a – have also been found in other arthropods, including some chelicerates, myriapods, and crustaceans (Eyun et al. 2017). We also identified several homologues of IR40a and IR75-like, one homologue of IR21a, and one putative homologue of IR68a in *P. argus*. Eyun et al. (2017) did not find IR40a in the 14 species of examined crustaceans.

The functions of IRs have been most explored in insects, especially *D. melanogaster*. In crustaceans, IR25a appears to be broadly expressed in all chemosensory neurons, but this is not the case in *D. melanogaster*. Although IR25a is broadly expressed in a subset of *D. melanogaster* antennal sensilla (coeloconic sensilla), it is not expressed in all ORNs in those sensilla. IR8a is also a co-receptor IR in *D. melanogaster* coeloconic sensilla. However, IR25a and IR8a are not expressed in the same neurons, with a few exceptions (Benton et al. 2009, Abuin et al. 2011). They form functional ionotropic receptors that respond to specific chemicals when co-expressed with specific divergent IRs (Abuin et al. 2011, Silbering et al. 2011, Rytz et al. 2013). These co-receptor IRs can play an integral role in the transport and insertion of IRs into the sensory ciliary membrane (Abuin et al. 2011). IR25a of *D. melanogaster* is a co-receptor not only in chemosensory neurons but also in neurons sensing humidity and cool temperatures (Enjin et al. 2016, Knecht et al. 2016, Ni et al. 2016). IR25a has also been implicated in mediating the circadian clock by itself through warmth-sensitivity without the co-expression of other IRs (Chen et al. 2015).

Another co-receptor IR found across arthropods, IR93a, is expressed in antennal neurons in *D. melanogaster*, and it functions in hygroreception together with IR25a and the conserved IRs IR40a and IR68a (Enjin et al. 2016, Knecht et al. 2016, Frank et al. 2017, Knecht et al. 2017). Another conserved IR, IR21a, plays a role in thermoreception when co-expressed with IR25a and IR93a (Knecht et al. 2016, Ni et al. 2016). We do not know the role of these conserved IRs in *P. argus* and other crustaceans. We detected a putative homologue to IR68a in *P. argus*, and found several homologues to IR40a and one to IR21a (Figure 2.2 and Figure 2.4), raising the possibility that in crustaceans IR25a and IR93a might play a role in reception of stimuli beyond chemicals. Previously, IR93a

was found only in antenna of insects, but we found this co-receptor IR not only in the LF but also in the dactyl and brain of *P. argus*. The expansion of IR40a homologues and the possible presence of IR68a in *P. argus* is particularly interesting since IR40a mediates 'dry' humidity sensing and IR68a mediates 'moist' humidity sensing in *D. melanogaster* when co-expressed with IR25a and IR93a (Enjin et al. 2016, Knecht et al. 2016, Frank et al. 2017, Knecht et al. 2017).

The co-receptor IR76b is broadly expressed in the antennal ORNs of *D. melanogaster* but is typically co-expressed with IR25a or IR8a. The co-expression of the conserved IR, IR41a, in IR76b expressing ORNs confers them with polyamine sensitivity (Hussain et al. 2016). IR25a and IR76b are not only expressed in the antenna of *D. melanogaster* but are also broadly expressed in the gustatory receptor neurons (GRNs) in sensilla on the labellum and tarsi (Croset et al. 2010, Zhang et al. 2013, Ganguly et al. 2017). Only IR76b appears to function as a co-receptor in these GRNs despite co-expression with IR25a. IR76b-expressing GRNs in the labellum and tarsi detect amino acids (Croset et al. 2016, Ganguly et al. 2017), salt (Zhang et al. 2013), and polyamines (Hussain et al. 2016). Furthermore, salt detection in tarsal GRNs is blocked when the divergent IR, IR20a, is co-expressed with IR76b (Croset et al. 2016, Ganguly et al. 2017). Given this, it is interesting to speculate that in CRNs of crustaceans, the divergent IRs co-expressed with IR76b might also confer a particular chemical sensitivity to that neuron.

In insects, IRs are preferentially sensitive to amines and acids, whereas ORs are more sensitive to esters and alcohols (Abuin et al. 2011, Silbering et al. 2011). This bias toward amines and acids for insect IRs is consistent with the evolution of IRs in ancestral protostomes from iGluRs, which are sensitive to L–glutamate. Abuin et al. (2011) proposed that IR25a, which is most similar in sequence to iGluRs and is likely the

ancestral IR, functioned as a glutamate detector in common protostome ancestors that lived in marine environments. With the expansion of IRs, IR25a and IR8a assumed co-receptor functions and the new divergent IRs allowed sensitivity to a broader array of environmental chemicals. What might be the ligands for the IRs of crustaceans? We know that for *P. argus* and other crustaceans, many olfactory and distributed chemoreceptors neurons are sensitive to amino acids, amines, and other small nitrogenous molecules (Atema and Voigt 1995, Schmidt and Mellon 2011, Derby and Weissburg 2014), which are of the same general classes though different specific molecules than detected by IRs of insects. Assuming that IRs are the major if not sole chemoreceptor proteins on ORNs and CRNs, then amino acids, amines, and other small nitrogenous molecules are among the ligands for IRs. Of course, identifying the specific ligands for each divergent IR will require future studies.

A total of 95 divergent IRs were identified in the LF and dactyls of *P. argus*. This is a conservative estimate, and the number of divergent IRs may be as high as 300. We found several unique sequences in the Parg transcriptome that had only one domain region (PF00060: transmembrane domains, pore, and S2 of LBD). All these sequences were not used for phylogenetic analyses and instead only a small subset that also had the S1 domain region were analyzed. Since we required that identified IRs have both domain regions (i.e., PF00060 and PF10613), our phylogenetic analyses did not include all sequences that had only one of the two domains regions. Interestingly, these 95 divergent IRs have different expression patterns in LF and dactyl: 51 are expressed only in LF, two are expressed only in the dactyl, 14 are expressed in both LF and dactyl, and the rest are expressed in LF and brain. One possible reason for this difference is that LF contains both olfactory and distributed chemoreceptor sensilla and neurons, while the dactyl

contains only distributed chemoreceptors. If the number of IRs associated with distributed chemoreception were the same in LF and dactyl, this would imply that most IRs in the LF are part of the olfactory pathway. We know that the olfactory pathway of *P. argus* mediates responses not only to food-related chemicals (as do dactyl chemoreceptors) but it also uniquely mediates responses to conspecific chemicals such as social cues, alarm cues, and most probably sex pheromones(Schmidt and Mellon 2011, Derby and Weissburg 2014, Derby et al. 2016). Despite the fact that both the LF and dactyl have sensilla that are innervated by distributed chemoreceptors, the overlap of IRs between these two organs is low. This finding suggests that all populations of distributed chemoreceptors on other appendages (Figure 2.1) express distinct sets of IRs, which makes the overall chemosensory system in *P. argus* unexpectedly complex. Alternative explanations are possible, of course, one of which is that distributed chemoreceptors express another class of currently unidentified chemoreceptor proteins not expressed in ORNs and that these unidentified proteins will significantly raise the total number of protein types expressed in CRNs to be equal to that of the IRs in ORNs.

How does the total number of IRs in *P. argus* compare to other species? By our conservative estimate, the number is about the same in insects: *Drosophila* spp. have 58-69 IRs, and other insect species are often in this range though the number in insects can be as few as 20 or as many as 150 (Rytz et al. 2013). Insects also have many ORs and GRs. Among crustaceans, *Daphnia* has 85 IRs and 58 GRs and thus seems to be highly equipped with chemoreceptor proteins despite its highly reduced olfactory system (both antennule and brain) relative to the complex lifestyle and use of chemical sensing in *P. argus*. The number of divergent IRs in hermit crabs has been reported to be lower: 16 in the hermit crab *Pagurus bernhardus* and 22 identified in *C. clypeatus* (Groh et al. 2013).

2.4.2 Cellular expression patterns and possible functions of crustacean IRs

2.4.2.1 IR25a in the periphery

In all sensilla-bearing appendages included in this study (lateral flagella of antennules, leg dactyls, second antennae), anti-HaIR25a exclusively labeled sensory neurons and no other tissue components. This high degree of selectivity is in line with the notion that the antibody indeed binds to authentic IR25a as was demonstrated for *P. argus* IR25a with Western blots (Corey et al. 2013) and for *H. americanus* IR25a through preabsorption controls with the antigen and Western blots (Stepanyan et al. 2004). Immunocytochemistry results for *P. argus* (Figure 2.5 and Figure 2.6; (Tadesse et al. 2011, Corey et al. 2013)) suggest that IR25a is expressed in most or all ORNs in the antennules and all CRNs but not MRNs in antennules, legs, and second antennae.

Intense labeling with anti-HaIR25a was previously reported for the clusters of ORN somata associated with the aesthetascs in *H. americanus, P. argus,* and the land hermit crab *Coenobita clypeatus*. This broad expression across the ORNs and CRNs in crustaceans is suggestive of a co-receptor function. In addition, a matching labeling pattern was found in *H. americanus, P. argus*, and *C. clypeatus* with *in situ* hybridization using specific antisense probes for IR25a (Stepanyan et al. 2004, Tadesse et al. 2011, Corey et al. 2013, Groh-Lunow et al. 2014). Our results corroborate the previous findings in *P. argus*. Furthermore, double labeling with anti-tubulin antibody, which appears to label all sensory neurons (albeit not selectively), demonstrates that in fact all ORN somata of a cluster are IR25a-positive, which also appears to be the case in *H. americanus* and *C. clypeatus*. The apparent ubiquitous and uniform (low variation in labeling intensity)

expression of IR25a in all ORNs in decapod crustaceans contrasts sharply with the expression pattern of chemoreceptor proteins in ORNs of *D. melanogaster* and other insects. First, olfactory sensilla of insects are differentiated into distinct morphological classes (and can occur on mouthpart appendages in addition to the antenna), and only one class (coeloconic sensilla) is innervated by IR-expressing ORNs whereas the majority of sensilla (trichoid and basiconic sensilla) are innervated by ORNs expressing ORs and/or GRs (Benton et al. 2009, Silbering et al. 2011). Furthermore, in *D. melanogaster*, IR-expressing ORNs are differentiated into some in which IR25a is expressed highly (intensely labeled by anti-DmIR25a - some additional ORNs are weakly labeled) and acts as co-receptor, and others in which IR8a is highly expressed (intensely labeled by anti-DmIR8a) and acts as co-receptor (Abuin et al. 2011). The different pattern suggests an evolutionary trajectory leading from a single class of ORNs in crustaceans characterized by expressing IRs, having IR25a as a co-receptor, and innervating a morphologically uniform sensillum type (aesthetascs), to multiple classes of ORNs in insects expressing chemoreceptor proteins of different types (IRs, ORs, GRs), having IR-expressing ORNs with different co-receptors (IR25a or IR8a), and innervating sensilla of distinct morphologies.

Clues about the possible functions of IR25a in the ORNs of *P. argus* are provided by the distribution of IR25a-like immunoreactivity in different cellular compartments. Of all labeled cellular compartments (outer dendritic segments, inner dendritic segments, somata, initial axon segments), the outer dendritic segments showed the highest intensity of anti-HaIR25a-like immunoreactivity, suggesting that they are the primary site of IR25a function. This finding corroborates a previous report that described IR25a expression in the outer dendritic segments of *P. argus* ORNs (Corey et al. 2013) and is in line with the

distribution of anti-DmIR25a-like immunoreactivity in *Drosophila* ORNs (Benton et al. 2009, Abuin et al. 2011). Since the outer dendritic segments are the sites of odorant binding and sensory transduction (Blaustein et al. 1993, Hatt and Ache 1994), identifying IR25a protein in this cellular compartment strongly suggests that it is part of the receptor complex (and hence directly contributes to sensory transduction) and/or it is involved in the trafficking of other IRs (divergent IRs) to the dendritic membrane as has been demonstrated for IR25a expressed in *Drosophila* ORNs (Abuin et al. 2011). If the outer dendritic segments are the primary location of IR25a in the ORNs, the weaker labeling of the somata and inner dendritic segments with anti-HaIR25a suggests that IR25a is synthesized in the somata and then transported through the inner dendritic segments to its final target. Labeling of the initial axon segments with anti-HaIR25a and the complete lack of labeling of the ORN axon terminals in the OLs of the brain matches the labeling pattern of *D. melanogaster* ORNs with anti-DmIR25a and anti-DmIR8a (Benton et al. 2009, Abuin et al. 2011). In contrast, ORN axon terminals are intensely labeled by anti-HaIR25a in *H. americanus* (Schmidt 2016), suggesting substantial interspecies differences in the function of IR25a in ORN axons of decapod crustaceans. In ORNs of mammals, ORs are expressed throughout the axons including their terminals (Barnea et al. 2004, Strotmann et al. 2004), and these axonal ORs contribute to axon targeting into select glomeruli (Mombaerts 2006). In insect ORNs, expression of ORs was detected in axons, but ORN axon targeting is OR-independent (Brochtrup and Hummel 2011) and it has not been determined if OR expression extends into the axon terminals in the antennal lobe (Elmore and Smith 2001, Gohl and Krieger 2006).

In addition to labeling ORNs, anti-HaIR25a intensely labeled other sensory neurons in all appendages or their parts (distal and proximal part of the lateral flagellum)

included in this study – a finding that has not been reported previously. In each case, 6–24 HaIR25a-posititve sensory neurons were located in spindle-shaped clusters of sensory neurons (delineated by anti-tubulin labeling) that contained 1–5 additional neurons at their apical pole that were not or only very weakly labeled by anti-HaIR25a. In number and location, these clusters of sensory neurons closely correspond to populations of diverse types of setae with morphological characteristics of bimodal chemo- and mechanosensory sensilla, but we could not unequivocally establish the connection of clusters of sensory neurons to particular setae, because of the long distance and weak labeling of the other dendritic segments. However, since we did not find any clusters of sensory neurons labeled by anti-tubulin that did not contain HaIR25a-positive neurons, we conclude that likely all bimodal chemo- and mechanosensory sensilla in the tested appendages of *P. argus* are innervated by a group of sensory neurons expressing IR25a and a few others that do not. Preliminary results from labeling other appendages (third maxillipeds, uropods, pleopods) with anti-HaIR25a strongly indicate that this finding can be extrapolated to the entire distributed chemoreception system of *P. argus* (M. Schmidt, S. Sparks, and C. Derby, unpublished data). Comparing the numbers of HaIR25a-positive and HaIR25a-negative sensory neurons in a cluster with the numbers of presumptive CRNs and MRNs innervating diverse bimodal chemo- and mechanosensory sensilla on the antennules (8–10 / 12–13 / 15 total number of sensory neurons, 3–4 of which are MRNs (Cate and Derby 2001, Cate and Derby 2002)) and mouthpart appendages (11–17 / 33–41, 1–4 of which are MRNs (Garm and Høeg 2006)) of *P. argus* indicates that anti-HaIR25a selectively labels CRNs while MRNs are HaIR25a-negative. All presumptive CRNs in the clusters of sensory neurons innervating bimodal sensilla on the second antenna and both parts of the LF are distinctly HaIR25a-positive (in parallel to the

situation in ORNs). However, this is not the case in the clusters of sensory neurons innervating dactyl sensilla. Here, 2–4 of the presumptive CRNs are only very weakly labeled by anti-HaIR25a, suggesting that they may express a different IR as primary co-receptor. In a subset of *D. melanogaster* IR-expressing ORNs, IR8a serves as co-receptor, and given the particularly intense bands of IR8a in the dactyl (Figure 2.8d), which though non-quantitative is consistent with high expression, allows for the possibility that IR8a may also be a co-receptor in dactyl sensilla.

Only in the dactyl did we identify single tubulin-positive bipolar sensory neurons that were HaIR25a-negative, suggesting that they represent MRNs innervating unimodal mechanosensory sensilla. In the tip region of dactyls of shore crabs, *Carcinus maenas*, unimodal mechanosensory sensilla innervated by two sensory neurons (called intracuticular sensilla) are known (Schmidt 1990), and it is possible that *P. argus* dactyls bear similar sensilla.

Our PCR results confirmed the expression of IR25a and two other co-receptors (IR8a, IR93a) in tissues containing ORNs and/or CRNs, including the proximal and distal regions of the LF, walking leg dactyls, and second antennae. PCR also confirmed the tissue specificity of one divergent IR expressed only in the LF (IR1028) and one IR expressed only in the dactyl (IR1074). Zbinden et al. (2017) used PCR to identify IR25a expression in LF and medial flagellum and second antennae of several species of shrimp, but contrary to our results in *P. argus*, they did not identify it in the walking legs or mouthparts of shrimp.

2.4.2.2IR25a in the CNS

In the brain, anti-HaIR25a exclusively and very distinctly labeled large cells in the lateral division of the antennular nerve (containing the axons of ORNs) and the axon sorting zone proximal to the OL. The axons and axon terminals of the ORNs, however, were completely devoid of labeling. This is unexpected, because the axons are clearly labeled further distally (at their origin from the ORN somata; see above) and because in *H. americanus*, the axon terminals of ORNs in the OL are distinctly HaIR25a-positive (Schmidt 2016). The distinct labeling by anti-HaIR25a allowed further characterization of the hitherto unknown morphological properties of these large cells. All cells possess at least one long, thin process, clearly identifying them as either neurons or glial cells (Schmidt and Derby 2011). The processes project in the direction of the OL, but because the labeling fades out rapidly (typically in less than 100 µm from the cell soma), the target of the processes and their terminal structures remain unknown. It is possible that thick fibers with multiglomerular arborizations in the outer fibrous layer of the olfactory lobe labeled by backfilling the antennular nerve that were interpreted as terminals of MRNs (Schmidt and Ache 1992) are in fact the projections of the large cells. Filling the large cells with an intracellular marker would be required to substantiate this speculation.

Even the seemingly simple question if the large cells are neurons or some type of glial cells has no straightforward answer at present. In his original description of these cells in the antennular nerve of the spiny lobster, *Palinurus vulgaris*, Herbst (1916) stated that they are neurons ('Ganglienzellen') based on being of similar size as neurons of the brain. However, since both neurons and glial cells of the spiny lobster brain are diverse in size (Schmidt and Derby 2011), this argument is not conclusive. Our data provide

additional evidence for a neuronal identity of the large cells in that their nuclei are almost spherical and have very loose heterochromatin as is typical of neuronal nuclei (Schmidt and Derby 2011). However, the failure of WGA, which has been established as a marker of neurons residing in the brain of *P. argus* (Schmidt and Derby 2011), to label the large cells strongly indicates that they do not represent CNS neurons. This leaves the possibility that they represent a specialized population of sensory neurons (which also fail to label with WGA: M. Schmidt, unpublished). A sensory function of these cells would well be in line with the expression of the chemoreceptor protein IR25a – but what this function may be is enigmatic and would have to be addressed by recording the activity of the large cells with electrophysiological or imaging methods. Sensory neurons with somata located in the CNS have been observed in arthropods, albeit very rarely (Bräunig and Hustert 1980, Bräunig 1982). In being located within a nerve root, the large cells to some extent resemble neurons located in nerve roots of the thoracic ganglia of *H. americanus* (Wallace et al. 1974, Evans et al. 1975). These cells are octopaminergic, have processes arborizing in the connective tissue sheath surrounding the nerves, and serve a neuromodulatory function. Given that in adult spiny lobsters new ORNs are continually being born and innervating the olfactory lobe (Schmidt 2014), it is tempting to speculate that these IR25a positive cells in the axon sorting zone might act in guiding or sorting of the axons of the new ORNs as they extend their processes into targets in the OL. This intriguing hypothesis deserved further examination.

In fact, our PCR identified iGluRs and IRs, including IR25a, not only in sensory organs and brain but also in the green gland of *P. argus*. The function of IRs in this excretory organ also needs further study, but this, together with our finding of IRs in the brain, suggests multiple roles for IRs.

2.4.3 Crustacean GRs

We identified a partial sequence of a GR in the LF of *P. argus* (Figure 2.9). GRs having expanded families and demonstrated or putative chemosensory function have now been identified in insects, some crustaceans (most notably *Daphnia*, with its 58 GRs), chelicerates (ticks), and myriapods (centipedes) (Peñalva-Arana et al. 2009, Chipman et al. 2014, Egekwu et al. 2014, Benton 2015, Robertson 2015, Saina et al. 2015, Eyun et al. 2017). The GRs and their non-arthropod homologues, the GRLs, have ancient origin, and they appear to have been lost in some groups and expanded in others (Benton 2015, Robertson 2015, Eyun et al. 2017). GRs are not just contact chemoreceptor proteins in *D. melanogaster* but are involved in other sensory functions such as promoting the detection of CO_2, sensing fructose in hemolymph, detecting light, and sensing warm temperatures. GR28B(D), together with TRPA1, mediates thermotaxis (Jones et al. 2007, Kwon et al. 2007, Thorne and Amrein 2008, Xiang et al. 2010, Miyamoto et al. 2012, Montell 2013, Ni et al. 2013). GRLs in ancestral protostomes appear to have roles in development (Saina et al. 2015), and their selective expansion in some clades appears to be related to chemosensory functions (Benton 2015, Robertson 2015, Eyun et al. 2017). GRs have been found in some crustaceans besides *Daphnia*, but not in high numbers – one in the barnacle *Amphibalanus amphitrite*, two in the copepod *Tigriopus californicus*, and six in the copepod *Eurytemora affinis* species (Eyun et al. 2017) – and all are of unknown function. Our identification of only one GR in *P. argus* suggests that while present in this species, it is unlikely to play a role in chemoreception, or at least in the discrimination of diverse chemical stimuli. Determining the cellular expression patterns of this GR will help in elucidating possible functions. In addition, transcriptomic analysis of additional

crustaceans and non-insect arthropods should help us understand the evolution and function of the GRL/GR gene family in this clade (Derby et al. 2016).

2.4.4 Crustacean TRP channels

Homologues of all subfamilies of TRP channels (Venkatachalam and Montell 2007, Venkatachalam et al. 2014) were found in our *P. argus* transcriptome (Figure 2.10). Here, we discuss only the four subfamilies of TRP channels that have known chemosensory functions in other species. First is the TRPA homologues, including those related to TRPA1, painless, TRPA5, pyrexia, and waterwitch of insects, found in LF, dactyl, and brain. The expansion of the TRPA subfamily in *P. argus* with nine different types is similar to insects. Second is the TRPV homologues, including those related to OSM-9 of *C. elegans*, and Nanchung and Inactive of *D. melanogaster*. Third is the TRPC homologues in LF, dactyl, and brain. Fourth is the TRPM homologues, including those related to insect TRPM channels and mammalian TRPM channels, in LF, dactyl, and brain.

Although the types and numbers of *P. argus* TRP channels are similar to insects, there are a few differences in crustaceans. The crustaceans *P. argus*, *H. americanus*, and *C. borealis* have 2–3 types of TRPM channels, whereas insects have only one (Figure 2.10; (Matsuura et al. 2009, Peng et al. 2015, McGrath et al. 2016, Northcutt et al. 2016)). While *Daphnia* also has two TRPM channels, they are both cluster with the insect TRPM channels (Figure 2.10; (Peng et al. 2015)). One TRPM channel expressed in the LF, dactyl, and brain of *P. argus* is homologous to TRPM channel of insects, while the other TRPM homologue, PargTRPMm, is highly expressed in the LF and dactyl and is more closely

related to mammalian TRPM channels. It is possible that with the inclusion of other decapod crustaceans in a phylogenetic analysis, PargTRPMm may turn out to be a crustacean specific TRPM channel and not a true homologue of mammalian TRPM channels.

The cellular expression patterns of TRP channels in *P. argus* were not explored, so their role in chemoreception remains speculative. We note that a sodium/calcium gated cation channel in ORNs of *P. argus* has certain physiological and pharmacological properties that resemble a TRPC channel (Bobkov and Ache 2005, Pezier et al. 2009, Bobkov et al. 2010). However, the sequence of this channel has not been reported, and the only sequences of crustacean TRP channels that we have found in public databases are for TRPA, TRPM, and TRPV channels from transcriptomes of the central nervous system of *C. borealis* and *H. americanus* (McGrath et al. 2016, Northcutt et al. 2016) and all the TRP channels from the *D. pulex* genome (Peng et al. 2015).

Many temperature sensitive, or thermo, TRPs, which include TRP channels TRPV1, TRPV2, TRPV3, TRPV4, and TRPM8 in mammals and TRPA1 across animals, are also activated by irritants or noxious chemicals such as reactive oxygen or nitrogen species (ROS or RNS) (Yoshida et al. 2006, Arenas et al. 2017), electrophiles such as allyl isothiocyanate (AITC), capsaicin, menthol, and camphor. The diversity of the structures of these thermo TRP activating molecules suggests various activation mechanisms that are not restricted to the receptor-ligand model that is commonly observed in most chemoreceptor proteins. TRPA1 channels can be activated following covalent modification of their cysteine residues by membrane permeable molecules such as AITC, nitric oxide, RNS, or H_2O_2 (Hinman et al. 2006, Yoshida et al. 2006, Bandell et al. 2007,

Macpherson et al. 2007, Arenas et al. 2017). In fact, Arenas et al. (2017) discovered that planarian and human TRPA1 channels can rescue *D. melanogaster* TRPA1 function in mutants despite planarians, flies, and humans being separated by millions of years of evolution, due to the shared commonality of being activated by ROS and H_2O_2. Although TRPC channels are not known to be thermo TRPs, they have been implicated in chemical sensing, and several TRPC channels are known to be sensitive to NO and H_2O_2 (Yoshida et al. 2006). Previous work has shown that H_2O_2 and other products found in opaline glands of the sea hare, *Aplysia californica*, are aversive and distasteful to *P. argus* (Aggio and Derby 2008). The receptor proteins mediating this aversion response are unknown, but RNS/ROS and H_2O_2 sensitive receptors such as TRPA1 and TRPC channels, both of which are in our Parg transcriptomes, are candidates.

2.4.5 Other chemoreceptor proteins in crustaceans?

We searched for homologues of other chemoreceptor proteins in our transcriptome. We did not find ORs in *P. argus*, which is consistent with past failures to identify ORs in non-insect arthropods, including crustaceans. Our results support the conclusion that ORs evolved after the origin of insects (Missbach et al. 2014, Missbach et al. 2015). We also searched for an expanded set of nicotinic acetylcholine receptors (nAChRs), such as is present in the *Octopus* genome and which speculatively might be a set of chemoreceptor proteins (Albertin et al. 2015). While we found several nAChR homologues, we did not find evidence of an expanded family. Another class of chemoreceptor proteins, vertebrate-like ORs, are 7TM GPCRs. While there are several rhodopsin-like GPCRs in the *P. argus* transcriptome, our initial InterProScan and BLAST

searches did not reveal any chemosensory GPCRs. Perhaps the inclusion of other decapods in phylogenetic analyses may resolve some of these 7TM GPCRs to be chemosensory.

2.5 Conclusions

The Caribbean spiny lobster, *P. argus*, has at least 108 IRs in two of its major chemosensory organs, the LF and dactyls of legs. Of these 108 IRs, four are co-receptor IRs and one is a conserved IR that are expressed in both LF and leg dactyl. The other 95 are divergent IRs, and most (51) are expressed only in the LF. Supporting the role of these IRs in chemoreception is the fact that the co-receptor IR25a is expressed in chemosensory cells in these organs – the ORNs of the LF and CRNs in the dactyls and LF – but they are not expressed in MRNs in either organ. It is interesting, though, that the co-receptor IRs and conserved IRs are expressed in other tissues including brain, which may be related to their demonstrated role in insects in non-chemosensory functions. Besides IRs, we found one GR and 18 TRP channels in *P. argus*, though any function in chemoreception in *P. argus* is unknown at this time.

Many questions remain unanswered. The much higher number of divergent IRs in the LF compared to dactyls is correlated with the unique role of LF in detecting waterborne conspecific cues including social, alarm, and sex cues, in addition to their detecting feeding cues also sensed by the dactyls. However, the chemical specificity of individual IRs needs to be determined to identify their broader functional roles. Related to this is the need to determine the expression patterns of IRs in individual cells. For example, how many IRs are co-expressed in individual ORNs or CRNs, and what is the

diversity of expression patterns in these cells? Each aesthetasc of *P. argus* is innervated by ca. 300 ORNs with a diversity of chemical specificities, such that the aesthetasc is considered to be a functional unit of olfaction (Mellon 1990, Mellon and Munger 1990, Steullet et al. 2000). Correlating IR expression pattern with chemical specificity for individual ORNs will be important in determining rules of olfactory coding in the periphery.

How many different types of chemoreceptor proteins exist in a crustacean? For *P. argu*s, we have sampled only two sensory organs, and the conservative answer is nearly 130, including IRs, GRs, and TRP channels. Given that the IR populations in *P. argus* are largely non-overlapping in LF and dactyls, it might be expected that when other chemosensory organs are analyzed, the total number of chemoreceptor proteins in *P. argus* will be much higher. How does this compare with other crustaceans? Analysis of the genome of *Daphnia* reveals at least 143 chemoreceptor proteins: 58 GRs and 85 IRs (Peñalva-Arana et al. 2009, Croset et al. 2010, Rytz et al. 2013, Benton 2015). Two species of copepods have been reported to have 8 IRs and 2-6 GRs (Eyun et al. 2017). Transcriptomes of the antennule from two species of hermit crabs have yielded up to 29 IRs and no ORs or GRs per species (Groh et al. 2013, Groh-Lunow et al. 2014). A better understanding of the olfactory logic in crustaceans, including differences associated with phylogeny, sex, or development, will benefit from an examination of more crustacean species.

Finally, important questions remain unanswered regarding how the chemoreceptor proteins are represented in the spiny lobster's central nervous system through the central projections of ORNs. The antennule's 300,000 ORNs with their ca. 96 divergent IRs project into ca. 1200 glomeruli in the OL of the brain. This suggests the

possibility of an olfactory wiring logic in decapod crustaceans that is significantly different than in insects or mammals, in which the ratio of the number of types of receptor molecules to glomeruli is 1:1 (Vosshall 2001, Galizia and Sachse 2010, Murthy 2011).

2.6 Figures and Tables

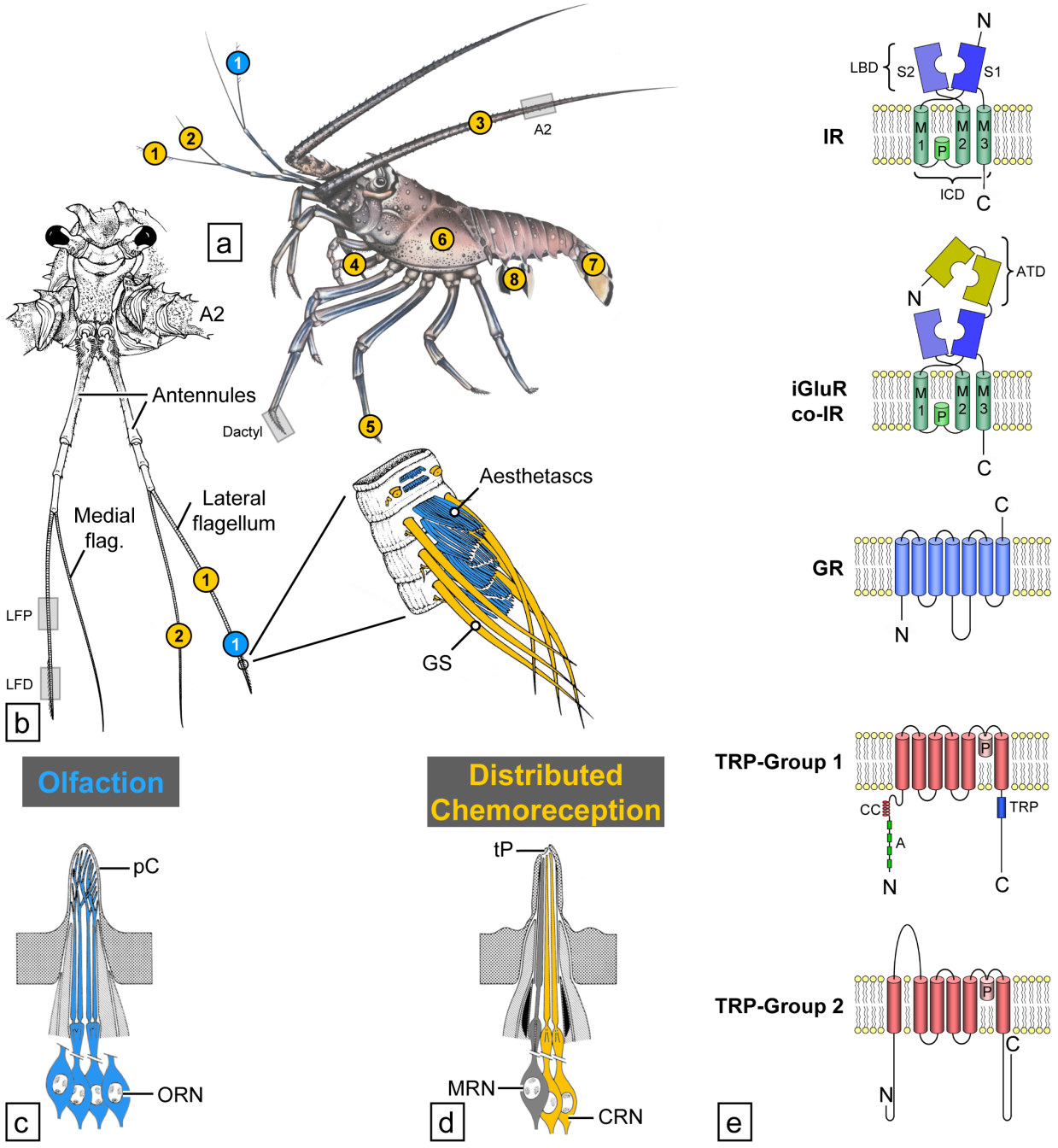

Figure 2.1 Overview of spiny lobster chemosensory systems.
Modified from Derby et al. (2016). **(a)** Location of aesthetascs mediating olfaction (blue dots) and bimodal chemo- and mechanosensory sensilla mediating distributed chemoreception (yellow dots) on different body parts and appendages of *P. argus* (1 -

lateral flagellum of antennule, 2 - medial flagellum of antennule, 3 - second antenna, 4 - mouthpart appendages, 5 - walking legs, 6 - gill chamber, 7 - tail fan, 8 - pleopods). Location of pieces of appendages used for immunocytochemistry and PCR indicated by gray boxes: dactyl, 2nd antenna (A2). **(b)** Location of aesthetascs and bimodal chemo- and mechanosensory sensilla on the antennules. Aesthetascs (blue) are restricted to a tuft of sensilla on the distal third of the lateral flagellum. Bimodal chemo- and mechanosensory sensilla (yellow) among them, guard setae (GS) are associated with the aesthetascs but also occur on the proximal part of the lateral flagellum and on the entire medial flagellum. Location of pieces of appendages used for immunocytochemistry and PCR indicated by gray boxes: lateral flagellum of antennule proximal (LFP), lateral flagellum of antennule distal (LFD). **(c)** Schematic drawing of the cellular organization of olfactory sensilla. Olfactory sensilla, called aesthetascs, are exclusively innervated by olfactory receptor neurons (ORN, blue). The outer dendritic segments of the ORNs (modified cilia) are highly branched and covered by extremely thin and permeable cuticle (pC). **(d)** Schematic drawing of the cellular organization of distributed chemosensilla. Distributed chemosensilla are bimodal chemo- and mechanosensory sensilla innervated by a few mechanoreceptor neurons (MRN) and several chemoreceptor neurons (CRN). Dendrites of CRNs are unbranched and extend to a terminal pore (tP) at the tip of the sensillum. **(e)** Schematic drawing of the molecular structure of chemoreceptor proteins (iGluRs and Co-IRs, IRs, GRs, and TRP channels). Transmembrane domains of iGluRs and IRs (M1 – M3), pore loop (P), ligand binding domains (S1, S2), amino terminal domain (ATD), coiled-coil domain (CC), ankyrin repeats (A), TRP domain (TRP).

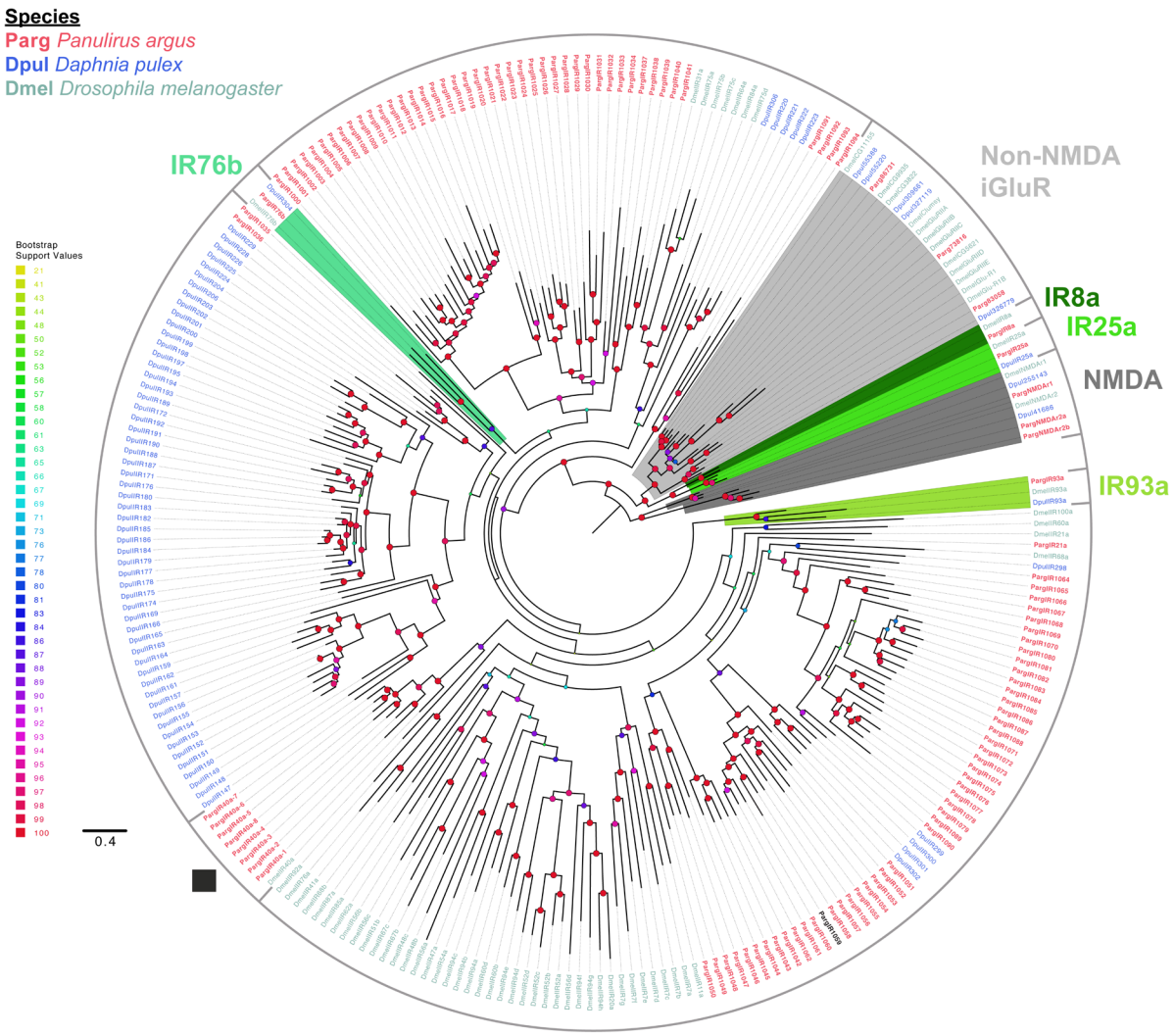

Figure 2.2 Maximum likelihood phylogenetic tree of iGluRs and IRs.

Colored areas of the tree represent iGluRs (shades of grey) and co-receptor IRs (shades of green). ■ indicates conserved IR sequences of the IR40a family. Sequence colored in black indicates the Parg divergent IR that has an amino acid substitution in the conserved glutamate binding site in the S1 region of the LBD. The tree was built under LG+F+G4 model of substitution with 1000 ultrafast bootstrap (UFBoot) replications and visualized on FigTree v1.4.2. Bootstrap values on some internal branches of divergent IRs are low due to incomplete sequences and high sequence divergence across the different species. The tree was unrooted but drawn with the NMDA clade as root. The scale bar represents expected number of substitutions per site.

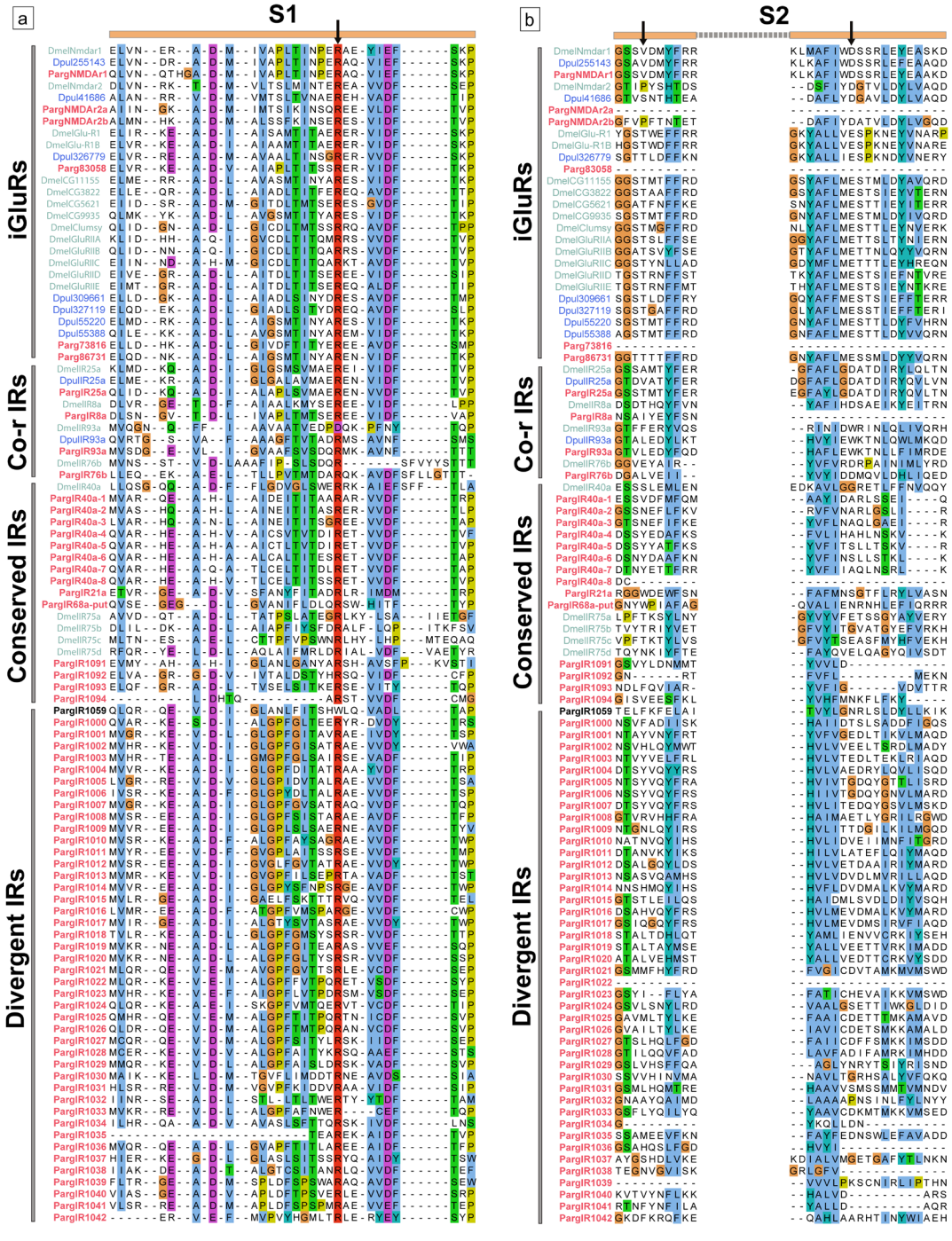

Figure 2.3 Multiple sequence alignment of iGluRs and IRs of Parg, Dmel, and Dpul. This analysis shows that the LBD is highly divergent within the IRs and in comparison to iGluRs. Parg (red), Dmel (mint), and Dpul (blue) sequences are organized based on sequence homology in Figure 2.2. Divergent IR, PargIR1059, an exception to 'R'

conservation in S1 lobe, is indicated in black. Glutamate binding sites are indicated with arrows. **(a)** S1 lobe of LBD. **(b)** S2 lobe of LBD. The sequences were aligned with MAFFT and visualized on Jalview. The residues were colored according to the Clustal X color scheme on Jalview. Criteria for the color scheme: http://www.jalview.org/help/html/colourSchemes/clustal.html

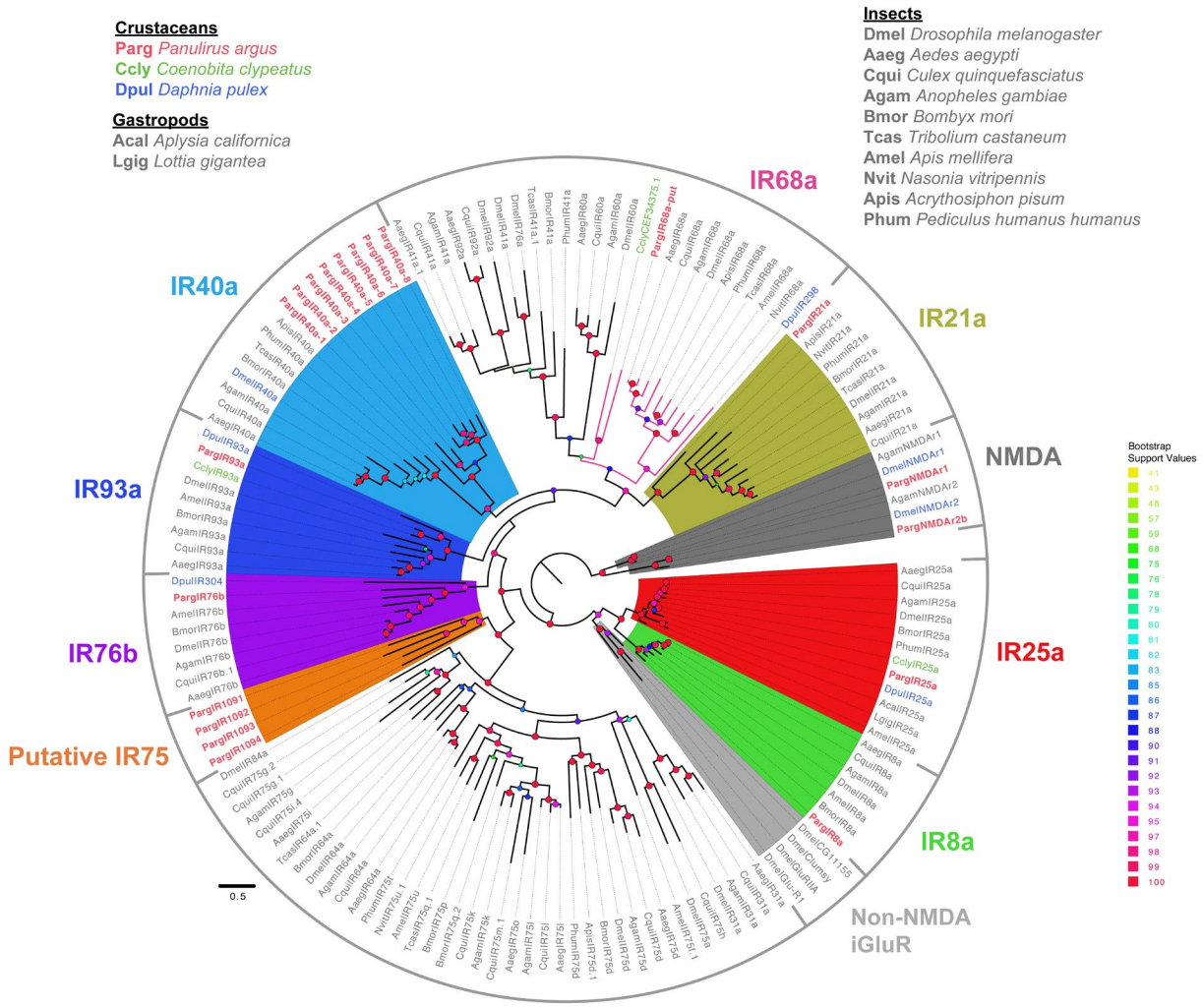

Figure 2.4 Maximum likelihood phylogenetic tree of homologous sequences of conserved IRs.

Only conserved IR groups with Parg homologues are colored. Pink branches represent homologous sequences of IR68a with the Ccly sequence, the putative PargIR68a-put, and DpulIR298 as crustacean representatives. Sequences were aligned with MAFFT and visualized on Jalview for editing. The tree was built on IQ-Tree under LG+F+G4 model of substitution with 1000 UFBoot replications and visualized on FigTree v1.4.2. Bootstrap values in some internal nodes are low due to incomplete sequences. The tree was unrooted but drawn with the NMDA clade as the root. The scale bar represents the expected number of substitutions per site.

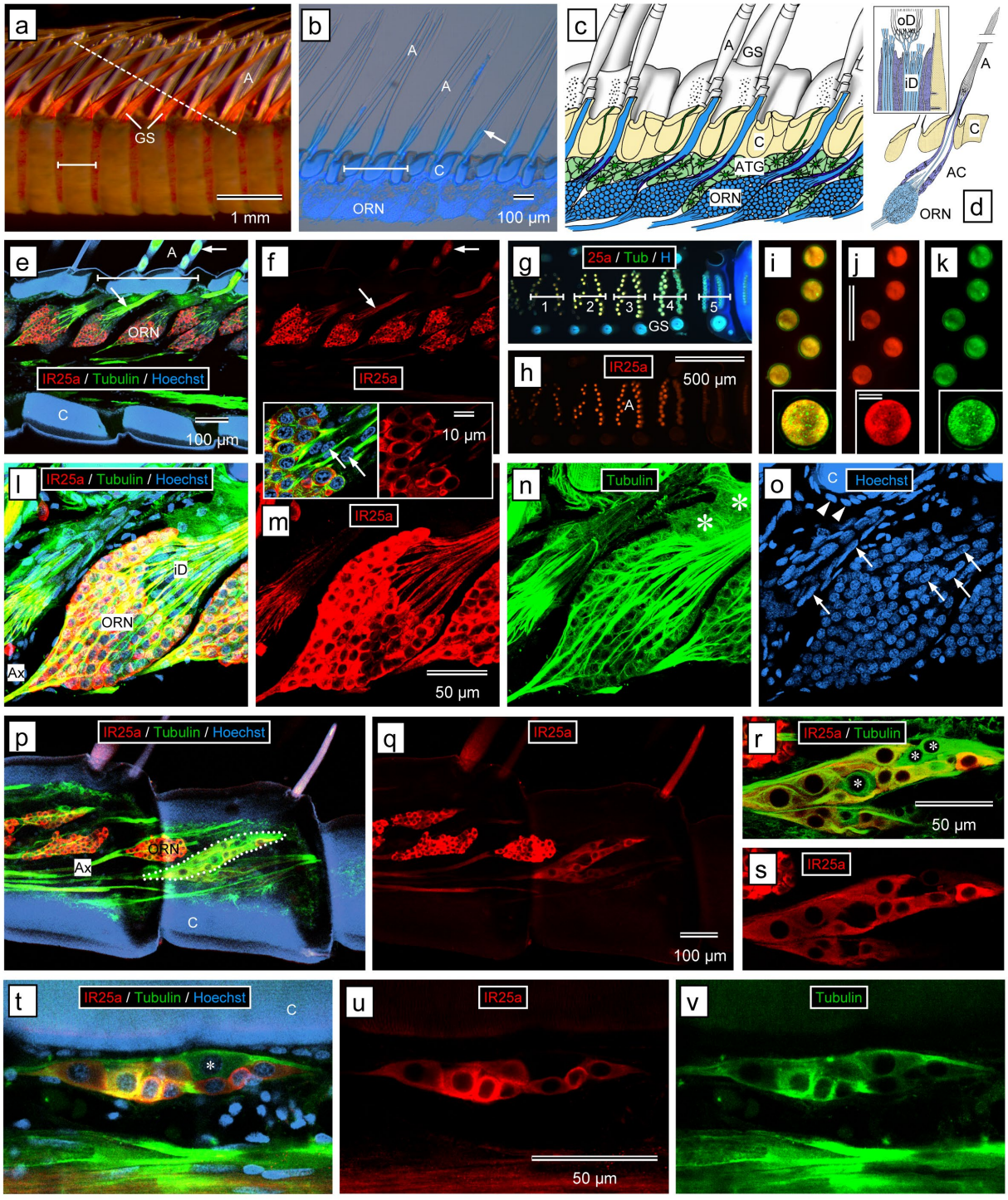

Figure 2.5 Immunolabeling with anti-HaIR25a in the lateral flagellum of the antennule. **(a–s)** Aesthetascs-bearing tuft region of the lateral flagellum. **(a)** Outer morphology - stereomicroscopical image. Each annulus (horizontal bar) bears two rows of transparent aesthetascs (A) accompanied by guard setae (GS). Dashed line shows direction of cross

section through aesthetascs in **(g–k)**. **(b)** Sagittal section labeled with Hoechst 33258 showing full length of aesthetasc setae (A). Each annulus (horizontal bar) bears two rows of aesthetascs. ORN clusters (ORN) are labeled by Hoechst whereas cuticle (C) is autofluorescent. Only the proximal 20% of the cuticle of the aesthetasc setae containing the inner dendritic segments of the ORNs is autofluorescent (arrow). **(c)** Sagittal view of the lateral flagellum (modified from Schmidt et al. (2006)). Each aesthetasc (A) is innervated by a cluster of ORNs (blue) whose inner dendritic segments traverse the cuticle (C) in a wide canal. ORN clusters are associated with subcuticular aesthetasc tegumental glands (ATG) whose thin drainage ducts terminate in pores at the base of the aesthetascs. Guard setae (GS) are located at the lateral margins of the aesthetasc rows. **(d)** Reconstruction of aesthetasc ultrastructure based on transmission electron microscopy; inset: region of transition from inner to outer dendritic segments within the base of aesthetasc setae at higher magnification (modified from Grünert and Ache (1988)). Each aesthetasc (A) is innervated by about 320 ORNs whose somata form a cluster (ORN) below the cuticle (C). Inner dendritic segments arising apically from the ORN somata are wrapped by auxiliary cells (AC). Inset: each inner dendritic segment (iD) gives rise to two highly-branched outer dendritic segments (oD). **(e, f)** Sagittal section through medial plane of the tuft region of a lateral flagellum labeled with anti-HaIR25a (red), anti-tubulin (green), and Hoechst 33258 (blue) at low magnification (confocal images). Scale bar in **(e)** also applies to **(f)**. **(e)** Overlay of all three fluorescence channels. **(f)** anti-HaIR25a channel. Two rows of aesthetascs setae (A) arise from the intensely autofluorescent (blue) cuticle of an annulus (horizontal bar in **e**). Each aesthetasc seta is associated with a clearly delineated cluster of ORN somata (ORN). The somata and inner dendritic segments (arrows) are intensely labeled by anti-HaIR25a and anti-tubulin. **(g–k)** Cross sections through aesthetasc setae (direction of section indicated by dashed line in **(a)**). **(g, h)** Low magnification epifluorescence images; scale bar in **(h)** also applies to **(g)**. **(g)** Overlay of all three fluorescence channels. **(h)** anti-HaIR25a channel. Aesthetasc setae (A) located on different annuli (horizontal bars with numbers) are cut at different levels systematically progressing from the tips (1) to the base emerging from the annulus cuticle (5). Note that the lumen of the aesthetasc setae containing outer dendritic segments (1, 2, 3, 4 - left aesthetasc row) is more intensely labeled by anti-HaIR25a than the lumen of aesthetasc setae containing only inner

dendritic segments (4 – right aesthetasc row, 5). **(i–k)** High magnification epifluorescence images of sections through aesthetasc setae on annulus 3; insets: very high magnification (confocal images); scale bars in **(j)** (top: 50 μm; bottom: 10 μm) also apply to **(i)** and **(k)**. **(i)** Overlay of anti-HaIR25a and anti-tubulin fluorescence channels. **(j)** anti- HaIR25a channel. **(k)** anti-tubulin channel. Note that entire lumen of aesthetasc setae is filled by outer dendritic segments of ORNs intensely labeled by anti-HaIR25a and anti-tubulin. **(l–o)** Sagittal section through ORN clusters labeled with anti-HaIR25a (red), anti-tubulin (green), and Hoechst 33258 (blue) at high magnification (confocal images – maximum intensity projection of entire stack of optical sections spanning about 60 μm). Scale bar in **(m)** also applies to **(l)**, **(n)**, and **(o)**. **(l)** Overlay of all 3 fluorescence channels. **(m)** anti-HaIR25a channel. **(n)** anti-tubulin channel. **(o)** Hoechst channel. Insets in **(m)**: Apical region of ORN cluster at high magnification; left, overlay of all 3 fluorescence channels; right, anti-HaIR25a channel (confocal images of a single optical section). The somata (ORN), axons (Ax), and inner dendritic segments (iD) of all ORNs are intensely labeled by anti-HaIR25a and anti-tubulin. Somata of auxiliary cells (nuclei indicated by arrows) are not labeled by either of the antibodies and epithelial cells (nuclei indicated by arrowheads) and aesthetasc tegumental glands (asterisks) are labeled by anti-tubulin but not anti-HaIR25a. Autofluorescent (blue) cuticle (C). **(p–s)** Sagittal section through lateral plane of tuft region of lateral flagellum labeled with anti-HaIR25a (red), anti-tubulin (green), and Hoechst 33258 (blue). **(p, q)** Section at low magnification (confocal images); scale bar in **(q)** also applies to **(p)**. **(p)** Overlay of all three fluorescence channels. **(q)** anti-HaIR25a channel. In addition to clusters of ORN somata, clusters of sensory neurons innervating bimodal sensilla accompanying the aesthetascs (one cluster outlined by white dots) are also labeled by anti-HaIR25a and anti-tubulin. The intensity of labeling with anti-HaIR25a is higher in ORN somata and axons (Ax) than in somata and axons of the other sensory neurons. **(r, s)** Section at high magnification (confocal images; scale bar in **(r)** also applies to **(s)**. **(r)** anti-HaIR25a and anti-tubulin channel. **(s)** anti-HaIR25a channel. All sensory neurons of the clusters are labeled by anti-tubulin, but three of them (asterisks) mostly located in the distal aspect of the cluster are not double-labeled by anti-HaIR25a. **(t–v)** Sagittal section through cluster of sensory neurons innervating a bimodal sensillum in the proximal part of the lateral flagellum labeled with anti-HaIR25a (red), anti-tubulin (green) and Hoechst

33258 (blue). **(t)** Overlay of all three fluorescence channels. **(u)** anti-HaIR25a channel. **(v)** anti-tubulin channel. All sensory neurons of the clusters are labeled by anti-tubulin but one of them (asterisk) located in the distal aspect of the cluster is not double-labeled by anti-HaIR25a. The intensity of labeling with anti-HaIR25a and anti-tubulin differs substantially but independent of each other between labeled somata.

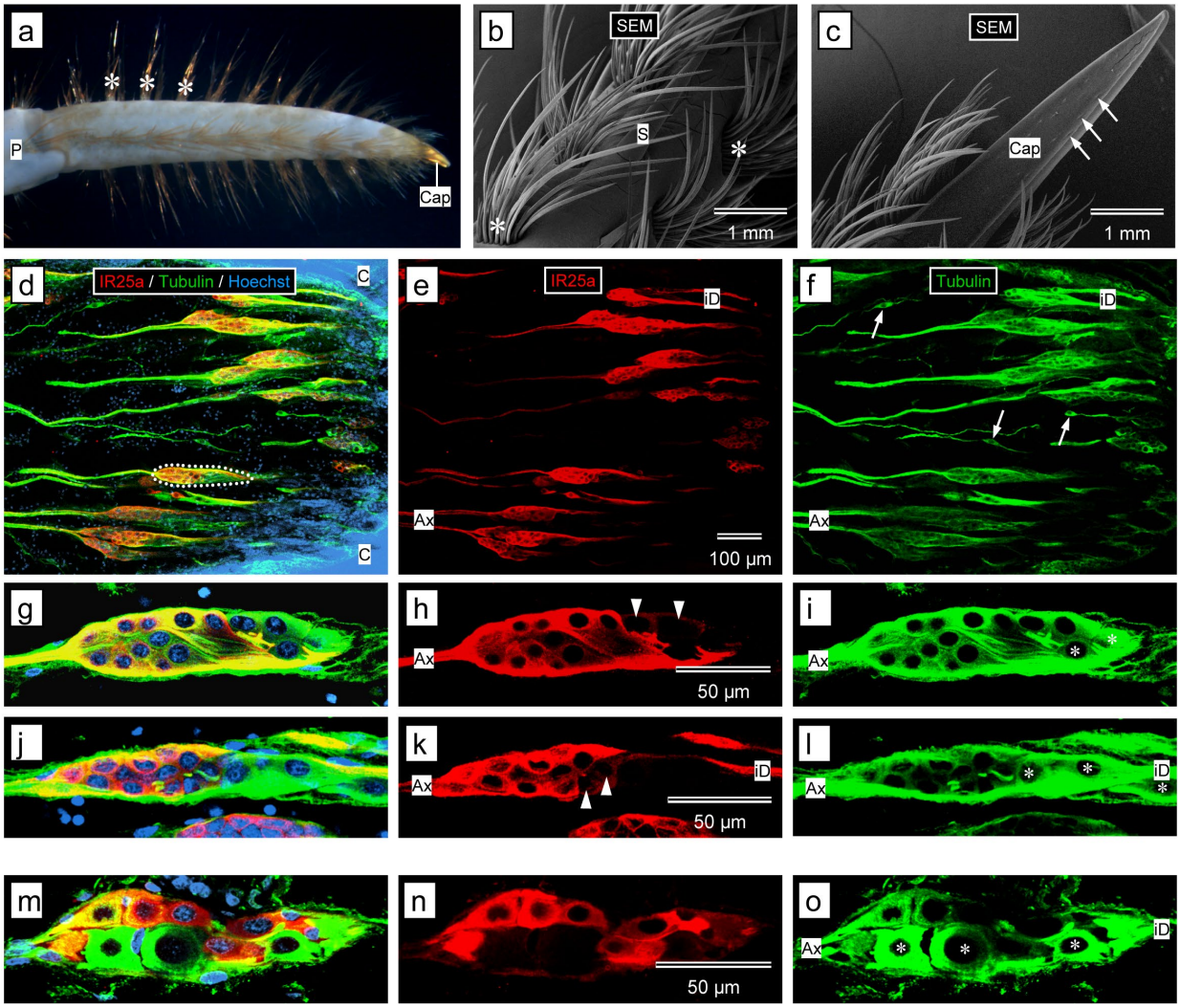

Figure 2.6 Immunolabeling with anti-HaIR25a in the walking leg dactyl and in the flagellum of the 2nd antenna.

(a – l) Walking leg dactyl. **(a)** Outer morphology of the dactyl of a third pereiopod of a late juvenile animal, shown in a stereomicroscopical image. The dactyl bears rows of evenly spaced bundles of smooth setae (asterisks), except on the epicuticular cap (Cap) at the tip. Propodus (P). **(b, c)** Outer morphology of the distal part of the dactyl of a second pereiopod of a late juvenile animal, shown in scanning electron micrographs (SEM). The main body of the dactyl bears rows of dense bundles of about 20 smooth setae (asterisks). Single smooth spines (S) are located between rows of bundled setae. The epicuticular cap (Cap) does not bear setae but instead holds numerous small depressions (arrows) that likely represent the outer structures of bimodal sensilla called funnel-canal organs. **(d –**

f) Sagittal section through the distal aspect of a third pereiopod dactyl (proximal to epicuticular cap) labeled with anti-HaIR25a (red), anti-tubulin (green), and Hoechst 33258 (blue) at low magnification (confocal images); scale bar in **(e)** also applies to **(d)** and **(f)**. **(d)** Overlay of all three fluorescence channels. **(e)** anti-HaIR25a channel. **(f)** anti-tubulin channel. Numerous spindle-shaped clusters of sensory neurons (one outlined by white dots), each innervating one of the smooth bundled setae are intensely labeled by anti-HaIR25a and anti-tubulin. Both antibodies label the somata of sensory neurons as well as their axons (Ax) and inner dendritic segments (iD). Overlay of all three channels reveals that in the clusters of sensory neurons, neurons that are labeled by anti-tubulin but not by anti-HaIR25a (and therefore appear green) are located at the distal pole of the clusters. Single bipolar sensory neurons labeled by anti-tubulin but not anti-HaIR25a (arrows) are interspersed between clusters of sensory neurons. **(g – l)** Clusters of sensory neurons labeled with anti-HaIR25a (red), anti-tubulin (green), and Hoechst 33258 (blue) at high magnification (confocal image); scale bar in **(h)** also applies to **(g)**, **(i)**; scale bar in **(k)** also applies to **(j)** and **(l)**. **(g, j)** Overlay of all three fluorescence channels. **(h, k)** anti-HaIR25a channel. **(i, l)** anti-tubulin channel. Each cluster contains about 15 bipolar sensory neurons, all strongly labeled by anti-tubulin. Neurons in the proximal part of the cluster are also intensely labeled by anti-HaIR25a, but two or three neurons located at the distal pole of the cluster are not labeled by anti-HaIR25a and two other neurons in the distal region are only weakly labeled by anti-HaIR25a (arrowheads). Axons (Ax), inner dendritic segments (iD). **(m – o)** Flagellum of 2nd antenna. Cluster of sensory neurons labeled with anti-HaIR25a (red), anti-tubulin (green), and Hoechst 33258 (blue) at high magnification (confocal images); scale bar in **(n)** also applies to **(m)** and **(o). (m)** Overlay of all three fluorescence channels. **(n)** anti-HaIR25a channel. **(o)** anti-tubulin channel. All sensory neurons of the cluster are labeled by anti-tubulin, although to different degrees. All but 3 sensory neurons are also labeled by anti-HaIR25a. Two of the HaIR25a-negative neurons (asterisks) are the largest neurons of the clusters suggesting that they are MRNs.

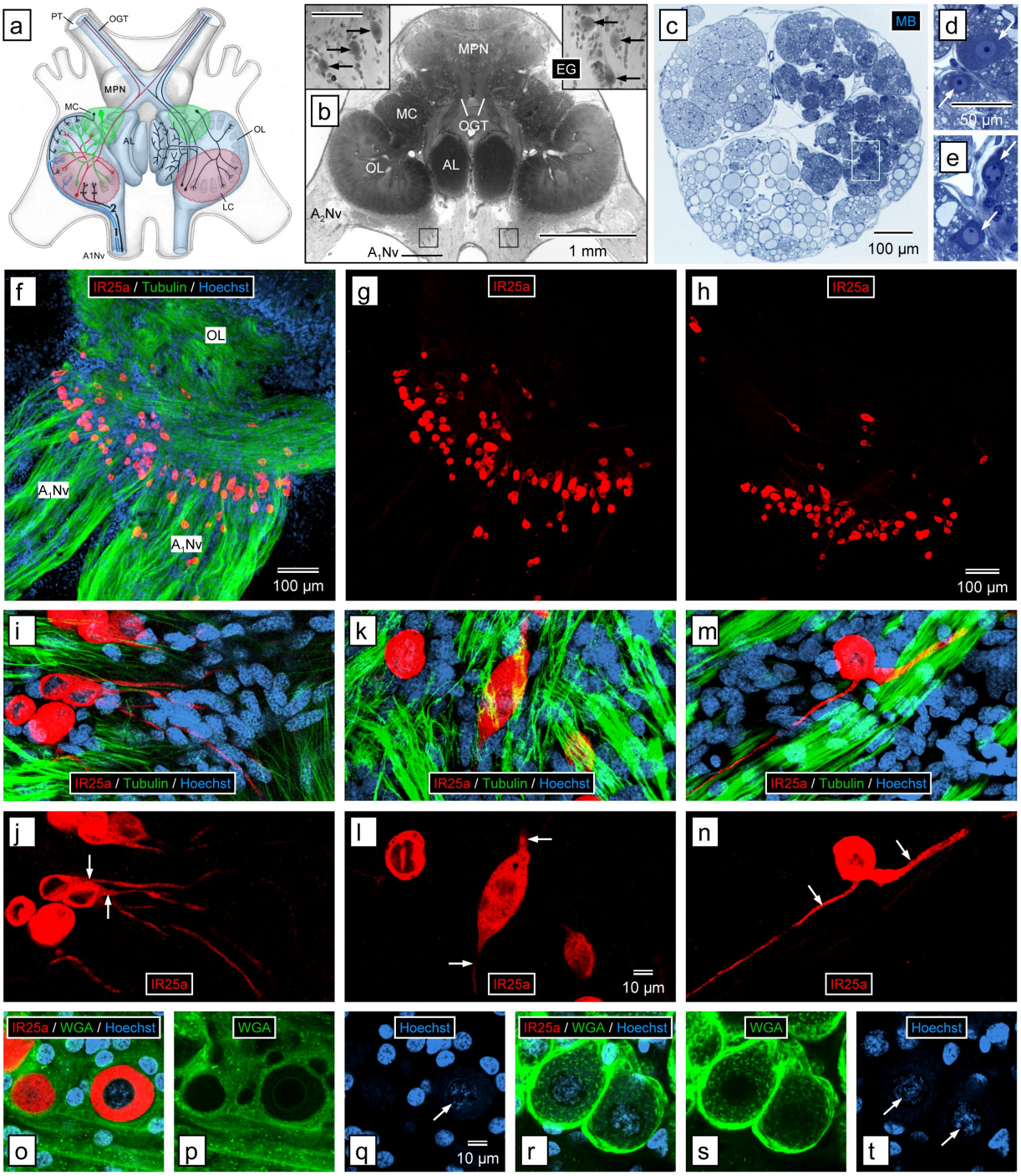

Figure 2.7 Immunolabeling with anti-HaIR25a in the brain.

(a) Schematic drawing of the olfactory pathway (light blue overlay) in the brain of *P. argus* (modified from Schmidt and Ache (1996)). Afferent axons of ORNs (1, blue) enter the brain via the antennular nerve (A_1Nv) and project to the ipsilateral olfactory lobe (OL) where they terminate in one of its cone-shaped glomeruli. The OL is closely linked to

another glomerular neuropil, the accessory lobe (AL). OL and AL are innervated by local interneurons (green) whose somata form the medial soma clusters (MC) and ascending projection neurons (red) whose somata form the lateral soma clusters (LC). Axons of projections neurons form the olfactory glomerular tracts (OGT) that run within the protocerebral tracts (PT) connecting the brain with the eyestalk ganglia. Median protocerebral neuropils (MPN). **(b)** Horizontal section through brain stained with ethyl gallate (EG). Large cells (arrows in insets) are located in the axon sorting zone of the A_1Nv before it reaches the OL. Location of insets is shown by black squares; scale bar in left inset is 100 µm and also applies to right inset. Antenna 2 nerve (A_2Nv). **(c – e)** Cross section through the antennular nerve where it enters the brain, stained with methylene blue (MB). **(c)** Low magnification. The axon fascicles in the antennular nerve form three large divisions. The lateral division is formed by axon fascicles from the lateral flagellum that are more intensely stained than other axon fascicles (because they contain numerous extremely thin ORN axons). Large, intensely stained cells selectively occur in the lateral division (white rectangle). **(d, e)** Large, intensely stained cells (arrows) in the lateral division of the antennular nerve at higher magnification. **(e)** Region highlighted in **(c)**. The large cells have voluminous cytosol and a spherical nucleus containing at least one dense nucleolus. **(f, g)** Confocal image of a sagittal section through brain labeled with anti-HaIR25a (red), anti-tubulin (green), and Hoechst 33258 (blue) at low magnification); scale bar in **(f)** also applies to **(g)**. **(f)** Overlay of all three fluorescence channels. **(g)** anti-HaIR25a channel. Anti-HaIR25a intensely labels a loose assembly of about 100 large cells located in the axon sorting zone of the A1Nv before it reaches the OL. Axons within the antennular nerve are intensely labeled by anti-tubulin but not by anti-HaIR25a, and HaIR25a-positive cells are not labeled by anti-tubulin. **(h)** Confocal image of a sagittal section through brain of different animal labeled with anti-HaIR25a at low magnification () shows a similar assembly of about 100 intensely labeled large cells in the axon sorting zone of the antennular nerve. **(i – n)** Morphology of single large cells in the axon sorting zone of the antennular nerve labeled with anti-HaIR25a (red), anti-tubulin (green), and Hoechst 33258 (blue) at high magnification (confocal images). Scale bar in **(l)** applies to all images. **(i, k, m)** Overlay of all three fluorescence channels. **(j, l, n)** anti-HaIR25a channel. **(i, j)** Unipolar cells. Most of the HaIR25a-positive cells have one process (arrows) projecting from the cell body. Generally, this process projects

toward the OL. **(k, l)** Bipolar cell. Some of the HaIR25a-positive cells have two processes (arrows) projecting from both poles of the cell body. **(m, n)** Pseudo-unipolar cell. Rarely HaIR25a-positive cells have two processes (arrows) projecting from the same region of the cell body. **(o – t)** Double-labeling with anti-HaIR25a (red) and WGA-AF488 (Hoechst 33258 - blue) at high magnification (confocal images). Scale bar in **(q)** applies to all images. **(o, r)** Overlay of all three fluorescence channels. **(p, s)** WGA-AF488 channel. **(q, t)** Hoechst 33258 channel. **(o–q)** Large cells in the axon sorting zone of the OL. **(r – t)** Neuronal somata in the medial soma cluster. Large cells in axon sorting zone are intensely labeled by anti-HaIR25a but not by WGA whereas somata in the MC are not labeled by anti-HaIR25a but are intensely labeled by WGA. Nuclei of large cells and neurons (arrows) are similar in shape (spherical) and in having very loose heterochromatin (Hoechst labeling of low intensity).

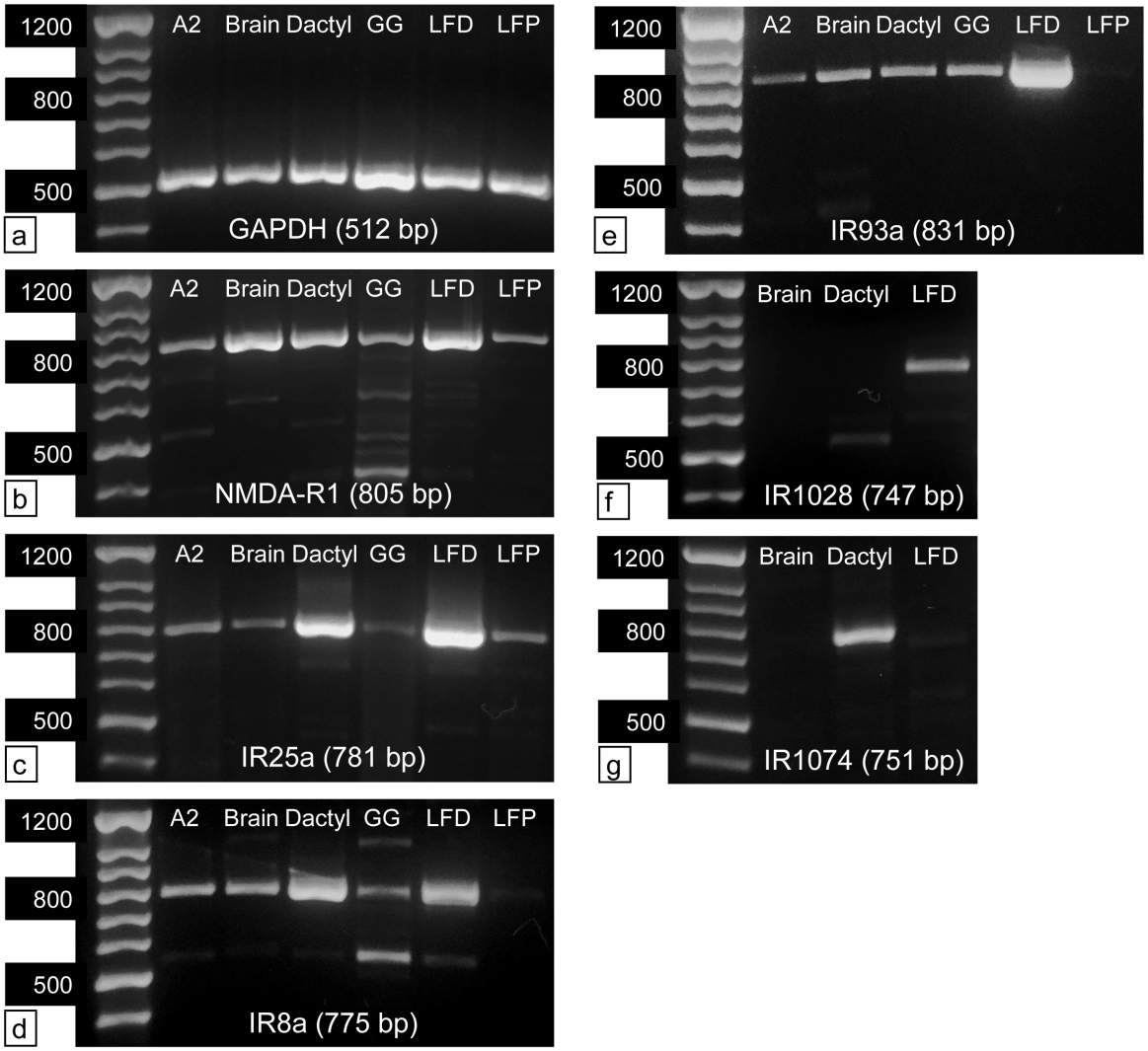

Figure 2.8 PCR results.

(a – e) Gel images of PCR products amplified from all target tissues, antenna 2 (A2), central brain (Brain), dactyl of second pereiopod (Dactyl), green gland (GG), aesthetasc-bearing tuft region of the distal lateral flagellum of antennule (LFD), and proximal region of the lateral flagellum of antennule (LFP), using specific primer pairs for GAPDH **(a)**, NMDA-R1 **(b)**, IR25a **(c)**, IR8a **(d)**, IR93a **(e),** PargIR1028 **(f)** found only in LF; PargIR1074 **(g)** found only in dactyl. The predicted amplicon length of the PCR product for each primer pair is given in parentheses. The left side of each gel shows a 100 bp DNA ladder.

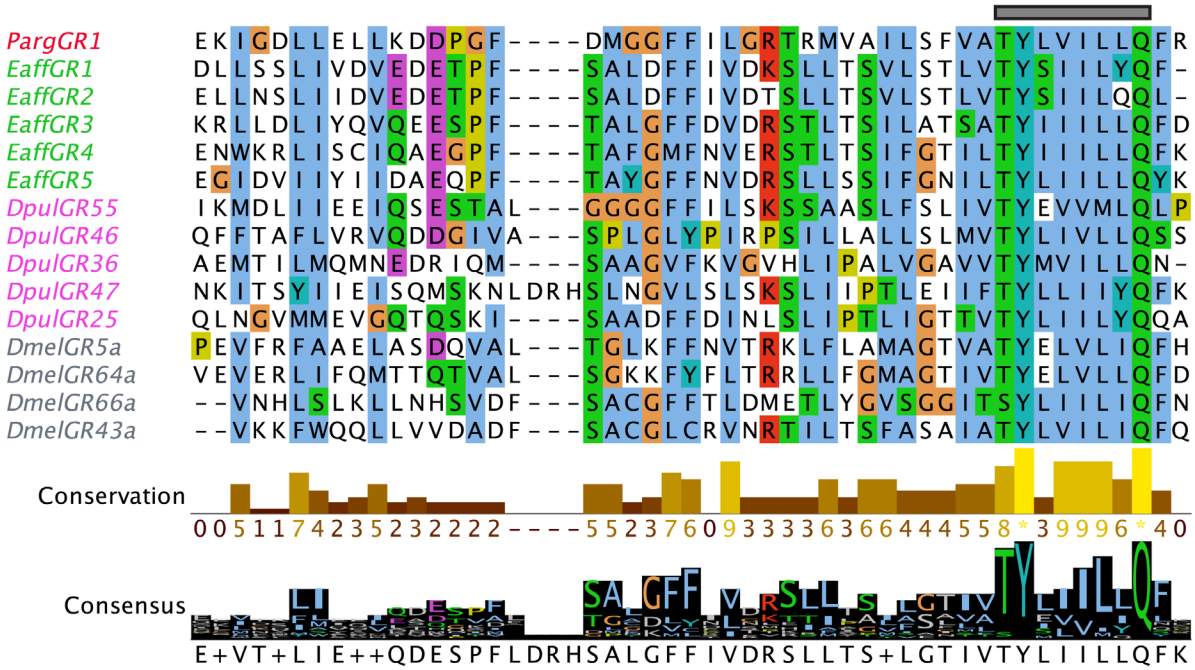

Figure 2.9 PargGR1 fragment sequence alignment.

The multiple sequence alignment of PargGR1 (red) with GRs from arthropods, Eaff (green), Dpul (pink), and Dmel (grey) shows the TM7 region of the 7tm_7 superfamily. Sequences were aligned using MAFFT and visualized on Jalview. Conservation of amino acids across GRs of various species is highest at the 'TYxxxxxQF' motif (grey bar) as shown in the consensus histogram. The residues were colored according to the Clustal X color scheme on Jalview.

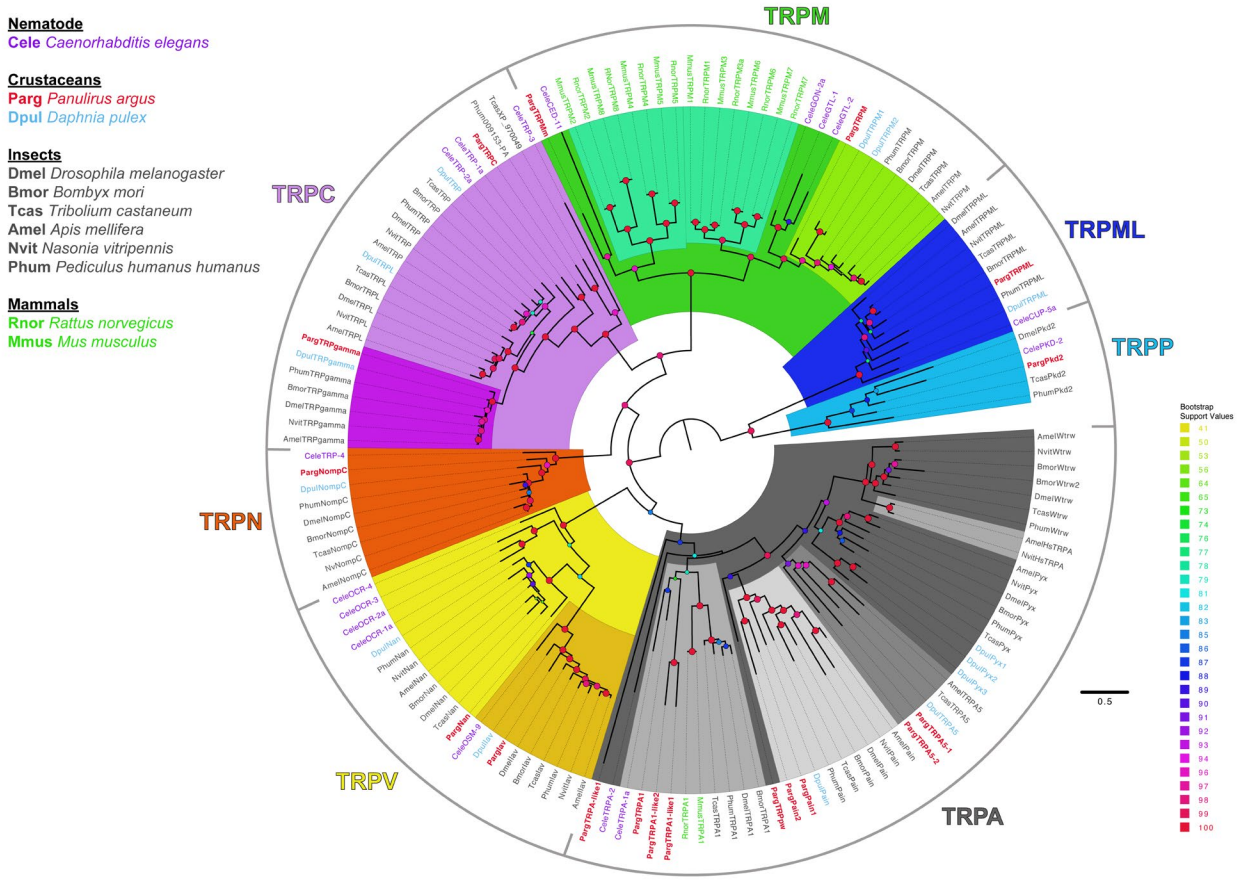

Figure 2.10 Maximum likelihood phylogenetic tree of TRP channels.

Various subfamilies of TRP channels are indicated by different colors where shades of each color indicates a class of TRP channels within a subfamily: TRPA subfamily (grey), TRPM (green), TRPV (yellow), TRPC (pink), TRPN (orange), TRPP (light blue), and TRPML (dark blue). Sequences were aligned with MAFFT and visualized on Jalview. The tree was built on IQ-Tree under LG+G4 model of substitution with 1000 UFBoot replications and visualized on FigTree v1.4.2. Tree was unrooted but drawn with Pkd2 and TRPML clades as the root. The scale bar represents expected number of substitutions per site.

Table 2.1 Number of unique iGluRs and IRs in P. argus*,* D. pulex*, and* D. melanogaster.

Species	iGluR	Co-receptor IR	Conserved IR	Divergent IR
P. argus	6	4	9	95
D. pulex	10	3	1	81
D. melanogaster	14	4	14	48

Only those Parg sequences that were selected for phylogenetic analyses are included here. The four groups of sequences are iGluRs (including NMDA and non-NMDA receptors), co-receptor IRs, conserved IRs, and divergent IRs. The numbers for *D. pulex* and *D. melanogaster* are acquired from sequenced genomes (Croset et al. 2010).

Table 2.2 PCR results on expression of iGluRs and IRs in different tissues of P. argus.

Tissue	Controls		Co-receptor IRs			Tissue-specific IRs	
	GAPDH	NMDAr1	IR25a	IR8a	IR93a	IR1028	IR1074
Dactyl	✓	✓	✓	✓	✓		
Brain		✓	✓	✓	✓		
LFD	✓		✓	✓	✓		
LFP		✓	✓			Not Tested	
A2			✓				
GG			✓	✓	✓		

Green box indicates that a PCR product with the expected size was identified. A green box with a check mark indicates that the PCR product with expected size was identified and when sequenced was found to have > 90% similarity to the expected sequence. Blue box indicates that a PCR product of expected size was identified and was detected only in expected tissue as predicted from the transcriptomes. Blank box indicates the PCR product of expected size was absent in the tissue.

2.7 Supplemental Figures and Table

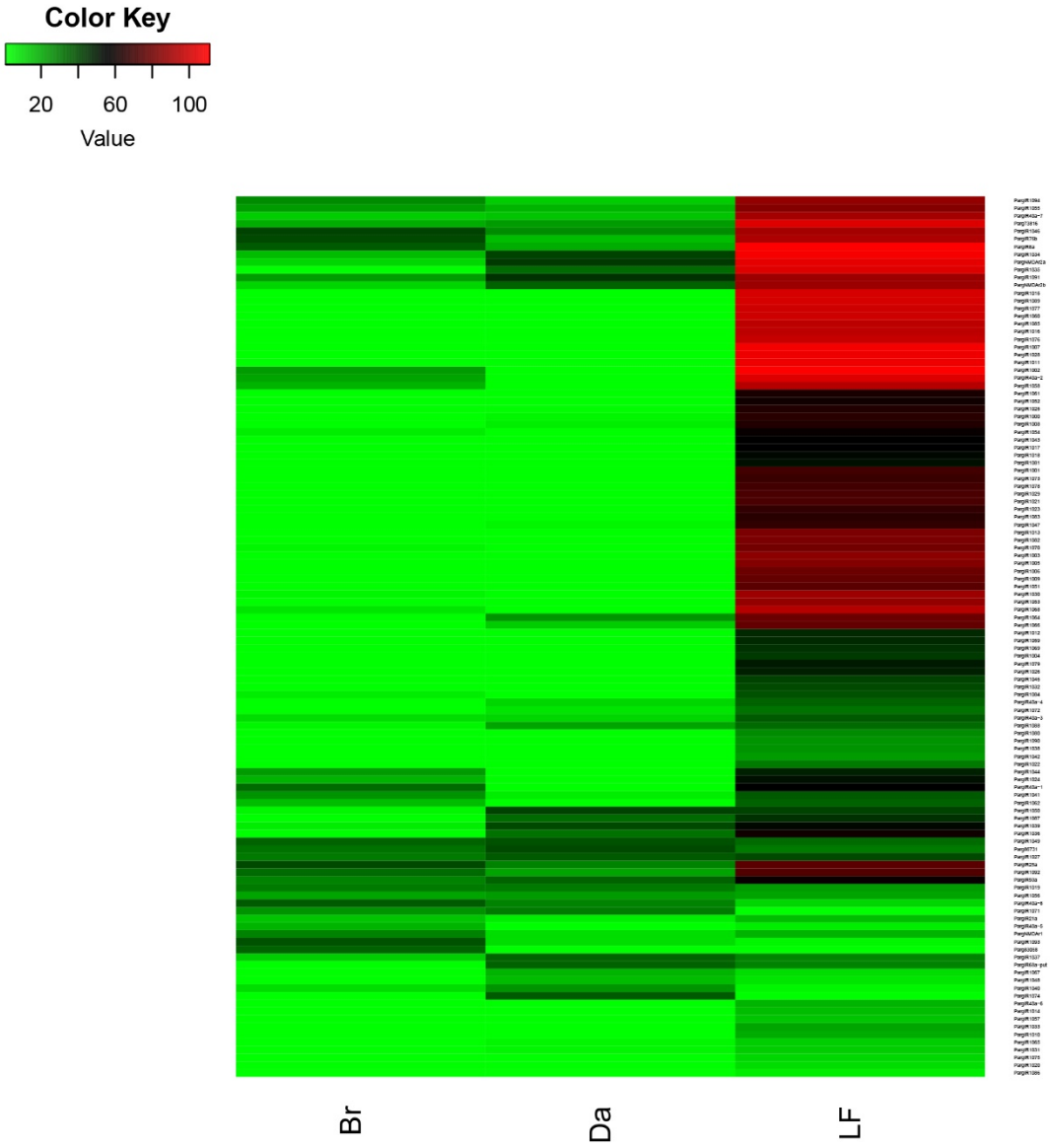

Figure 2.11 Heatmap of abundance of IRs in the LF, Dactyl, and Brain based on not normalized raw counts.

A qualitative representation of raw counts from RSEM.gene.counts.matrix generated by RSEM perl script in Trinity. Plot was created with heatmap.2 function from R gplot package.

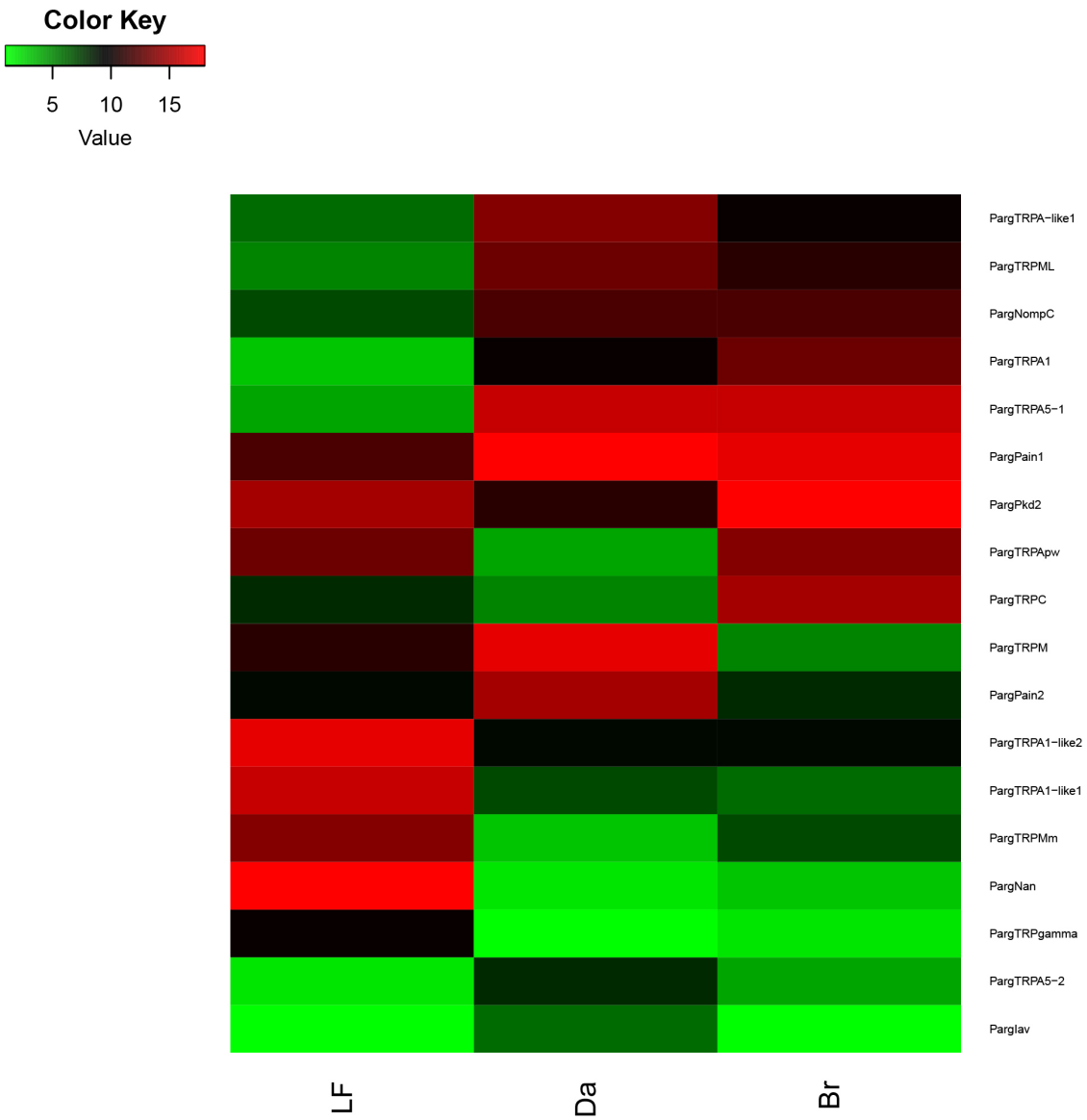

Figure 2.12 Heatmap of abundance of TRP channels in the LF, Dactyl, and Brain based on not normalized raw counts.

A qualitative representation of raw counts from RSEM.gene.counts.matrix generated by RSEM perl script in Trinity. Plot was created with heatmap.2 function from R gplot package.

Table 2.3 Primers.

Target	Primer Sequence
Parg IR8a FWD	TGG TGC CGC AGT TTA TGT
Parg IR8a REV	AGC ATA AGC ACC ACG AAG AG
Parg IR25a FWD	AAT GCT GAT TCC GTC GCT
Parg IR25a REV	CTA CAA GGA TGA CGA CGA GAA G
Parg IR93a FWD	GAC GAC GGG TTT GAG TGT TA
Parg IR93a REV	TCC ATC GTA GAG GTC GTA GTA G
Parg NMDAr1 FWD	ACG CTG GGC TGT GTA CTT A
Parg NMDAr1 REV	CAT CAG AGG CGT TGA CAA G
Parg GAPDH 2 FWD	GAG AAC TTC GAG ATC GTT GAG G
Parg GAPDH 2 REV	CCA TCA ACC TTC TGC ATG TGC T

3 COMPARISON OF TRANSCRIPTOMES FROM TWO CHEMOSENSORY ORGANS IN FOUR DECAPOD CRUSTACEANS REVEALS HUNDREDS OF CANDIDATE CHEMORECEPTOR PROTEINS

Previously published as: Kozma, M. T., H. Ngo-Vu, Y. Y. Wong, N. S. Shukla, S. D. Pawar, A. Senatore, M. Schmidt and C. D. Derby (2020). PLOS ONE **15**(3): e0230266.

Author contribution: Kozma, M.T. designed experiments, collected and analyzed data, and wrote the paper. Ngo-Vu, H collected data and contributed to experimental design and data analysis. Wong, Y.Y., Shukla, N.S., and Pawar, S.D. contributed to data analysis. Senatore, A. contributed to data analysis and editing of manuscript. Schmidt, M. designed experiments, collected and analyzed immunocytochemistry data, and wrote the paper. Derby, C.D. designed experiments and wrote the paper.

Supplemental Information available at: https://doi.org/10.1371/journal.pone.0230266

Acknowledgments: We thank Dr. Timothy McClintock for sharing his anti-HaIR25a antibody, Matthew Rump for phylogenetic analysis of GPCRs, and Drs. Daniel Cox and Paul Katz for many helpful discussions. Funding was provided by a seed grant from the Georgia State University's Brains & Behavior program and fellowships from Georgia State University's 2CI Neurogenomics program, Center for Neuromics, and Neuroscience Institute.

3.1 Introduction

Sensing environmental chemicals is critical for animals because it informs them of the presence and location of important resources. The process of chemoreception begins with an animal acquiring and detecting chemical cues. The detection is performed by chemosensory cells, whose receptor proteins bind stimulus molecules which in turn lead to a cascade of transduction events that results in activation of those cells. Our understanding of these receptor proteins has significantly evolved since the discovery by Buck and Axel (1991) that a major class of chemoreceptor proteins in the rodent olfactory system are modified rhodopsin-like type A G-protein coupled receptors (GPCR), called Odorant Receptors (OR). Since then, many more classes of chemosensory GPCRs have been discovered in mammals and other vertebrates, including trace amine-associated receptors (TAAR), vomeronasal receptor type 1 and 2 (V1R, V2R), formyl-peptide receptors (FPR), and taste receptor type 1 and 2 (T1R, T2R). Vertebrates also have a different class of important though numerically less abundant chemoreceptor proteins – ionotropic receptors – which include transient receptor potential (TRP) channels, epithelial sodium channels (ENaC), and MS4A receptors (Greer et al. 2016)

Unlike the vertebrates, most metazoans including the arthropods, which are the most abundant group of animals, rely more on ionotropic receptors than GPCRs for chemical sensing (Rytz et al. 2013, Benton 2015, Joseph and Carlson 2015, Robertson 2015, Derby et al. 2016, Robertson 2019). In the major and best studied group of arthropods, the insects, ionotropic chemoreceptors include two classes of seven transmembrane ionotropic receptors called Gustatory Receptors (GRs) and Odorant Receptors (ORs), a class of three-transmembrane heterotetrameric receptors called variant Ionotropic Receptors (IRs), TRP channels, and ENaCs. Their structures with the exception of ENaCs

are shown in Figure 3.1. GRs appear to be an ancient lineage, with GR-like receptors (GRLs) being traced possibly even to plants (Benton 2015, Robertson 2015, Saina et al. 2015), though the role of GRLs in chemoreception is unknown (Saina et al. 2015). The ORs likely evolved in insects from GRs, and so far there is no evidence of their existence in non-insect arthropods (Benton 2015, Robertson 2015, Eyun et al. 2017, Brand et al. 2018, Kozma et al. 2018, Vizueta et al. 2018). The IRs evolved from ionotropic glutamate receptors (iGluRs), and they are organized as heterotetramers (Benton 2015, Joseph and Carlson 2015, Robertson 2015) containing two classes of IRs that are different in structure and function: co-receptor IRs, for which there are four (IR25a, IR8a, IR93a, and IR76b); and tuning IRs, for which there are many more (Rytz et al. 2013, van Giesen and Garrity 2017, Sánchez-Alcañiz et al. 2018, Abuin et al. 2019, Robertson 2019). The co-receptors IR25a and IR8a differ from co-receptors IR93a and IR76b as they have high sequence identity to each other and to iGluRs. IR25a and IR8a retain the amino-terminal domain (ATD) of iGluRs that is also largely absent in most tuning IRs and have a distinctive co-receptor extra loop (CREL) region in their ligand binding domain (LBD) that is lacking in all other IRs (Abuin et al. 2019). IR25a and IR8a are also necessary for targeting the IRs to the dendritic membrane to make a functional receptor-channel and may also contribute to the tuning specificity of the cell (van Giesen and Garrity 2017). In contrast, the roles of IR93a and IR76b as co-receptors are still ambiguous, other than being necessary for forming functional receptor channels in several chemosensory neurons. Tuning IRs, on the other hand, only confer specific sensitivity to the IRs heterotetramer by virtue of the particular combination of types. There is evidence in *Drosophila* that the heterotetrameric IR in some olfactory sensory neurons (OSNs) is composed of two co-receptor IRs and two tuning IRs (Abuin et al. 2019).

The IRs have various lineages that have appeared over evolutionary history. The co-receptor IRs differ in their phylogenetic conservation (Abuin et al. 2011, Abuin et al. 2019): IR25a is "protostome conserved," having been identified in all protostomes studied to date but not in deuterostomes or animals ancestral to the protostomes and deuterostomes. IR8a, IR93a, and IR76b are reported only in the arthropods, i.e. "arthropod-conserved" (Eyun et al. 2017, Poynton et al. 2018). The tuning IRs have evolved more recently (Croset et al. 2010, Eyun et al. 2017) and are more specific to phyla or individual species (Abuin et al. 2019). In fact, IRs are not just chemoreceptors but can participate as "environmental sensors," detecting temperature and humidity (Enjin et al. 2016, Knecht et al. 2016, Frank et al. 2017, Knecht et al. 2017).

Regarding other types of chemoreceptor proteins in the arthropods, the ionotropic TRP channels and ENaCs have been shown to be important in insects (Rytz et al. 2013, Benton 2015, Joseph and Carlson 2015, Robertson 2015, Arenas et al. 2017, Ng et al. 2019). GPCRs have not been identified as candidates in arthropods, although pheromone transduction in the hawkmoth *Manduca sexta* may be mediated by a metabotropic signal transduction cascade (Gawalek and Stengl 2018). On the other hand, a large family of chemosensory GPCRs are found in nematodes, and also appear to be in gastropod and cephalopod molluscs and asteroid echinoderms (Cummins et al. 2009a, Albertin et al. 2015, Benton 2015, Eyun et al. 2017, Roberts et al. 2017, Kozma et al. 2018, Roberts et al. 2018).

Crustaceans (Figure 3.2) are more poorly studied regarding chemoreceptor molecules, compared to insects. Chemoreceptor expression has been examined in two chemosensory organs of the Caribbean spiny lobster, *Panulirus argus* (Kozma et al. 2018). One organ is the lateral flagellum of the antennule (LF), which mediates both

olfaction (due to unimodal olfactory sensilla called aesthetascs that are innervated exclusively by OSNs) and distributed chemoreception (due to bimodal sensilla that are innervated by chemosensory neurons (CSNs) and mechanosensory neurons (MSNs)). The other chemosensory organ studied in *P. argus* is the walking leg dactyl, the distal-most segments of the walking legs and which mediate only distributed chemoreception (Kozma et al. 2018). *P. argus* and other crustaceans appear to rely most heavily on IRs for chemoreception, a claim that is based solely on IRs being found in chemosensory organs of all species examined and with the co-receptor IR25a being expressed in most or all OSNs and CSNs (Hollins et al. 2003, Stepanyan et al. 2004, Tadesse et al. 2011, Corey et al. 2013, Groh et al. 2013, Groh-Lunow et al. 2014, Kozma et al. 2018). Besides IRs, GRs are also prevalent in at least some crustaceans, though no ORs are reported (Peñalva-Arana et al. 2009, Robertson 2019). For example, the amphipod *Hyalella azteca* has 155 GRs, the branchiopod *Daphnia pulex* has 59 GRs, and the copepod *Eurytempora affinis* has 67 GRs (Poynton et al. 2018). Other crustaceans have been reported to have very limited GR representation (Eyun et al. 2017, Kozma et al. 2018, Abramova et al. 2019). Homologues of insect chemoreceptive TRP channels have also been identified in crustaceans (Kozma et al. 2018). Many IRs, one GRL, and representatives from each of the major types of TRP channels were found in *P. argus* (Kozma et al. 2018).

The goal of the current study is to extend our analysis of chemoreceptor proteins of decapod crustaceans by studying three additional decapod species commonly used in studies of chemoreception: American lobster *Homarus americanus*, crayfish *Procambarus clarkii*, and blue crab *Callinectes sapidus*. These decapod species were chosen for their phylogeny (Figure 3.2), habitat, and lifestyle. *H. americanus* is similar to *P. argus* with its "lobster"-like body form, long life span, and life in a complex marine

environment, but it is phylogenetically distant from *P. argus*. Its use of chemoreception in feeding, social behavior, and sexual behavior is well described (Atema and Voigt 1995, Atema 2018). *P. clarkii* is phylogenetically close to *H. americanus*, but unlike the lobster it is a freshwater crustacean with a shorter lifespan. Its chemical senses and chemical sensing are also well described (Hatt 1984, Derby and Blaustein 1988, Hazlett 1990, Breithaupt 2001, Sandeman and Sandeman 2003, Moore and Bergman 2005, Fedotov 2009, Mellon 2012). *C. sapidus* is a brachyuran crab, phylogenetically most dissimilar from the other three decapods, with a euryhaline life that allows it to live in both marine and essentially freshwater environments. Its chemoreception has been studied with respect to feeding (Keller and Weissburg 2004, Aggio et al. 2012, Poulin et al. 2018), and sexual behavior and pheromones (Gleeson 1991, Kamio et al. 2008, Kamio et al. 2014). Our goal was to describe the diversity and phylogenic relationship of the chemoreceptor proteins of these decapods relative to other arthropods. We discovered a broad expansion of variant IRs, TRP channels from all subfamilies, and a few GRLs in the chemosensory organs of these four decapod crustaceans.

3.2 Materials and Methods

3.2.1 Animals

Male and female Caribbean spiny lobsters, *Panulirus argus*, American lobsters, *Homarus americanus*, red swamp crayfish, *Procambarus clarkii*, and blue crabs, *Callinectes sapidus* were used. Specimens of spiny lobsters were collected in the Florida Keys and kindly provided by the Florida Keys Marine Laboratory and Dr. Donald Behringer (University of Florida). Specimens of *H. americanus*, *P. clarkii*, and *C. sapidus*

were obtained from H Mart (Doraville, GA), with *H. americanus* having been collected in New England, *P. clarkii* in Louisiana, and *C. sapidus* in the U.S. southeast coast. Animals were held at Georgia State University in communal 800-L aquaria or in individual 10-L aquaria containing aerated, recirculated, filtered artificial seawater (Instant Ocean, Aquarium Systems, Mentor, OH) for the marine species and dechlorinated tap water for *P. clarkii* in a 12-hr:12-hr light:dark cycle. They were fed shrimp or squid three times per week.

3.2.2 Tissue collection and RNA isolation for generating transcriptomes

Animals was anesthetized on ice prior to tissue collection. For *P. argus*, tissues were collected from four adult animals (three females and one male), as described in Kozma et al. (2018). For *H. americanus*, tissues were collected from an adult male (weight 484 g, carapace length 114 mm) and an adult female (weight 463 g, carapace length 109 mm). For *P. clarkii*, nine animals were used (five adult females and four adult males, mean weight 37.1 g and mean carapace length 36.5 mm), with five animals being used for extracting RNA from antennules, two for dactyls (one from each sex), and two for brains (one from each sex). For *C. sapidus*, ten animals (eight adult males and two adult females, mean weight 161.5 g, mean carapace length 143 mm) were used for extracting RNA from antennules, and one male and one female from the same group of animals were used for dactyl tissue extraction.

Three tissues were collected: aesthetasc-bearing region of both antennular lateral flagella (LF); sensilla-bearing dactyl of the second walking legs (dactyl); and supraesophageal ganglion (brain) (only in *P. argus*, *H. americanus*, and *P. clarkii*).

In *P. argus* and *H. americanus*, the soft tissue within the cuticle of LF and dactyls was dissected out from each animal and pooled together according to tissue type for each species. In *P. clarkii* and *C. sapidus*, the soft tissue of the dactyl was dissected out of the cuticle, while the LFs were collected with the tissue still within the cuticle due to their small size. Cuticle was removed by homogenizing tissue, centrifugation, and filtration of supernatant prior to RNA extraction. Collected tissues were frozen instantly in liquid nitrogen and stored at -80° C until RNA extraction (Kozma et al. 2018).

To extract total RNA, the same methods as described in Kozma et al. (2018) were used. Frozen tissues were homogenized in Tri-Reagent (Sigma-Aldrich, St. Louis, Missouri). Sequential centrifugation with chloroform and ethanol was used to precipitate RNA from the tissue. DNA, protein, and carbohydrate contaminants were removed by reconstituting RNA in diethylpyrocarbonate (DEPC)-treated water and then precipitating again in lithium chloride. Potassium acetate was used to precipitate other possible contaminants such as sodium dodecyl sulfate (SDS) and SDS-bound proteins and leave RNA in solution, which was then precipitated out of solution using ethanol. Total RNA reconstituted in DEPC-treated water was tested for concentration and purity using NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts). Total RNA extracted for each tissue type of each species was frozen over liquid nitrogen and stored in aliquots at –80°C.

3.2.3 RNA sequencing, de novo *assembly, and transcript abundance estimation*

Quality assessment on Agilent Bioanalyzer2000 and TapeStation of total RNA extracted, mRNA specific cDNA synthesis, and cDNA paired-end sequencing on the

Illumina HiSeq 2500 high-throughput sequencer were performed by Beckman Coulter Genomics (now part of GENEWIZ, South Plainfield, New Jersey) similar to Kozma et al. (2018). For *P. argus*, the read length was 2x100 (base pair reads) for LF and dactyl, and 2x125 (base pair reads) for brain. For *H. americanus*, the read length was 2x125 (base pair reads) for all three tissues. For *P. clarkii*, the read length was 2x125 (base pair reads) for LF and dactyl and 2x100 (base pair reads) for brain. For *C. sapidus*, the read length was 2x125 (base pair reads) for both tissues. The number of reads per sample was > 120 million. Adapter sequences tracking Illumina reads from multiplexed samples were removed prior to delivery. The reads were deposited to NCBI under BioProject accession PRJNA596786, with SRA accessions SRR10874089, SRR10874088, SRR10874086, SRR10874085, SRR10874084, SRR10874083, SRR10874082, SRR10874081, SRR10874080, SRR10874079, and SRR10874087. Raw reads for each species were concatenated prior to transcriptome assembly. For example, for *P. argus*, left reads from LF, dactyl, and brain were concatenated into one file, and right reads were concatenated into one file. Eight independent *de novo* assemblies were first generated for each species using different transcriptome assembly software, as described below. A reference transcriptome for each species was then generated by using the EvidentialGene pipeline (https://f1000research.com/posters/5-1695).

3.2.3.1 Trinity de novo *assemblies*

Unnormalized Trinity *de novo* Assembly: The first *de novo* transcriptome was assembled via Trinity v.4.0 (Grabherr et al. 2011). Raw reads were compiled into left and right read files respectively, and processed through Trinity-Trimmomatic v.4.0 (Grabherr et al.

2011, Bolger et al. 2014) for trimming of 3'-ends of the sequenced reads. This *de novo* transcriptome was generated without normalization of reads.

Normalized Trinity *de novo* Assembly: The second transcriptome was then assembled with Trinity v.4.0 (Grabherr et al. 2011). Raw reads were compiled into left and right read files respectively, and processed through Trinity-Trimmomatic v.4.0 (Grabherr et al. 2011, Bolger et al. 2014) for trimming of 3'-ends of the sequenced reads. This *de novo* transcriptome was generated with the default Trinity v2.4.0 normalization of reads.

3.2.3.2 Other de novo *assemblies*

Normalization: The remaining six transcriptomes were all generated using normalized reads. The trimmed raw reads (processed through Trinity-Trimmomatic v.4.0, obtained from the trimming process of the first transcriptome assembly mentioned above) were normalized with FastUniq (Xu et al. 2012) to remove redundancy in reads data.

Normalized TransAbyss *de novo* Assemblies: The third, fourth, and fifth transcriptomes were generated using Trans-Abyss v1.5.3 (Robertson et al. 2010), with K-mer sizes 63, 87, and 111, respectively.

Normalized Velvet/OASES *de novo* Assemblies: The sixth, seventh, and eighth transcriptomes were generated using Velvet v1.2.10 and OASES v0.2.09, with K-mer-sizes 63, 87, and 111, respectively.

3.2.3.3 EvidentialGene pipeline

All eight transcriptome assemblies for each species were input to EvidentialGene (EVG) pipeline to give a single refined transcriptome for each species with transcript and protein-coding gene counts shown in Supplemental Table 3.3.

The TransDecoder program (http://transdecoder.github.io/) was used to predict protein sequences from transcripts based on open reading frames (ORF). CD-Hit (Li et al. 2001) was performed on each transcriptome to remove redundancy. All further analyses were performed on the cd90 datasets generated by CD-Hit. BUSCO v3 was run on the transcriptomes to analyze the completeness of the assembly using the *Arthropoda odb9* lineage (Creation date: 2017-02-07, number of species: 60, number of BUSCOs: 1066) (Simao et al. 2015, Waterhouse et al. 2017, Waterhouse et al. 2019). The BUSCO output for *P. argus* was **C**:91.7% [**S**:89.1%, **D**:2.6%], **F**:1.0%, **M**:7.3%, **n**:1066; for *H. americanus*, **C**:91.5% [**S**:89.5%, **D**:2.0%], **F**:1.0%, **M**:7.5%, **n**:1066; for *P. clarkii*, **C**:92.8% [**S**:90.0%, **D**:2.8%], **F**:0.9%, **M**:6.3%, **n**:1066; for *C. sapidus*, **C**:96.2% [**S**:85.8%, **D**:10.4%], **F**:1.3%, **M**:2.5%, **n**:1066; where **C**=complete BUSCOs, **S**=complete and single-copy BUSCOs, **D**=complete and duplicated BUSCOs, **F**=fragmented BUSCOs, **M**=missing BUSCOs, and **n**=total BUSCO groups searched (Supplemental Table 3.4).

Following the removal of redundancy, the abundance of transcripts for each transcriptome was estimated using RSEM (Li and Dewey 2011) bundled into the Trinity v2.8.2 package for each tissue type and a counts matrix was generated. Custom 'R' script (Supplemental S1 File: https://doi.org/10.1371/journal.pone.0230266.s014) for DESeq2 (Love et al. 2014) was used to measure fold differences of transcript abundance within each transcriptome to predict tissue specificity for a given transcript of interest. Transcripts whose expression was $\log_2$[fold change] ≥ 1.5 or $\log_2$[fold change] ≤ -1.5,

according to DESeq2 (~ 2.8 fold actual change) were considered to have a higher level of expression in one tissue compared to the other (i.e. LF vs. dactyl). Transcripts whose fold change was in between this range were considered to be expressed at the same level in both tissues.

3.2.4 IR identification, sequence alignment, and phylogenetic analysis

Screening for IRs and iGluRs was performed with TMHMM v2.0 for transmembrane domain prediction and domain region screened with InterProScan 5 (v5.28-67.0) (Jones et al. 2014) for conserved Pfam (Finn et al. 2016) and InterPro (Finn et al. 2017) domains on high performance computing systems at Georgia State University (Sarajlic et al. 2016, Sarajlic et al. 2017). IRs have several distinctive domains: an extracellular amino-terminal domain (ATD) involved in assembly of the heteromeric channel; an extracellular ligand binding domain (LBD) consisting of two half-domains (S1 and S2) to which agonists bind; an ion channel domain (ICD) that forms the ion channel, consisting of three transmembrane domains (M1, M2, M3) and a pore loop (P); and an intracellular carboxyl-termination domain (CTD). The ICD domain and S2 of the LBD were predicted by the presence of the Pfam domain PF00060 (which contains M1, P, M2, S2, and M3). S1 of the LBD was predicted by the presence of the Pfam domain PF10613. All predicted protein sequences from the transcriptomes that contained both PF00060 and PF10613 domain regions were considered putative IRs and selected for phylogenetic analyses. IRs and iGluRs from *P. argus* that were previously identified (Kozma et al. 2018) were used as reference sequences. In *P. argus*, following transcriptome assembly with the EvidentialGene pipeline, we identified more complete putative IR sequences than predicted previously (Kozma et al. 2018). Phylogenetic trees were built as described in

Kozma et al. (2018). Selected sequences were aligned using default settings for MAFFT (Katoh et al. 2002, Katoh and Standley 2013). Alignments were visualized and trimmed on Jalview (Waterhouse et al. 2009) to remove gaps and regions of aligned sequences with low amino acid conservation. Sequences that had large gaps in the LBD and ICD regions were removed, with some exceptions. IQ-Tree (Nguyen et al. 2015, Trifinopoulos et al. 2016) was used for constructing maximum likelihood phylogenetic trees, along with ModelFinder (Kalyaanamoorthy et al. 2017) integrated into IQ-Tree to automatically determine the best model of substitution (see selected models in figure legends). Ultrafast bootstrap (UFBoot) (Minh et al. 2013) integrated into IQ-Tree was used to generate confidence values for the trees. The phylogenetic trees were visualized using FigTree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/), and color schemes were edited on Adobe Illustrator CS6, San Jose, CA.

Phylogenetic analysis of conserved IRs also included conserved IR sequences from several arthropod species (Supplemental S12 Table: https://doi.org/10.1371/journal.pone.0230266.s013). These include chelicerates [horseshoe crab *Limulus polyphemus* (Lpol), mites [*Galendromus occidentalis* (Gocc), *Tetranychus urticae* (Turt), *Leptotrombidium deliense* (Ldel), *Dinothrombium tinctorium* (Dtin), and *Tropilaelaps mercedesae* (Tmer)], tick *Ixodes scapularis* (Isca), scorpion *Centruroides sculpturatus* (Cscu), spider *Parasteatoda tepidariorum* (Ptep)]; myriapod [*Strigamia maritima* (Smar)]; crustaceans [thecostracan barnacle *Amphibalanus improvisus* (Aimp), copepod *Eurytemora affinis* (Eaff), isopod *Armadillidium vulgare* (Avul), amphipod *Hyalella azteca* (Hazt), decapod *Litopenaeus vannamei* (Lvan), branchiopod *Daphnia pulex* (Dpul)]; insects [*Drosophila melanogaster* (Dmel), *Aedes aegypti* (Aaeg), *Culex quinquefasciatus* (Cqui), *Anopheles*

gambiae (Agam), *Bombyx mori* (Bmor), *Tribolium castaneum* (Tcas), *Apis mellifera* (Amel), *Nasonia vitripennis* (Nvit), *Acrythosiphon pisum* (Apis), and *Pediculus humanus humanus* (Phum)]; and two gastropod molluscs [*Aplysia californica* (Acal) and *Lottia gigantea* (Lgig)].

3.2.5 IR and iGluR nomenclature

IRs and iGluRs were named as previously described in Kozma et al. (2018). Sequences from each species were given one of the following prefixes: Parg (*P. argus*), Hame (*H. americanus*), Pcla (*P. clarkii*), or Csap (*C. sapidus*). Newly identified Parg IR sequences from the EvidentialGene pipeline were arbitrarily assigned numbers increasing from 1095 following the nomenclature in Kozma et al. (2018), e.g. PargIR1095, PargIR1096, and so on. Some newly identified sequences were named by assigning suffixes "b", "c," and on so to previously assigned sequence numbers in order to maintain continuity of numbers within a cluster of closely related IRs. Sequences were given numbers increasing from 2000 for *H. americanus*, 3000 for *P. clarkii*, and 4000 for *C. sapidus*. Sequences from Hame, Pcla, and Csap with homologues to Parg sequences were given the same number as the Parg sequence. Sequences from Pcla and Csap with homologues to only Hame sequences were given the same number as the Hame sequence. Sequences from Csap with homologues to only Pcla sequences were given the same number as the Pcla sequence. Whenever there were multiple homologues to a sequence in one species, suffixes "a", "b," and on so on were attached to each homologue. Therefore, only conserved IRs across the four decapod species share the same numbers. NMDA iGluRs were named according to their homologues in other species. Non-NMDA iGluRs were assigned arbitrary numbers, where the same number across species indicates

homologues (Supplemental S12 Table: https://doi.org/10.1371/journal.pone.0230266.s013).

3.2.6 GRL identification and sequence alignment

The transcriptomes were screened using InterProScan for the Pfam domain family, 7tm_7 (PF08395), since this family includes GRs and ORs found in insects. Randomly selected GRs from *Drosophila melanogaster*, *Daphnia pulex*, and *Eurytemora affinis* were used as reference sequences for multiple sequence alignment using MAFFT and visualized on Jalview. GRL numbers for each species do not indicate that they are homologues across other species (e.g. PargGR1 and HameGR1 are not homologous).

3.2.7 TRP channels identification

Similar to putative IR and GR sequences, TRP channels were identified using InterProScan and screening for Pfam domain regions that are typically found in the different subfamilies of TRP channels: PF06011, PF08344, PF00520, PF12796, PF00023, PF16519, and PF08016. Multiple sequence alignments were constructed using MAFFT. Reference TRP channel sequences included in the alignments were from *D. melanogaster*, *B. mori*, *T. castaneum*, *A. mellifera*, *N. vitripennis*, *P. humanus humanus*, *D. pulex*, *Caenorhabditis elegans*, *Rattus norvegicus*, and *Mus musculus*. Maximum likelihood trees were constructed using IQ-Tree with confidence values generated by 1000 bootstraps using UFBoot, where the model of substitution was predicted by ModelFinder and shown in the figure legend. The trees were visualized on FigTree v.1.4, and color schemes were edited on Adobe Illustrator.

3.2.8 Immunocytochemistry

LF and dactyls of second and third pereiopods of male and female *P. argus* (carapace length 34–65 mm, weight 45–250 g, n = 6) and of male and female *H. americanus* (carapace length 109–122 mm, weight 458–569 g, n = 4) were dissected after anesthetizing the animals on ice for about 20 min. LF were cut into 8-annuli long pieces as described previously (Schmidt et al. 2006), and dactyls were cut into 2 or 3 pieces. Tissue was fixed for 6–24 hr at room temperature in 4% paraformaldehyde in 0.1 M Sörensen phosphate buffer (SPB) containing 15% sucrose. Tissue was then decalcified by incubation in 10% EDTA in SPB for about one week (pieces of LF) or 2 weeks (pieces of dactyls) with several changes of the medium and then stored in 0.02 M SPB with 0.02% sodium azide at 4° C. For sectioning, tissues were embedded in 300-bloom gelatin with some modifications to method described in detail previously (Schmidt et al. 2006). According to this procedure, tissue pieces were first incubated for at least 1 hr in warm (60° C) gelatin to facilitate penetration of the entire internal tissue (according to Long (2018)), then pieces were embedded in warm gelatin in a small disposable paraffin mold, the gelatin was hardened by cooling it on ice, and finally the gelatin was hardened by fixing it with 4% paraformaldehyde at 4° C overnight. Tissue pieces in hardened gelatin were cut on a vibrating microtome (VT 1000 S; Leica, Wetzlar, Germany) into 80–100 µm thick sagittal sections.

Free-floating sections were incubated overnight at room temperature with an affinity-purified polyclonal rabbit antiserum against IR25a of *H. americanus* (anti-HaIR25a - courtesy of Dr. Timothy McClintock, University of Kentucky) diluted 1:750 in SPB containing 0.3% Triton-X-100 (TSPB). Anti-HaIR25a (previously annotated as anti-GluR1) was generated using two non-overlapping peptides ($P1_{Ha}$: TGEGFDIAPVANPW;

$P2_{Ha}$: REYPTNDVDKTNFN) from the C-terminus of *H. americanus* IR25a (originally annotated as OET-07; Genbank accession #AY098942) (Hollins et al. 2003, Stepanyan et al. 2004). Sequence alignments of P1 and P2 with the deduced amino acid sequences of all IRs and iGluRs of *P. argus* identified in our transcriptome sequencing project showed close matches for both peptides only in *P. argus* IR25a, with $P1_{Ha}$ at 79% identity and $P2_{Ha}$ at 86% identity. For *P. clarkii*, $P1_{Ha}$=57% identity and $P2_{Ha}$=43% identity, and for *C. sapidus*, $P1_{Ha}$=71% identity and $P2_{Ha}$=36% identity. Consequently, the anti-HaIR25a yielded results for *H. americanus* and *P. argus* but not *P. clarkii* or *C. sapidus*.

Anti-HaIR25a was combined with a mixture of two mouse monoclonal antibodies against modified α-tubulin isoforms that are enriched in neurons (Fukushima et al. 2009) to achieve labeling of all sensory neurons. These tubulin antibodies were anti-tyrosine tubulin (T9028, clone TUB-1A2, Sigma-Aldrich, St. Louis, Missouri) diluted 1:2000 and anti-acetylated tubulin (sc-23950, Santa Cruz Biotechnology, Dallas, Texas) diluted 1:200. Our previous immunocytochemical studies on *P. argus* (Kozma et al. 2018) showed that this mixture of anti-tubulin labeled all bipolar sensory neurons in the LF, dactyl, and second antennae, and that while epithelial cells and the walls of hemolymph vessels could also be labeled, these cells are easily distinguishable from sensory neurons based on location and morphology.

After incubation in primary antibodies, sections were rinsed 4 x 30 min in TSPB and then incubated in a mixture of two secondary antibodies, goat anti-rabbit CY3 (111-165-003, Jackson ImmunoResearch, West Grove, Pennsylvania) diluted 1:400 and goat anti-mouse DyLight-488 (35502, Thermo Fisher Scientific, Waltham, Massachusetts) diluted 1:100 in TSPB. After rinsing 3 x 30 min in TSPB, sections were incubated for 20 min in Hoechst 33258 diluted 1:150 in TSPB from a stock solution of 1 mg/ml to stain nuclei.

After a final rinse in SPB, sections were mounted on slides in 1:1 glycerol:SPB containing 5% DABCO (diazabicyclol[2.2.2]octane) to prevent photobleaching. Coverslips were secured with nail polish, and slides were stored at 4° C or at −20° C for extended storage time.

Labeled sections were viewed and imaged at low magnification in an epifluorescence microscope equipped with color CCD camera (AxioScope FL LED with Axiocam 503, Carl Zeiss Microscopy, Thornwood, New York) and imaged at higher magnification in a confocal microscope (LSM 700, Carl Zeiss Microscopy) using the associated software package ZEN. Stacks of optical sections each 0.3–1.0 µm thickness covering from several µm to the entire section thickness of 80 µm were collected. LSM Image Browser software (version 4.2.0.121, Carl Zeiss MicroImaging GmbH, Jena, Germany) was used to select sub-stacks of optical sections (0.5–3.0 µm total thickness) and collapse them to two-dimensional images using maximum-intensity projection. PaintShopPro 6 (Jasc Software, Eden Prairie, Minnesota) was used to optimize brightness and contrast of the images and to filter out pixel noise. Final image plates were assembled in Adobe Illustrator CS6 (Adobe Systems Inc., San Jose, California).

To scrutinize the specificity of anti-HaIR25a to label IR25a of *P. argus*, we did a pre-absorption control using the corresponding *P. argus* IR25a peptides ($P1_{Pa}$: GGDGYDIAPVANPW; $P2_{Pa}$: REYPTNDVDKSNFT). We incubated 20 µl of anti-HaIR25a with 1 mg of each peptide in 800 µl SPB at 4°C overnight (according to Stepanyan et al. 2004 (Stepanyan et al. 2004)) and in parallel prepared a control antibody (20 µl anti-HaIR25a in 800 µl SPB at 4°C overnight). Then we used the pre-absorbed and control antibodies at 1:750 final dilution in TSPB to label alternating 50 µm-thick vibratome sections through an 8-annuli long section of distal (aesthetasc-bearing) LF of *P. argus*.

Confocal images of sections labeled with pre-absorbed anti-HaIR25a show a lack of specific labeling in all parts of OSNs (outer dendritic segments: oDS, inner dendritic segments: iDS, and somata), while adjacent sections labeled with control anti-HaIR25a (and collected at the same intensity setting of the confocal channel for CY3) show specific labeling in all parts of OSNs (Figure 3.3). As reported previously (Kozma et al. 2018), oDS were labeled with highest intensity, followed by iDS, and then OSN somata. The result of the pre-absorption control confirms that anti-HaIR25a binds to and therefore labels *P. argus* IR25a. In general, anti-HaIr25a has some limits in its ability to penetrate tissue, so it is most effective in labeling cells on the surface of the tissue.

3.3 Results

We previously identified IRs, GRLs, and TRP channels expressed in two chemosensory organs – LF and dactyl – of the Caribbean spiny lobster, *P. argus*, based on sequence homology to receptors in other species (Kozma et al. 2018). In this paper, we expanded our analysis to include three additional species: *H. americanus*, *P. clarkii*, and *C. sapidus*. We used the EvidentialGene pipeline to generate a single refined *de novo* transcriptome for each species, using reads generated from RNA-Seq of two chemosensory organs (LF and dactyl) for each species and brain for three species (*P. argus*, *H. americanus*, and *P. clarkii*). Furthermore, we estimated the abundance of transcripts for each transcriptome using RSEM and report the fold differences of abundance (LF vs. dactyl) to predict tissue specificity for transcripts of interest. Our findings are described in the following sections according to receptor type.

3.3.1 IRs

The total numbers of sequences having the PF00060 domain (consisting of M1, P, M2, S2, and M3 region: see Figure 3.1), PF10613 domain (consisting of the S1 region), or both domains of iGluR and IR for the transcriptomes (generated from LF, dactyl, and brain) for each of the four species are shown in Table 3.1. All sequences containing PF00060 and PF10613, respectively, are in Supplemental S2-S9 Files (https://doi.org/10.1371/journal.pone.0230266). Using only sequences with both domains generates a conservative estimate of the number of iGluRs and IRs, and these values range from 96 for *P. clarkii* to 252 for *P. argus*. The estimated numbers of IRs are much higher when based on sequences having only one of the two domains.

We performed phylogenetic analyses of sequences with both domains of iGluRs and IRs (referred as 'selected sequences') from the four decapod crustaceans (Figure 3.4; Supplemental S10-S12 Files: https://doi.org/10.1371/journal.pone.0230266) and other arthropods (Figure 3.5; Supplemental S13-S15 Files: https://doi.org/10.1371/journal.pone.0230266) to better determine iGluR and IR homologues. Table 3.2 summarizes these findings.

The phylogenetic tree of IRs and iGluRs in decapods is shown in Figure 3.4 and Supplemental Figure 3.1 in polar tree and radial tree configurations, respectively. IRs and iGluRs in decapods are distributed into nine broad clades (Supplemental Figure 3.10). One clade has iGluRs, co-receptor IRs, IR75-family (IR1091–IR1095 and IR1034), IR1035, and IR1036. The IR40a family forms its own clade. The other seven clades are all tuning IRs only. Most of these clades have conserved tuning IRs, which are described below.

3.3.1.1 iGluRs

The four decapod species expressed a number of iGluRs in their transcriptomes, ranging from six in *C. sapidus* to eleven in *H. americanus*. These iGluRs included NMDA receptor homologues NMDAr1, NMDAr2, and NMDAr3, and non-NMDA receptor homologues (Table 3.2). The identity and distribution of these iGluRs are shown in Supplemental Table 3.5, and their phylogenetic relationships are shown in Figure 3.4.

3.3.1.2 Co-receptor IRs

The four co-receptor IRs – IR25a, IR8a, IR93a, and IR76b – were found in both chemosensory organs of all four decapod crustacean species (Figure 3.4, Table 3.2). According to RSEM analysis (Supplemental S4-S7 Tables: https://doi.org/10.1371/journal.pone.0230266), IR25a is the most abundantly expressed IR in all four species.

DESeq2 analysis was used to identify quantitative differences in expression of these four co-receptor IR transcripts between tissue types for a given species. A difference in expression of transcripts between tissue types that is greater than ~ 2.8 fold (i.e., $\log_2$ [fold change] that is ≥ 1.5 or ≤ -1.5) is considered to be higher expression in a particular tissue type. The results of this analysis (Supplemental S8-S11 Tables: https://doi.org/10.1371/journal.pone.0230266) show that IR25a is more highly expressed in LF than dactyl in all four species: in *P. argus*, 16 fold difference; in *H. americanus*, 356 fold difference; in *P. clarkii*, 121 fold difference, and in *C. sapidus*, 103 fold difference.

Like IR25a, IR93a is more abundantly expressed in LF than dactyl in all four species: in *P. argus*, 385 fold difference; in *H. americanus*, 1323 fold difference; in *P. clarkii*, 154 fold difference; and in *C.* sapidus, 4211 fold difference.

IR8a is more highly expressed in LF than dactyl in two of the species – *P. clarkii* and *C. sapidus*: in *P. clarkii*, 9.3 fold difference; and in *C. sapidus*, 5.8 fold difference. But in *P. argus* and *H. americanus*, there is no difference in expression in LF and dactyl.

IR76b has a more varied expression pattern between the tissue types in each species. In *P. clarkii*, IR76b has greater expression in LF than in dactyl by 34 fold. In *H. americanus*, there is no difference in expression of IR76b between the tissues. In *P. argus* and *C. sapidus*, IR76b has greater expression in dactyl than LF: 23 fold in *P. argus* and 3.0 fold in *C. sapidus*.

3.3.1.3 Tuning IRs

Each species expresses many tuning IRs in the LF and dactyl: *P. argus* has 254 tuning IRs, *H. americanus* has 181, *P. clarkii* has 92, and *C. sapidus* has 186 (Table 3.2). It is important to note that these conservative estimates are based only on the number of sequences found in the transcriptome of each species that had both domain regions (PF00060 and PF10613) that define a variant IR. Some incomplete sequences were also included in this analysis due to homology to IRs with both domain regions from another species. If all the sequences that have only one of these domain regions are also taken into account, then the number of tuning IRs in each species almost doubles or more (Table 3.1).

Under the assumption that genes have not been missed in our transcriptome sequencing and assembly, some tuning IRs appear to be conserved phylogenetically,

while some are unique to one of the four decapod species (Figure 3.4 and Figure 3.5, Table 3.2). Therefore, there exist sub-classes of 'conserved tuning IRs.' Seventeen IRs have homologues in all four species; these are IR1001, IR1018, IR1020, IR1021, IR1029, IR1033, IR1037, IR1038, IR1039, IR1044, IR1046, IR1053, IR1057, IR1064, IR1065, IR1097, and IR1155. Out of these seventeen IRs, five IRs also have homologues in the decapod crustacean, *Litopenaeus vannamei* (Pacific white shrimp), three also have homologues in the amphipod, *Hyalella azteca*, one (IR1067) has homologues in both *L. vannamei* and *H. azteca*, and one (IR1020) has homologues in *L. vannamei*, *H. azteca*, and *Armadillidium vulgare* (Figure 3.5). Thus, we consider eleven IRs as "decapod-conserved" tuning IRs (IR1001, IR1018, IR1021, IR1029, IR1037, IR1044, IR1046, IR1053, IR1057, IR1097, IR1155) and five IRs as "crustacean-conserved" tuning IRs (IR1020, IR1038, IR1064, IR1066, IR1067) (Figure 3.5). Another IR, IR1069, detected in three species (*P. argus*, *P. clarkii*, and *C. sapidus*) also has a homologous sequence in *L. vannamei* and is therefore considered "decapod-conserved." Four tuning IRs families that are conserved in insects – IR21a, IR40a, IR68a, and IR75 – were previously identified in *P. argus* (Kozma et al. 2018). While *P. clarkii*, *H. americanus*, and *C. sapidus* have homologues to IR21a, IR40a, and IR75-family, only *H. americanus* and *C. sapidus* also have homologues to IR68a (Figure 3.4 and Figure 3.5). Similar to *P. argus* (Kozma et al. 2018), *P. clarkii*, *H. americanus*, and *C. sapidus* have an expanded family of IR40a homologues (Figure 3.4 and Figure 3.5). In all four decapods, IR21a is more abundantly expressed in LF compared to dactyl, and all IR40a homologues are more abundantly expressed in LF compared to dactyl with three exceptions: IR40a-3 in *P. argus* is more highly expressed in dactyl compared to LF; IR40a-8 in *P. argus* and IR40a-e in *H. americanus* have similar expression in both LF and dactyl (Supplemental S4-S11

Tables: https://doi.org/10.1371/journal.pone.0230266). IR68a homologues are more abundantly expressed in LF compared to dactyl in all three species that express it (*P. argus*, *H. americanus*, and *C. sapidus*). We consider sequences that are numbered IR1091–IR1095 and IR1034 in the decapod crustaceans to be homologous to the insect IR75-family of genes. For the IR75-family, in *P. argus*, IR1091–IR1093 and IR1034 have higher expression in dactyl compared to LF, and IR1094 and IR1095 have similar expression in LF and dactyl. In *H. americanus*, IR1034, IR1092, and IR1095 have higher expression in LF compared to dactyl, and IR1034b, IR1091, and IR1094 have similar expression in LF and dactyl. In *P. clarkii*, both IR1093 and IR1094 have higher expression in LF compared to dactyl. In *C. sapidus*, IR1091, IR1093, and IR1095 have higher expression in LF compared to dactyl, while IR1092 has higher expression in dactyl than LF (Supplemental S4-S11 Tables: https://doi.org/10.1371/journal.pone.0230266). We also found homologues to these four conserved IRs (IR21a, IR40a, IR68a, and IR75-family) in *L. vannamei* and *H. azteca*.

Among chelicerates, we found homologues to IR25a, IR76b, and IR93a in several species (Figure 3.5). However, we found IR8a only in *Limulus polyphemus*, similar to Vizueta et al. (2018). We also identified putative homologues to IR40a in *Parasteatoda tepidariorum* and *Centruroides sculpturatus*, and putative homologues to IR68a in *Galendromus occidentalis*. We identified putative homologues to the tuning IR, IR1039, which has homologues in all four decapod crustaceans and in *H. azteca*, *L. polyphemus*, *P. tepidariorum*, and *C. sculpturatus* (Figure 3.5). As IR21a, IR40a, IR68a, and IR75-family are found in at least two major groups of arthropods – Crustacea and Insecta (Pancrustacea), we consider these tuning IRs as "arthropod-conserved." Similarly, IR1039 is also considered an "arthropod-conserved," since it is found in crustaceans and

chelicerates. The remaining tuning IRs from the four decapod crustaceans, which include most of them, were expressed in one to four of the decapod species examined in our study. Thirty-four IRs have homologues in at least three decapod species, and many IRs have homologues in at least two species (Figure 3.4, Table 3.2). IRs with homologues in all four decapod species might be a conservative estimate of "decapod-conserved" IRs while IRs with homologues in two of the four decapods is a more liberal estimate of "decapod-conserved" IRs. Species-specific IRs are operationally defined as those expressed in only one of the four decapod species and without homologues being found in other arthropod species (Figure 3.4, Table 3.2).

The DESeq2 analysis showed that for all four species, most of the tuning IRs are more highly expressed (> ~ 2.8 fold difference) in one or the other chemosensory organ (Figure 3.4b), and almost without exception, they were more highly expressed in the LF than dactyl (Supplemental S4-S11 Tables: https://doi.org/10.1371/journal.pone.0230266). None of the dactyl-enriched tuning IRs (Figure 3.4) have homologues in all four species. The distribution and expression of IRs across the tissue types and species are diverse. The only consistent pattern among the four decapod species is that almost all IRs found so far are more highly expressed in the LF than dactyl (Supplemental S4-S11 Tables: https://doi.org/10.1371/journal.pone.0230266).

Some IRs previously identified in various insect species (e.g. IR31a, IR60a, and IR64a) (Croset et al. 2010, Rytz et al. 2013) were searched for but not identified in these decapod transcriptomes, and thus they are probably "insect-conserved" tuning IRs.

3.3.1.4 Immunolocalization of IR25a in LF and dactyl of P. argus *and* H. americanus

We compared the expression of IR25a in sensory neurons in the LF and walking leg dactyls of *P. argus* and *H. americanus*. The general organization of aesthetascs and the cells associated with them is similar for *P. argus* (Figure 3.6a, (Kozma et al. 2018)) and *H. americanus* (Figure 3.6e). Each aesthetasc is innervated by ca. 300 OSNs, whose somata form a cluster near the base of the seta, with dendrites extending into the sensilla and axons projecting to the brain. Besides OSNs, there are two other cell types associated with the aesthetascs, and these are easily distinguishable based on their position and shape (Schmidt et al. 2006). Auxiliary cells ensheathe the bundle of OSN inner dendritic segments, forming a strand of flat lenticular nuclei arranged in a tube-like fashion around those dendrites (Figure 3.6c, g: arrows). A second cell type is the tegumental gland cells, which are located in the spaces between the OSN somata clusters and the dendritic bundles (Figure 3.6a, c, e: asterisks), and which form distinctive rosettes with a duct projecting from the gland to the cuticular surface.

For the LF, in *P. argus*, anti-HaIR25a intensely labeled the clusters of OSN somata associated with the aesthetascs (Figure 3.6a-d). Simultaneous labeling with anti-tubulin (to preferentially label neurons – see Methods) and Hoechst 33258 (to label nuclei of all cells) revealed that all or close to all OSN somata of a cluster are HaIR25a-positive as long as they are not so deep in the cluster that the antibody cannot penetrate to react with them (see Methods) while adjacent auxiliary cells are HaIR25a-negative (Figure 3.6c, d). The HaIR25a-like immunoreactivity is present in the cytosol and cell membrane of the OSN somata, as well as in the inner- and outer dendritic segments (Figure 3.6d). In *H. americanus*, anti-HaIR25a also labeled the clusters of OSN somata associated with the

aesthetascs of the LF, but with slightly lower intensity (Figure 3.6e-h). Like in *P. argus*, simultaneous labeling with anti-tubulin and Hoechst 33258 revealed that all or close to all OSN somata of a cluster are HaIR25a-positive while auxiliary cells, some of which are located within the cluster of OSN somata, are HaIR25a-negative (Figure 3.6g, h). In the OSN somata, HaIR25a-like immunoreactivity is most uniform and intense in the cell membrane, whereas labeling of the cytosol is more variable in intensity (Figure 3.6h). As in *P. argus*, HaIR25a-like immunoreactivity extends into the dendrites of the OSNs where it is distinctly more intense in the outer- compared to the inner dendritic segment (Figure 3.6i, j).

For the walking leg dactyls, in *P. argus*, anti-HaIR25a labeled the fusiform clusters of somata of sensory neurons associated with the clustered smooth setae (Figure 3.7a-f). Simultaneous labeling with anti-tubulin and Hoechst 33258 revealed that almost all somata of putative CSNs (nuclei characterized by dense heterochromatin) in the proximal part of each sensory neuron cluster are intensely labeled by anti-HaIR25a, while the remaining 1–3 putative CSNs are less intensely labeled. In addition to the putative CSNs, each cluster of sensory neurons contains 2–3 larger somata of putative MSNs (nuclei characterized by very loose heterochromatin) at its distal pole that are HaIR25a-negative (Figure 3.7c-f). In *H. americanus*, anti-HaIR25a also labeled the fusiform clusters of somata of sensory neurons associated with the tufts of smooth setae of the walking leg dactyls (Figure 3.7g-k). Simultaneous labeling with anti-tubulin and Hoechst 33258 revealed that in contrast to the situation in *P. argus*, only about half of the somata of putative CSNs (nuclei characterized by dense heterochromatin) in the proximal part of each sensory neurons cluster are intensely labeled by anti-HaIR25a while the remaining half are only lightly labeled (Figure 3.7i, j). As in *P. argus*, each cluster of sensory neurons

contains 2–3 larger somata of putative MSNs (nuclei characterized by very loose heterochromatin) at its distal pole that are HaIR25a-negative (Figure 3.7k).

In both species, each cluster contains about 20 bipolar sensory neurons. In *P. argus*, all sensory neurons are strongly labeled by anti-tubulin, but in *H. americanus*, the labeling with anti-tubulin is not as intense and more diffuse. In *P. argus*, all sensory neurons located in the proximal part of the cluster are also intensely labeled by anti-HaIR25a, but three particularly large neurons located at the distal pole of the cluster are not labeled by anti-HaIR25a (nuclei labeled by asterisks), and one additional neuron in the distal region is only weakly labeled by anti-HaIR25a (arrowhead). In contrast, in *H. americanus*, only some of the sensory neurons in the proximal part of the cluster are intensely labeled by anti-HaIR25a and about equally many are weakly labeled (arrowheads). Two particularly large neurons located at the distal pole of the cluster are not labeled by anti-HaIR25a (nuclei labeled by asterisks). Note that in both species, the nuclei of the large HaIR25a-negative neurons at the distal pole of the cluster (asterisks) are characterized by very loose heterochromatin, whereas the heterochromatin of the other sensory neurons is considerably denser.

3.3.2 TRP channels

Homologues to all seven subfamilies in both groups of TRP channels (see Figure 3.1) were previously found in the LF and dactyl transcriptomes for *P. argus* (Kozma et al. 2018). Homologues of all seven subfamilies were also found in the three additional crustacean transcriptomes analyzed here (Figure 3.8, Supplemental S16-S18 Files: https://doi.org/10.1371/journal.pone.0230266). Within each subfamily of TRP channels, those from decapod crustaceans form their own cluster to corresponding

homologous clusters of TRP channels from other species. In Group 1 TRP channels of arthropods, the TRPA subfamily has an expanded family of channels, and a similar expansion was found in these decapod crustaceans. Similar to insects, the TRPA family in the four decapods is expanded with homologues to TRPA1, painless, TRPA5, and pyrexia/waterwitch. *H. americanus* has one additional sequence (HameTRPApw-1) that is more closely related to the pyrexia/waterwitch clade, but with low bootstrap support, which was not detected in the other decapod crustaceans. Additionally, the four decapod crustaceans also have TRPA sequences that do not have homologues in insects (TRPA1-like1 and TRPA1-like2). Unlike insects, there are two genes that belong to the TRPM subfamily in crustaceans. Homologues to the insect TPRM channel were found in all four species. *P. argus*, *H. americanus*, and *P. clarkii* have an additional TRPM channel (TRPMc – previously denoted as TRPMm in Kozma et al. (2018)) (Figure 3.8), which clusters away from the arthropod conserved group of TRPM channel sequences, while *C. sapidus* has only the insect-like TRPM channel. The branchiopod *Daphnia pulex* has two TRPM channels as well; however, both are more closely associated with the insect TRPM channel. Therefore, it is possible that the TRPMc channel is specific to decapods, but this needs to be resolved by the inclusion of TRPM channels from other crustaceans and protostomes in analyses.

Among the TRPC subfamily, there were two homologues in the decapods. Homologues to TRPgamma were detected in *H. americanus* and *P. clarkii* along with the homologue previously discovered in *P. argus*. Homologues to the TRPC channel previously detected in *P. argus* were also detected in the other three decapods. In the TRPN subfamily, homologues to NompC were detected in all four species, along with an additional TRPN channel that was found only in *P. clarkii*. In the TRPV subfamily,

homologues to Inactive (Iav) and Nanchung (Nan) were found in all four species. An additional Nanchung homologue was also detected in *P. clarkii* and *C. sapidus*.

Homologues of Group 2 TRP channels – TRPP (Pkd2) and TRPML – were also found in all four decapod species. Two TRPML sequences were detected in *P. clarkii*: TRPML1 and TRPML2.

Based on DESeseq2 analysis, in *P. argus*, TRPA5-1 and TRPgamma have higher expression in LF compared to dactyl, while NompC, Iav, and Nan have higher expression in dactyl compared to LF. In *H. americanus*, NompC, TRPgamma, and TRPC have higher expression in LF compared to dactyl, while TRPApw and Iav have higher expression in dactyl than LF. In *P. clarkii*, almost all TRP channels have differential expression between the LF and dactyl tissues. TRPML1, TRPML2, NompC, and NompC1 have higher expression in dactyl than LF, while TRPA5-1, TRPA5-2, Pain2, TRPA-like, TRPA1, Iav, Nan, Nan2, TRPC, and TRPMc have higher expression in LF compared to dactyl. In *C. sapidus*, Nan2 has higher expression in dactyl than LF, while TRPA1-like, TRPA1-like1, and TRPC have higher expression in LF compared to dactyl. The remaining TRP channels in all four decapod species studied here have similar expression in LF and dactyl.

3.3.3 GRs and GRLs

Using InterProScan for domain search of the 7tm_7 domain region (PF08395) and the criteria for identification of GRs and GRLs as described by Robertson (2015), we detected GRLs in each of the four decapod species (Figure 3.9, Supplemental S19–S20 Files: https://doi.org/10.1371/journal.pone.0230266). The number of GRLs range from one to four across the four decapods. We did not detect homologues to these GRLs in other crustaceans or insects. There is one GRL, PargGR1, identified previously in *P. argus*

(Kozma et al. 2018). Although PargGR1 has low expression in the transcriptome, it is three-fold more abundant in the LF than dactyl (Supplemental S4 and S8 Tables: https://doi.org/10.1371/journal.pone.0230266). In *H. americanus*, there are four GRLs: HameGR1 is expressed in both tissues and is 10 fold more abundant in dactyl than LF; HameGR2 has similar expression in both tissues; HameGR3 has low expression in LF; and HameGR4 expression is low but similar in both tissues. In *P. clarkii*, there are two GRLs: PclaGR1 has low level of expression in both tissues without much difference in abundance; and PclaGR2 has a low level of expression and only in the dactyl. *C. sapidus* has only one GRL, CsapGR1. CsapGR1 was identified through InterProScan in the transcriptome before redundancy was removed; however, it was not subsequently detected in the non-redundant transcriptome.

3.3.4 Expression of putative chemoreceptor proteins in the brain

We previously discovered expression of several IRs including IR25a, IR8a, and IR93a in the brain transcriptome of *P. argus* and found evidence of expression of these co-receptor IRs in the brain through PCR (Kozma et al. 2018). Several other conserved IRs and tuning IRs with low expression were also detected in the brain transcriptome. Using immunocytochemistry, we discovered that IR25a is only localized to very large and conspicuous cells in the axon sorting zone in the lateral division of the antennular nerve entering the brain in *P. argus* (Kozma et al. 2018). Here, we generated brain transcriptomes and detected expression of all four co-receptor IRs in *H. americanus* and *P. clarkii*, similar to *P. argus*. We also detected expression of IR21a, IR40a-family, and IR75-family in the brains of these three species. In *H. americanus*, we also detected

expression of IR68a. Twenty to fifty tuning IRs are expressed in the brain transcriptomes of *P. argus*, *H. americanus*, and *P. clarkii*, with *P. clarkii* on the lower end of the range and *H. americanus* on the higher. In *P. argus*, there are no IRs that have higher expression in the brain compared to LF with the exception of IR1093 from the IR75-family: expression of IR1093 in the brain is 28 fold higher than LF and 4.9 fold higher than dactyl. In *H. americanus*, IR2067 has 21 fold higher expression in brain than LF, IR2032 in the brain has 4.3 fold higher expression than LF and 3.1 fold higher expression than dactyl, IR2005 in the brain has 7.9 fold higher expression than dactyl, and IR1034, IR1092, and IR1095 of the IR75-family have 8.6 fold, 6.1 fold, and 8.7 fold higher expression respectively in the brain compared to dactyl. In *P. clarkii*, IR1094 of the IR75-family has 94 fold higher expression in the brain than dactyl.

All TRP channels from all subfamilies that were detected in the LF and dactyl of *P. argus*, *H. americanus*, and *P. clarkii* were also detected in their brain transcriptomes with varying levels of expression (Supplemental S4-S6 Tables: https://doi.org/10.1371/journal.pone.0230266, Supplemental S1 File: https://doi.org/10.1371/journal.pone.0230266.s014). The only exception was the TRPN channel of *H. americanus*, NompC, which had little to no expression in the brain. In *H. americanus,* TRPgamma had higher expression in brain than LF and dactyl by 935 fold and 5000 fold respectively, and TRPC was 32 fold more abundant in brain than dactyl. In *P. clarkii*, TRPML1 and TRPML2 were 3.5 fold and 3 .0 fold more abundant in brain than LF, Pain2 was more abundant in brain than dactyl by 8.6 fold, TRPA1-like2 in brain was more highly expressed than LF and dactyl by 5.3 fold and 6.2 fold respectively, TRPV channel Nan was expressed more highly in the brain than dactyl by 3.9 fold, TRPgamma

was more abundantly expressed in brain than LF and dactyl by 352 fold and 358 fold respectively, and TRPC was more highly expressed in the brain than dactyl by 23 fold.

GR expression was found in the brain of *H. americanus* for HameGR1, HameGR3, and HameGR4. In fact, HameGR4 was much more abundantly expressed in the brain than in the LF or dactyl by 74 fold and 36 fold respectively.

3.4 Discussion

Animals have a diversity of types of chemoreceptor proteins with expression patterns that differ phylogenetically. Arthropods are a major animal phylum, for which its largest clade – the insects – has been the focus of research on the molecular identity of chemoreceptors (Robertson 2019). The other major clades of arthropods, including crustaceans, have received much less attention. Our work contributes to our understanding of the evolution of chemoreceptor proteins and helps support future research on crustacean chemoreception by a comparative analysis of chemoreceptor proteins in four species of decapod crustaceans that are used as models of chemoreception: Caribbean spiny lobster *Panulirus argus*, American lobster *Homarus americanus*, red swamp crayfish *Procambarus clarkii*, and blue crab *Callinectes sapidus*.

3.4.1 Evolution and function of crustacean IRs

IRs are an ancient group of chemoreceptor proteins, being present in all protostomes and well represented in the decapod crustaceans. A conservative estimate of the number of different IRs expressed in the two chemosensory organs of the four species examined in our study, based on sequences that have both of the major domains of IRs, is ca. 250

for *P. argus*, ca. 170 *for H. americanus*, ca. 100 for *P. clarkii*), and ca. 180 for *C. sapidus*. Since IRs, like the iGluRs from which they evolved, are heterotetramers, these IRs exist in combinations to form functional channels. One or two of the constituent subunits are co-receptor IRs and the other two to three subunits are tuning IRs (Abuin et al. 2019).

Homologues to four co-receptor IRs – IR25a, IR8a, IR76b, and IR93a – exist in the transcriptomes of both chemosensory organs of the four species of decapod crustaceans examined here. IR25a is the most ancient of IRs, being a protostome conserved IR that is absent from the deuterostomes and clades ancestral to the protostome-deuterostome split (i.e., Placozoa, Porifera, Cnidaria, and Ctenophora: (Croset et al. 2010), (Eyun et al. 2017), (Poynton et al. 2018)). IR8a is an arthropod conserved co-receptor IR. There is limited information about the expression of IR25a and IR8a in the crustaceans, unlike in the insects, for which OSNs are known to express either IR25a or IR8a, and sometimes both. IR93a is expressed only in the antenna of *Drosophila*, whereas IR76b is found in all of *Drosophila*'s chemosensory tissues (Sánchez-Alcañiz et al. 2018). More specifically, IR93a is co-expressed in 10 to 15 neurons surrounding the sacculus on the antenna of the fly. The tuning IRs IR40a, IR68a, and IR21a are co-expressed with IR93a in these neurons, and this combination of IRs specifies hygro- and thermosensation rather than chemoreception (Enjin et al. 2016, Knecht et al. 2016, Frank et al. 2017, Knecht et al. 2017). In crustaceans, studies of cellular expression using immunocytochemistry and/or *in situ* hybridization show that IR25a is broadly expressed in OSNs, being present in many or most OSNs in *P. argus*, *H. americanus*, and *C. clypeatus* (Hollins et al. 2003, Stepanyan et al. 2004, Tadesse et al. 2011, Corey et al. 2013, Groh-Lunow et al. 2014, Kozma et al. 2018). IR25a is also broadly expressed in CSNs of various chemosensory organs of *P. argus* (Kozma et al. 2018). Our immunocytochemical results on *P. argus*

show that IR25a is expressed in all or most OSNs and chemosensory neurons (CSNs) in the dactyls but not in mechanosensory neurons (MSNs). Furthermore, the labeling is strongest in the outer dendrites of the OSNs, where receptor proteins are highly expressed (Blaustein et al. 1993, Hatt and Ache 1994). Our results on *H. americanus* are largely similar, except that only about half of the dactyl CSNs are labeled. The reason for this difference in extent of expression of IR25a in dactyl CSNs between *P. argus* and *H. americanus* is not clear. Considering IR8a, its cellular expression pattern has not been described, though IR8a is known through PCR experiments to be expressed in several chemosensory and non-chemosensory organs in *P. argus* (Kozma et al. 2018). Furthermore, organ-level expression levels for IR8a appear to be lower than for IR25a, suggesting that IR8a is expressed in a more limited number of cells than is IR25a. IR93a is reported to be expressed in all or most OSNs of *P. argus*, based on in situ hybridization studies (Corey et al. 2013). In insects, IR25a and IR8a are essential for targeting IRs to the OSN dendritic membrane and thus becoming a functional channel, and are not necessary for expression of the channels' chemical specificity (Abuin et al. 2011, Abuin et al. 2019). It might be expected that crustacean co-receptor IRs act similarly, but that has not been studied.

The tuning IRs make up the vast majority of the IR repertoire in crustaceans, similar to insects. Conservatively, this is at least 254 in *P. argus*, 181 in *H. americanus*, 92 in *P. clarkii*, and 186 in *C. sapidus*. *Daphnia pulex* has ca. 150 tuning IRs, *Hyalella azteca* has 114, and *Eurytempora affinis* has 18 (Croset et al. 2010, Rytz et al. 2013, Poynton et al. 2018). Some tuning IRs in crustaceans are more phylogenetically conserved than others. For example, to date, IR21a, IR40a, and IR75 have been found to have homologues only in pancrustaceans and thus are likely expressed only in crustaceans and insects. IR1039

is conserved in crustaceans and chelicerates, but is not detected in insects. Many tuning IRs have limited phylogenetic expression, where homologues are detected only in species of the same order and in some cases only within a particular species. Using phylogeny as the basis, we introduced an IR classification system, where tuning IRs were labeled as "protostome conserved," "arthropod conserved," "pancrustacean conserved," "crustacean conserved," and "decapod conserved." This classification scheme can be expanded to other clades across protostomes. Although we performed a search for homologues of IRs across non-hexapod arthropods, it was by no means exhaustive. A broader and more thorough phylogenetic examination of IRs across protostomes would reveal lineages of conserved IRs and possibly give more insight into their function and evolutionary history across species. From our data, there appear to be candidate crustacean conserved tuning IRs (e.g. IR1020, Figure 3.5), decapod conserved IRs (e.g. IR1001, IR1018, IR1020, IR1021, IR1029), and arthropod conserved IRs (e.g. IR1039, IR21a, IR40a, IR75). Other tuning IRs appear to be insect conserved (e.g. IR31a, IR60a, IR64a) (Croset et al. 2010, Rytz et al. 2013, Eyun et al. 2017). Finally, specific-specific IRs have been found in crustaceans as well as insects (Croset et al. 2010, Rytz et al. 2013, Kozma et al. 2018, Robertson et al. 2018). Although the chemical specificity of tuning IRs has been described for many IRs in *Drosophila* and other insect species (Silbering et al. 2011), there have been no functional studies of IRs in crustaceans. So, while one might assume that a given IR expressed in different species confers the same response specificity, we do not have any data at this time to verify these assumptions. One particular case in point, IR40a, has been shown to mediate hygrosensation in fruit flies. Not only are there homologues to IR40a in all four aquatic decapods that we examined, but there are multiple homologues

to IR40a in each species. If it is true that a given IR retains its specificity across species, then the nature of that specificity awaits future study.

How does the total number of IRs in these decapod crustaceans compare to other crustaceans and insects? Interestingly, *Daphnia pulex* has at least 210 types of chemoreceptor proteins in its genome – 154 IRs and 56 functional GRs – but where in the body those are expressed is not known. This is a relatively high number of chemoreceptor proteins, despite *Daphnia*'s highly reduced olfactory system. The copepod *Eurytempora affinis* has at least 83 types of chemoreceptor proteins – 22 IRs and 61 GRs (Eyun et al. 2017, Poynton et al. 2018). The antennular chemoreceptor proteins in two species of hermit crabs have been studied, and they have up to 29 IRs and no ORs or GRs per species (Poynton et al. 2018). Insect species have been described as typically having 20 to 150 IRs (Rytz et al. 2013), but they also have ORs and GRs often in numbers equal to the IRs. For example, *D. melanogaster* has ca. 60 IRs and 68 GRs, plus ca. 60 ORs, for a total of nearly 200 functional chemoreceptor proteins. Thus, these four decapod crustaceans appear to have a relatively high number of IRs compared to many other crustaceans, and this high diversity plus the heterotetrameric organization of IRs allows many combinations and thus diverse chemical response spectra. We do not yet know the number of different IRs expressed in individual receptor neurons, but multiple types of heterotetramers are theoretically possible in single cells.

The levels of expression of IRs differ between the two chemosensory organs, the LF and dactyl. For most IRs, the expression level is higher ($\log_2$ fold change ≥ 1.5, or ≤ -1.5) in one chemosensory organ than the other, and in almost all of these cases of disproportionate expression levels, the LF has higher expression than the dactyl. What is the function of these "LF enriched/specific IRs" vs. the "LF-dactyl shared IRs?" This is

not known since functional expression studies have not yet been done. However, it is tempting to speculate that the IRs shared by the LF and dactyl are sensitive to the compounds that both appendages are known to detect – food-related compounds such as amino acids, amines, nucleotides, nucleosides (Derby and Weissburg 2014). On the other hand, the LF-specific IRs might be expected to detect chemicals that only the LF senses – and these are pheromones, including sex, social, and alarm chemical signals and cues (Derby and Weissburg 2014). The legs are rarely described as being involved in the detection of pheromones (Derby and Weissburg 2014), though there are exceptions (Belanger and Moore 2006). This correlates with a lack of dactyl rich expression of IRs. In any case, the trend toward greater expression of IRs in LF vs. dactyl may be due to their expression in higher numbers in olfaction (i.e. in OSNs, which are found only in LF) than in distributed chemoreception (i.e. in CSNs, which are found in LF and dactyl).

3.4.2 Chemoreception beyond IRs

The IRs appear to be the major chemoreceptor proteins in most crustaceans, but others are likely to contribute to their total chemosensory repertoire. Most likely are TRP channels and to some extent GRs.

3.4.2.1 TRP channels

Previously, we found in *P. argus* members of every sub-family of the two groups of TRP channels, including those that are known to be chemoreceptors in other species including mammals, insects, and nematodes. In this study, we also found homologues of all these TRP channels in the other three decapod species. This includes the four

subfamilies of TRP channels known to have chemosensory functions in other species: TRPA, including those related to TRPA1, painless, TRPA5, pyrexia, and waterwitch of insects. Similar to *P. argus*, there were no direct homologues to pyrexia or waterwitch with high branch support. Instead, there was a homologue in each decapod that clustered with the clade containing both pyrexia and waterwitch, and one additional sequence in *H. americanus* that clustered more closely with the pyrexia/waterwitch clade with low bootstrap support (Figure 3.8); TRPV, including those related to OSM-9 of *C. elegans* and Nanchung and Inactive of *D. melanogaster*; TRPC channels; and TRPM, including those related to insects and mammalians, and these seem to be particularly expanded in the crustaceans. Decapods appear to have an additional TRPM channel compared to insects. This additional TRPM channel (TRPMc) that was found in *P. argus*, *H. americanus*, and *P. clarkii*, clearly clusters away from the TRPM channel homologue found in insects. While *Daphnia* also has multiple TRPM channels, all of these cluster closer to the insect TRPM channel. We can only speculate about the function of these crustacean TRP channels, as sequence homology does not necessarily signify functional homology for TRP channels. However, it is reasonable to speculate that several of the crustacean TRP channels may function as chemical sensors, due to the multimodal nature and known chemo-sensitivity of TRP channels across animals.

3.4.2.2 GRs and GRLs

Using multiple sequence alignment of GRs and InterPro domain search, we confirmed our prior identification of one GRL in *P. argus*, and extended this by finding GRLs in the other three decapod species. These were found in both the LF and dactyl transcriptomes. GRs and/or GRLs have been found in other crustacean species, usually

in low numbers (Eyun et al. 2017, Kozma et al. 2018). However, a recent paper by Poynton et al. (2018) (Poynton et al. 2018) provides new insight into the potential importance of GRs in crustaceans. They annotated the GR gene family of *Hyalella azteca*, generated improved models of the genome assemblies for two crustacean species, and showed considerable expression of GRs in these species. The amphipod *H. azteca*, the branchiopod *Daphnia pulex*, and the copepod *Eurytempora affinis* have 155 (of which 41 are pseudogenes), 59 (3 pseudogenes), and 67 (6 pseudogenes) GRs respectively. Among crustaceans, it is not known whether GRLs are expressed in chemosensory cells, and no functional studies have been done. Our results show that GRLs also exist in the brain (at least for *H. americanus*), and GRLs have also been found in the transcriptome of the Y-organ, an endocrine gland, in the decapod crustacean *Gecarcinus lateralis* (blackback land crab) (Tran et al. 2019). There is some evidence in species other than crustaceans that the GRLs may play roles in development (Saina et al. 2015), and any role that GRLs may have in chemoreception is speculative.

3.4.2.3 Others?

Beyond IRs, TRP channels, and possibly GRLs, other classes of chemoreceptor proteins have been identified in arthropods and other protostomes. One major class is the ORs, which appear to have evolved from GRs and to date have been identified only in insects (Benton 2015, Joseph and Carlson 2015, Robertson 2015). We did not find any evidence for ORs in crustaceans in our work. Another class of chemoreceptor proteins are GPCRs, which have been shown or suggested to function as chemoreceptor proteins in some protostomes (Cummins et al. 2009a, Albertin et al. 2015, Roberts et al. 2017,

Roberts et al. 2018). We found many rhodopsin-like GPCRs in our decapod transcriptomes, but most could be identified as homologues of known classes of GPCRs that are not mammalian ORs, and to date we have not found expanded families of orphaned GPCRs as might be expected if they function as chemoreceptor proteins (M. Rump, M. Kozma, and C. Derby unpublished results). Epithelial sodium channels (ENaC) are a class of ionotropic receptors that have been shown to be used by *Drosophila* for detecting salt (Zhang et al. 2013), water (Cameron et al. 2010, Chen et al. 2010), and pheromones (Lu et al. 2012, Pikielny 2012, Thistle et al. 2012), in at least one case by activating downstream of IRs, ORs, or GRs in a calcium-dependent amplification step (Ng et al. 2019). Though we found ENaCs in all four decapod transcriptomes, we did not find homologues of ppk23 and ppk28, the chemosensory *pickpocket* genes that are ENaC homologues in insects. Clearly, more analysis is necessary in order to fully evaluate the role of GPCRs, ENaCs, and other classes of receptor proteins in chemical sensing in crustaceans.

3.4.3 Olfactory logic in decapod crustaceans

Given the heterotetrameric combinatorial nature of IRs and that there are over 100 co-receptor and tuning 100 IR units in the chemosensory organs of each of these four decapod species, and given the other candidate chemoreceptor proteins in these tissues including TRP channels, GPCRs, and ENaCs expressed, the breadth and scope of receptor molecules in these systems is potentially very large. To understand the olfactory logic in decapod crustaceans, one must know the patterns of expression of these receptor molecules in single OSNs, as well as the central projections of OSNs with defined expression patterns. For example, the antennule of *P. argus* and *H. americanus* has *ca.*

300,000 OSNs, and those OSNs project into *ca*. 1,200 glomeruli in the olfactory lobe (Beltz et al. 2003). Electrophysiological studies in lobsters suggest potentially dozens to hundreds of different physiological classes of chemosensory neurons (Derby 2000, Steullet et al. 2000, Atema 2018), which might lead to the expectation of a high diversity of receptor expression patterns in the population of OSNs. This scenario raises the possibility of an olfactory logic in decapod crustaceans that is significantly different than in insects or mammals, in which the ratio of the number of types of receptor molecules to glomeruli is often *ca*. 1:1, with some exceptions (Vosshall 2001, Galizia and Sachse 2010, Murthy 2011). To explore the olfactory logic in crustaceans, we need to perform single cell transcriptomics on hundreds of OSNs whose chemical sensitivities are defined, then classify these cells based on their receptor expression and physiological response profiles, and describe their patterns of innervation in the olfactory lobe.

3.5 Conclusions

Decapod crustaceans have hundreds of candidate chemoreceptor proteins in their olfactory and distributed chemoreception systems. IRs are certainly major chemoreceptor molecules in crustaceans, though there is almost no functional analysis of their roles. More work is necessary to determine the chemical sensitivity of different families of IRs or specific IRs. TRP channels of decapod crustaceans are likely to include some chemoreceptor proteins, as crustacean homologues of TRP channels with chemosensory functions in other species are identified. Still, experimental evidence for their roles in chemoreception are completely lacking. The role of GRs and GRLs in crustaceans in general is more difficult to evaluate, in part because the extent of their expression seems to vary tremendously across the clades of crustaceans, from more than

100 in branchiopods and amphipods to one or few in decapods. Other classes of receptor proteins, including suspected (GPCRs and ENaCs) and others not identified, also need to be further considered as possible candidates. Future studies should use single cell transcriptomics to understand the combinatorial expression patterns of chemoreceptor proteins in single chemosensory neurons, examine function of receptor proteins by examining receptor expression in single cells whose chemical sensitivities are defined, and by experimentally determining the chemical sensitivities of specific receptor proteins through heterologous expression of combinations of receptor molecules and/or through regulation of receptor expression levels in chemosensory neurons.

3.6 Figures and Tables

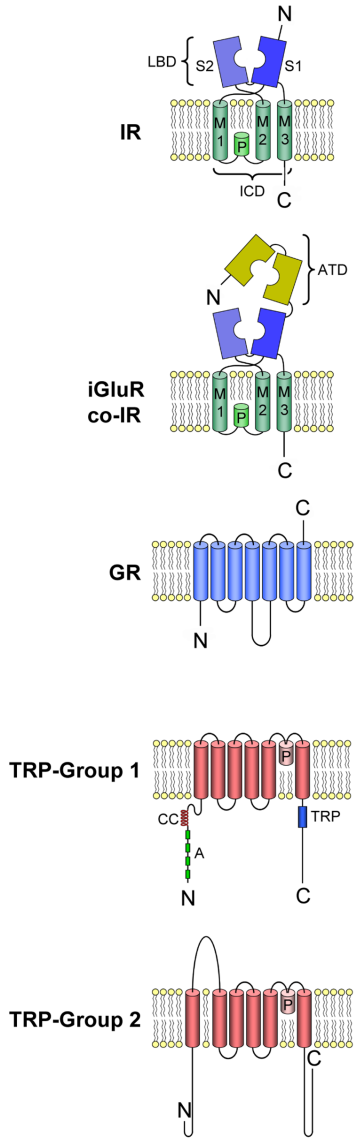

Figure 3.1 Schematic drawing of the molecular structure of putative chemoreceptor proteins co-receptor IRs and iGluRs, tuning IRs, GRs, and TRP channels.

IRs, co-receptor IRs, and iGluRs contain the following domains: extracellular amino terminal domain (ATD); ion channel domain (ICD) that forms the ion channel, consisting of three transmembrane domains (M1, M2, M3) and a pore loop (P); ligand binding domain (LBD) consisting of two half-domains (S1, S2). TRP channels contain the following domains: coiled-coil domain (CC), ankyrin repeats (A), TRP domain (TRP). Adapted from (Kozma et al. 2018).

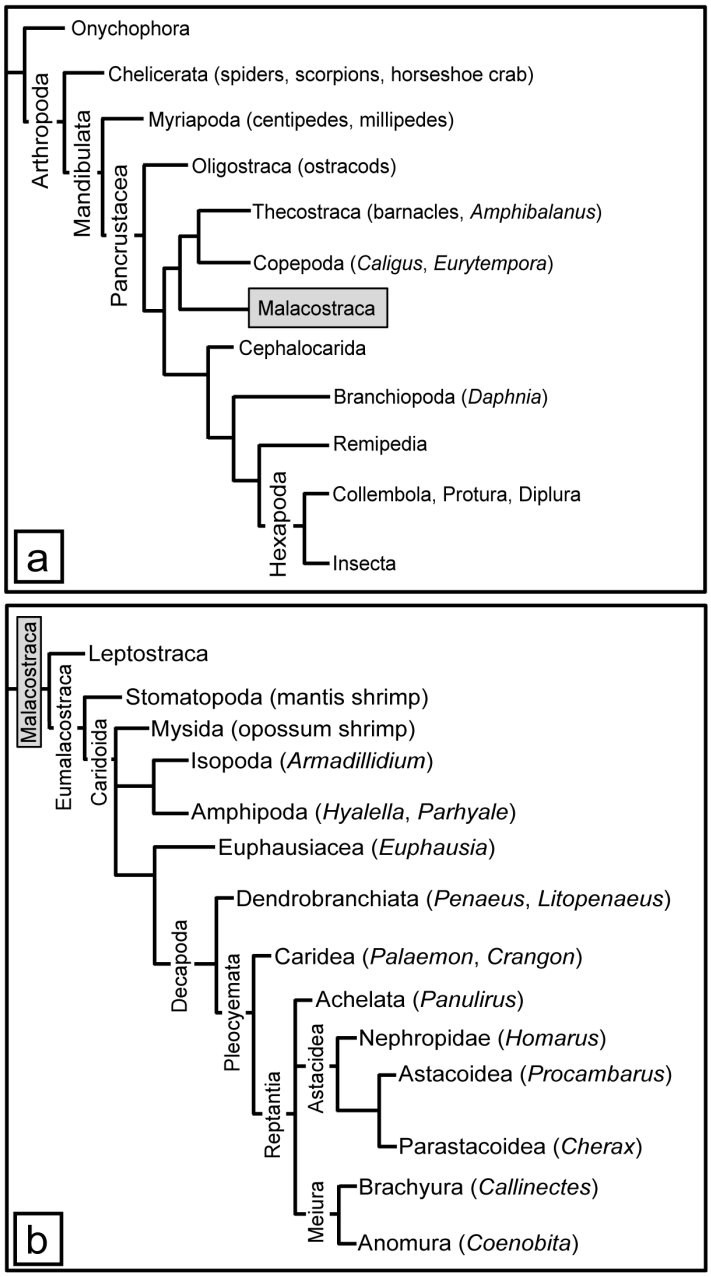

Figure 3.2 Arthropod phylogeny.

Panel **(a)** shows all arthropod groups. Panel **(b)** shows an expanded view of the Malacostraca. Names of species used in our analysis are included. Based on (Meusemann et al. 2010, Regier et al. 2010, Giribet and Edgecombe 2012, Schmidt 2016, Wolfe et al. 2019).

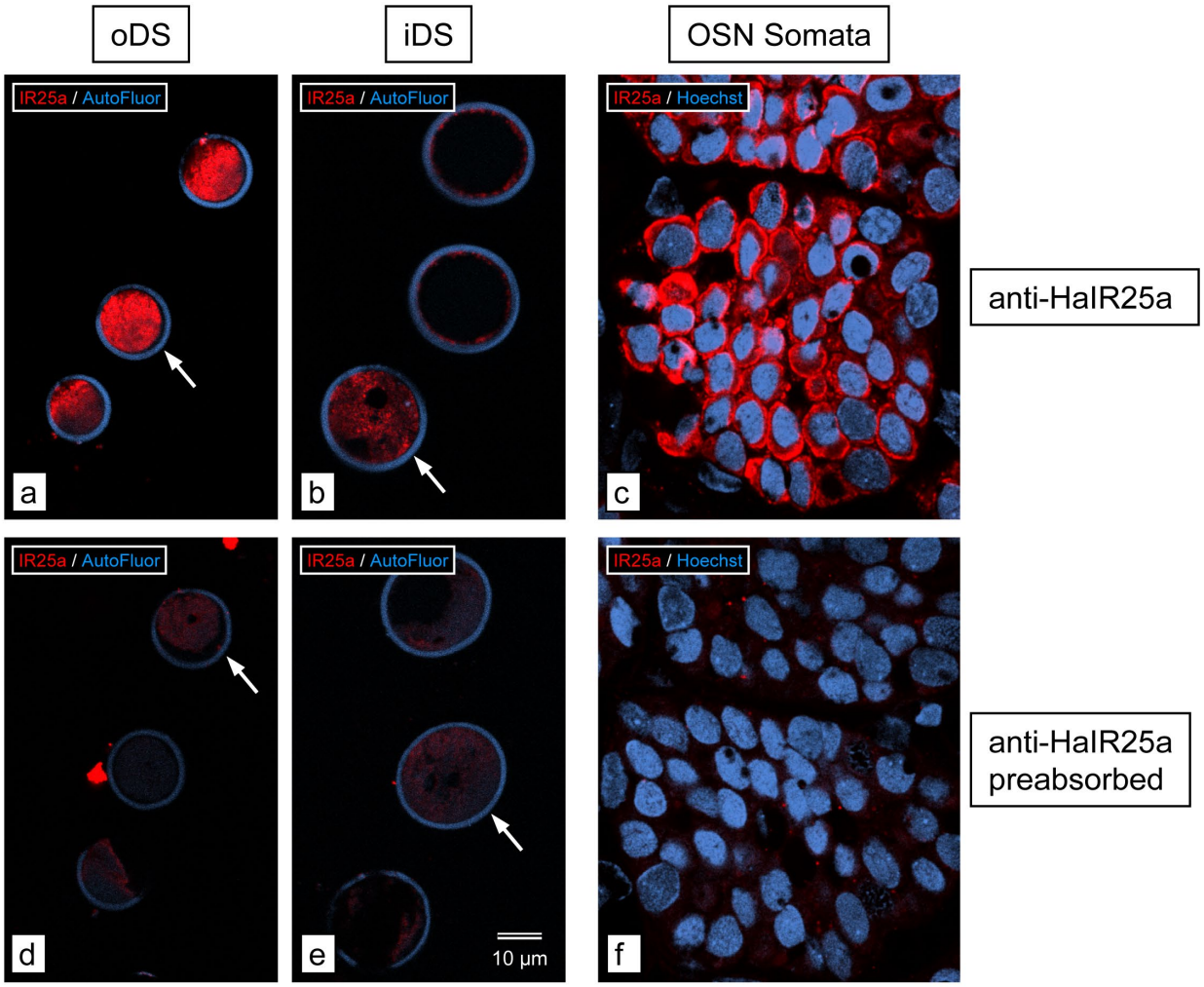

Figure 3.3 Pre-absorption control for anti-HaIR25a.

Cross sections through aesthetascs of *P. argus*; single optical sections of 0.5 µm thickness. **(a – c)** Sections labeled with control anti-HaIR25a (red) and Hoechst 33258 (blue). **(d – f)** Sections labeled with preabsorbed (with $P1_{Pa}$ and $P2_{Pa}$) anti-HaIR25a (red) and Hoechst 33258 (blue). Scale bar in **(e)** applies to all images. Arrows in **(a)**, **(b)**, **(d)**, and **(e)** point to cross-sections of aesthetascs in which anti-HaIR25a labeling is captured at the very surface of the section (highest intensity). Images **(a)**, **(b)**, **(d)**, and **(e)** were collected with the same intensity setting of the red fluorescence channel of the confocal microscope; images **(c)** and **(f)** were collected at a higher (but between **(c)** and **(f)** consistent) intensity setting to compensate for the fact that labeling intensity of anti-HaIR25a is much higher in the inner dendritic segments (iDS) and outer dendritic segments (oDS) of OSNs than in the somata as was previously reported (Kozma et al.

2018). Labeling intensity is high in **(a)** oDS, **(b)** iDS, and **(c)** somata of OSNs with the control anti-HaIR25a, but below detectability in **(d)** oDS), **(e)** iDS, and **(f)** somata of OSNs with preabsorbed anti-HaIR25a, demonstrating the specificity of anti-HaIR25a for IR25a in *P. argus* in addition to that demonstrated by Stepanyan et al. (2004) for IR25a in *H. americanus*.

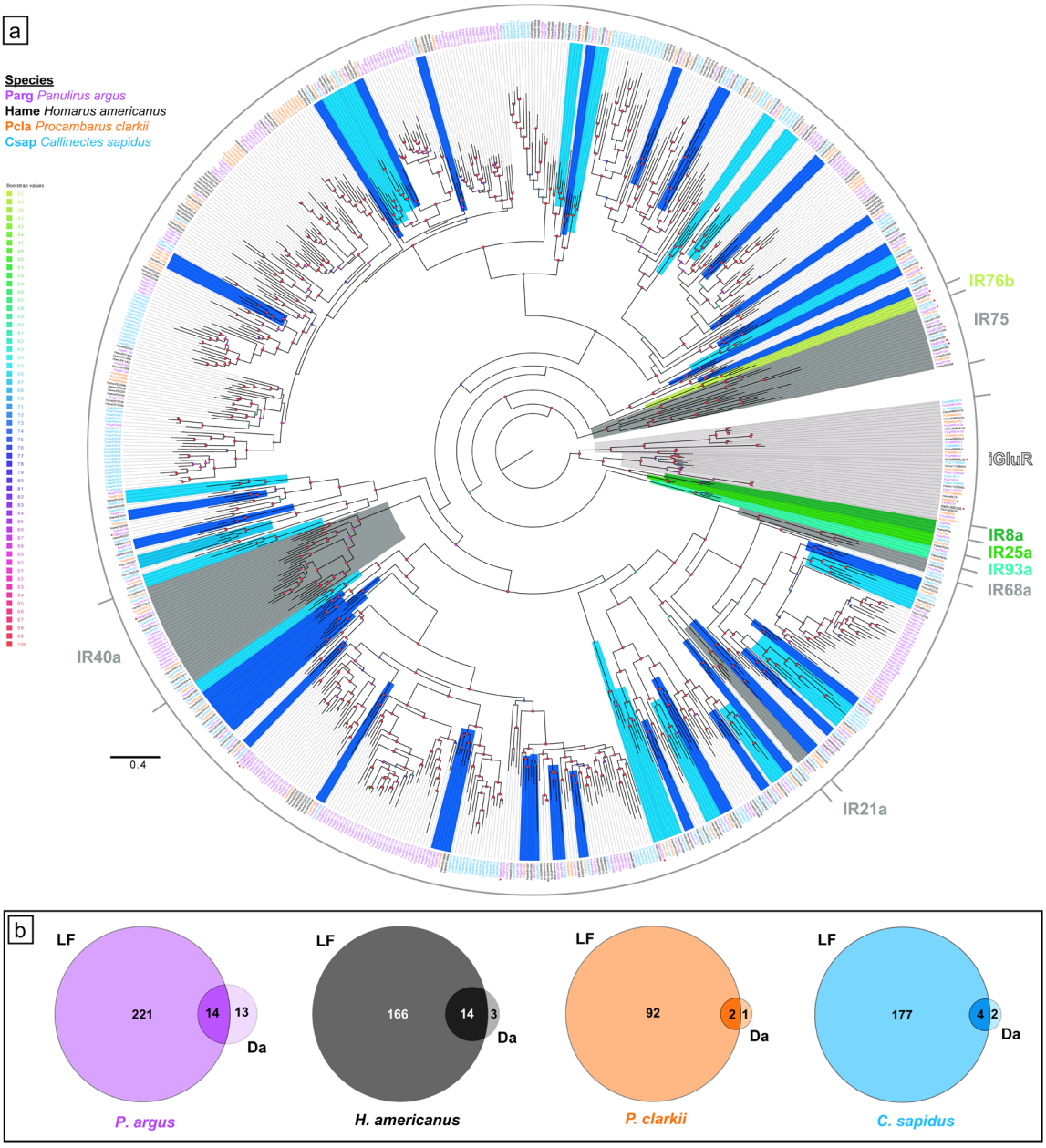

Figure 3.4 Phylogenetic tree of IRs in four decapod species and tissue expression.

(a) Maximum likelihood phylogenetic tree of IRs and iGluRs from four decapod crustaceans. Clades with co-receptor IRs (IR25a, IR8a, IR76b, and IR93a) are colored in shades of green; clades with tuning IRs that are conserved across crustaceans and insects (IR21a, IR40a, IR68a, and IR75-family) are colored dark grey; clades with tuning IRs that are conserved across all four decapod crustaceans are colored light blue; clades with tuning IRs that are conserved in at least three decapod crustaceans are colored dark blue;

clades with iGluRs are colored light grey. * with underline indicates higher expression in dactyl than LF. The tree was built using IQ-Tree with 1000 UFBoot replications under the WAG+F+G4 model of substitution according to BIC as selected by ModelFinder. The tree was visualized on FigTree v.1.4.4. The tree is unrooted but the root is drawn at the iGluR/IR25a/IR8a clade. Scale bar represents expected number of substitutions per site. **(b)** Venn diagrams showing tissue specific differential expression (**LF** – lateral flagella of antennules, **Da** – dactyls of walking legs) of IRs and iGluRs in each decapod crustacean as calculated by DESeq2, where a ~ 2.8 fold difference or greater in expression (i.e. $\log_2$[fold change] ≥ 1.5 or $\log_2$[fold change] ≤ -1.5) between tissue types is considered higher expression in one tissue.

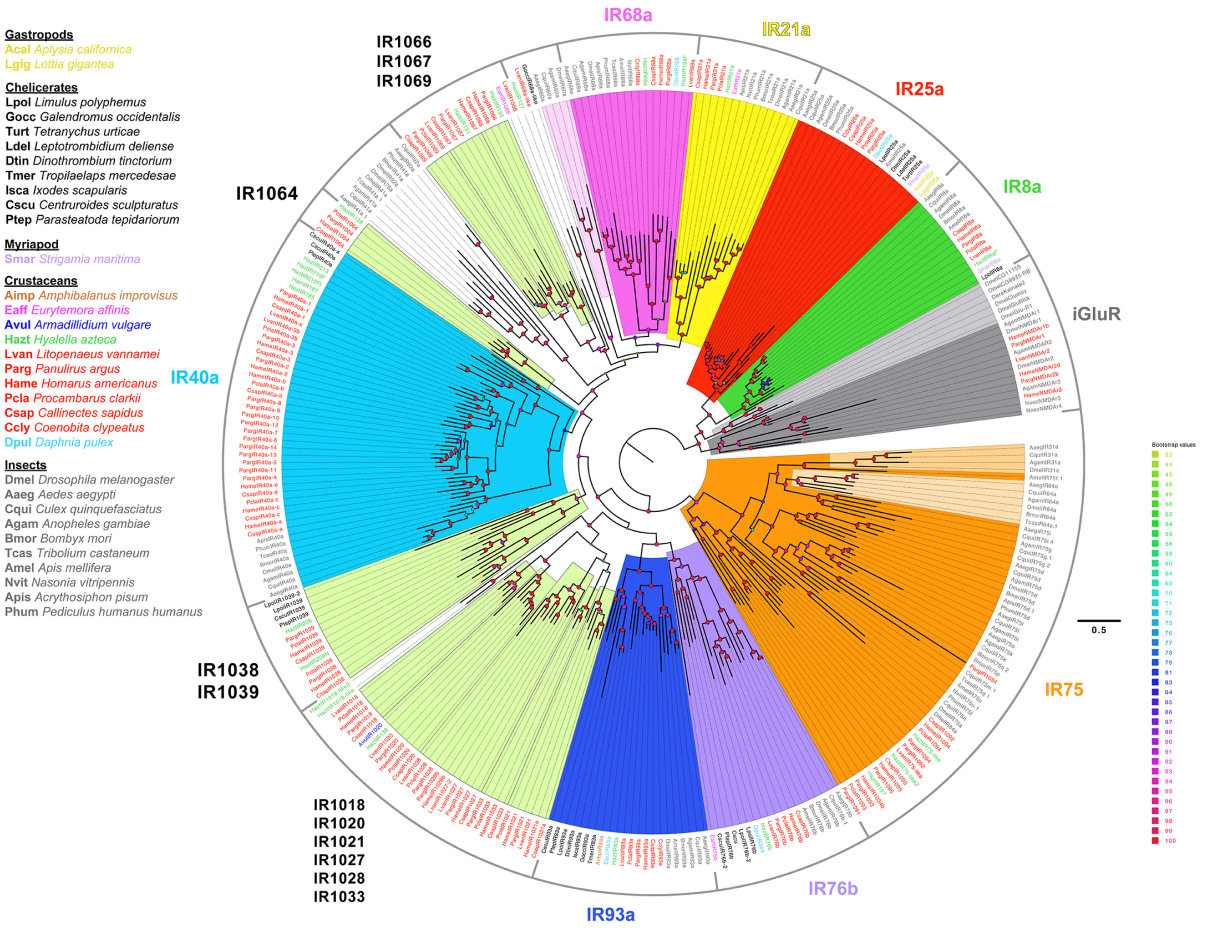

Figure 3.5 Phylogenetic tree of conserved IRs across arthropods.

Maximum likelihood phylogenetic tree shows the different IRs that are conserved across major groups of arthropods: chelicerates, myriapods, crustaceans, and insects. IR25a sequences from two gastropods are also included. Among crustaceans, species are colored by their subclass as follows: thecostracan – brown; copepods – pink; isopods – navy blue; amphipods – green; decapods – red; branchiopod – fluorescent blue. The tree was built using IQ-Tree with 1000 UFBoot replications under the LG+F+G4 model of substitution according to BIC as selected by ModelFinder. The tree was visualized on FigTree v.1.4.4. The tree is unrooted but the root is drawn at the iGluR/IR25a/IR8a clade. Scale bar represents expected number of substitutions per site.

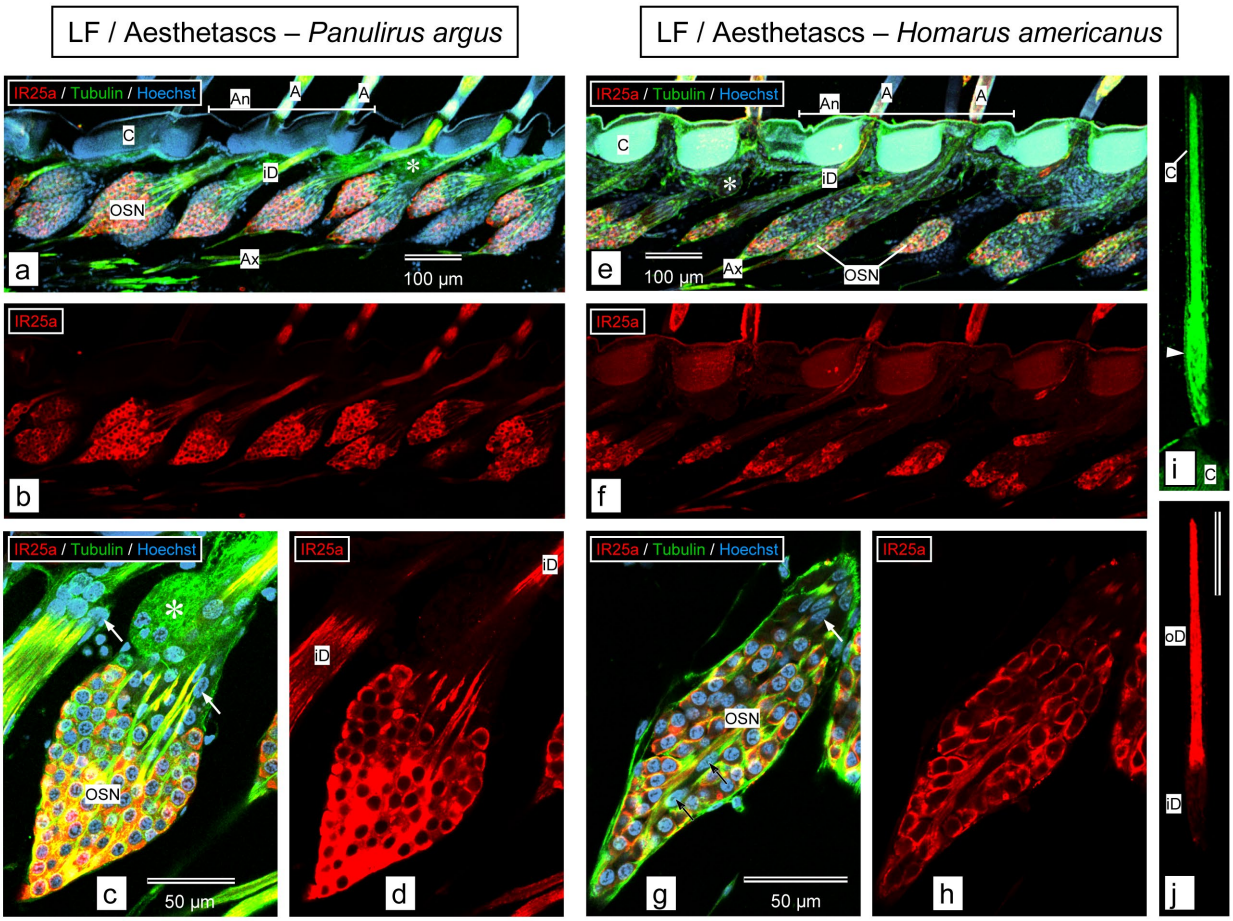

Figure 3.6 Immunolabeling with anti-HaIR25a in the aesthetasc-bearing tuft region of the lateral flagellum of the antennule.

(a – d) *Panulirus argus*. **(e – j)** *Homarus americanus*. **(a)**, **(b)**, **(e)**, **(f)** Sagittal sections through the medial plane of the tuft region of lateral flagellum labeled with anti-HaIR25a (red), anti-tubulin (green), and Hoechst 33258 (blue) at low magnification (maximum intensity projections of confocal image stacks). Scale bar in **(a)** also applies to **(b)**, scale bar in **(e)** also applies to **(f)**. **(a)** and **(b)** adapted from Kozma et al. (2018). **(a)**, **(e)** Overlay of all 3 confocal channels. **(b)**, **(f)** anti-HaIR25a channel. Overall organization of aesthetascs is similar in both species: two rows of aesthetasc setae (A) arise from the autofluorescent (blue in *P. argus*, blue and green in *H. americanus*) cuticle (C) of an annulus (An + horizontal bar). Each aesthetasc seta is associated with a large cluster of olfactory sensory neuron (OSN) somata which are distinctly labeled by anti-HaIR25a and labeled with moderate intensity by anti-tubulin. Bundles of inner dendritic segments (iD) arising at the apical pole of the OSN clusters are labeled with moderate intensity by both

antibodies. Bundles of axons (Ax) arising at the basal pole of OSN clusters are intensely labeled by anti-tubulin labeled with moderate intensity by anti-HaIR25a. Tegumental glands (asterisks) are located between bundles of inner dendritic segments. OSN clusters in *H. americanus* are more elongated than in *P. argus* and contain fewer OSNs. **(c)**, **(d)**, **(g)**, **(h)** Sagittal section through OSN clusters labeled with anti-HaIR25a (red), anti-tubulin (green), and Hoechst 33258 (blue) at high magnification (confocal images – maximum intensity projection of confocal image stacks with a total thickness of about 1 µm). Scale bar in **(c)** also applies to **(d)**, scale bar in **(g)** also applies to **(h)**. **(c)**, **(g)** Overlay of all 3 confocal channels. **(d)**, **(h)** anti-HaIR25a channel. The somata (OSN) of all OSNs identified by having almost spherical nuclei are distinctly labeled by anti-HaIR25a. Note that the overall shape of OSNs is close to spherical in *P. argus* but more elliptical in *H. americanus*. Somata of auxiliary cells (identified by having flat, elongated nuclei - arrows) are not labeled by anti-HaIR25a. In *H. americanus*, auxiliary cells are not only present at the apical pole of the OSN cluster (white arrows) but also in its center (black arrows). **(i)**, **(j)** Horizontal section through an aesthetasc of *H. americanus* labeled with anti-HaIR25a (red) and anti-tubulin (green) (confocal images of one optical section of 1 µm thickness). Scale bar in **(j)** (100 µm) also applies to **(i)**. Note that anti-tubulin non-specifically labeled cuticle (C) in addition to dendrites enclosed in the thin cuticular tube of the aesthetasc seta. The bulge at the bottom of the seta (arrowhead) indicates the transition region between inner dendritic segments (iD) and outer dendritic segments (oD). Note that labeling intensity of anti-HaIR25a is considerably higher in oD compared to iD.

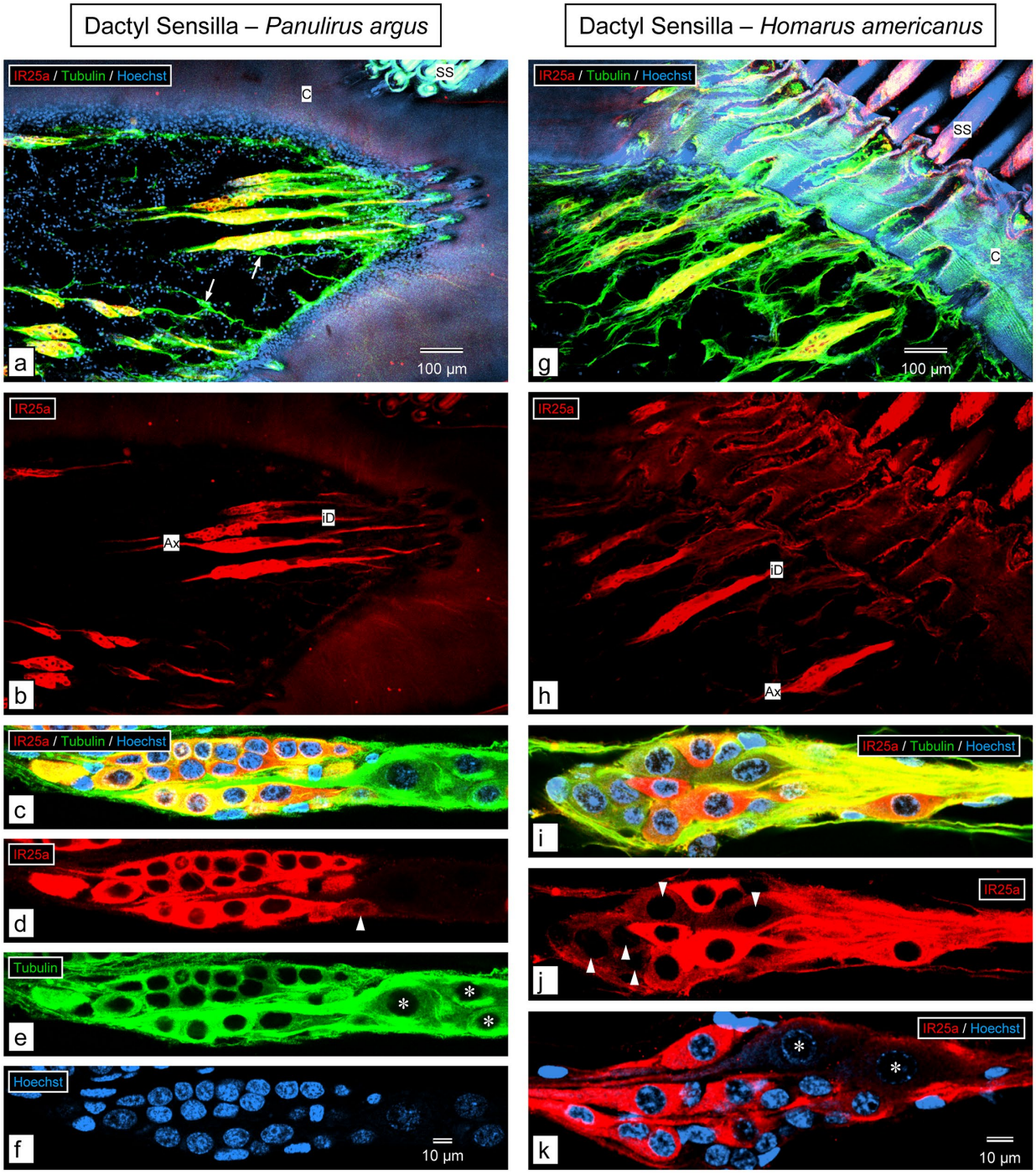

Figure 3.7 Immunolabeling with anti-HaIR25a in the walking leg dactyl.

(a – f) *Panulirus argus*. **(g – k)** *Homarus americanus*. **(a)**, **(b)**, **(g)**, **(h)** Sagittal sections through distal part (excluding the epicuticular cap) of the dactyl of the 3rd pereiopod labeled with anti-HaIR25a (red), anti-tubulin (green), and Hoechst 33258 (blue) at low magnification (maximum intensity projections of confocal image stacks that

are 5–10 μm thick). Scale bar in **(a)** also applies to **(b)**, scale bar in **(g)** also applies to **(h)**. **(a)**, **(g)** Overlay of all 3 confocal channels. **(b)**, **(h)** anti-HaIR25a channel. In both species, the numerically dominant sensilla are smooth setae (SS) organized into large, distinct groups. Each smooth seta is innervated by an elongated cluster of sensory neurons that is more than 200 μm long and about 50 μm in diameter and intensely labeled by anti-HaIR25a and anti-tubulin. Both antibodies label the somata of sensory neurons as well as their axons (Ax) and inner dendritic segments (iD). Note that in *P. argus*, but not in *H. americanus*, single bipolar sensory neurons labeled by anti-tubulin but not anti-HaIR25a (arrows) are interspersed between the double-labeled clusters of sensory neurons. **(c – f)**, **(i – k)** Two examples of clusters of sensory neurons labeled with anti-HaIR25a (red), anti-tubulin (green), and Hoechst 33258 (blue) at high magnification (maximum intensity projections of two adjacent optical sections of 0.4 μm thickness); scale bar in **(f)** also applies to **(c – e)**; scale bar in **(k)** also applies to **(i)** and **(j)**. **(c)**, **(i)** Overlay of all three channels. **(d)**, **(j)** anti-HaIR25a channel. **(e)** anti-tubulin channel. **(f)** Hoechst channel. **(k)** Overlay of anti-HaIR25a and Hoechst channel.

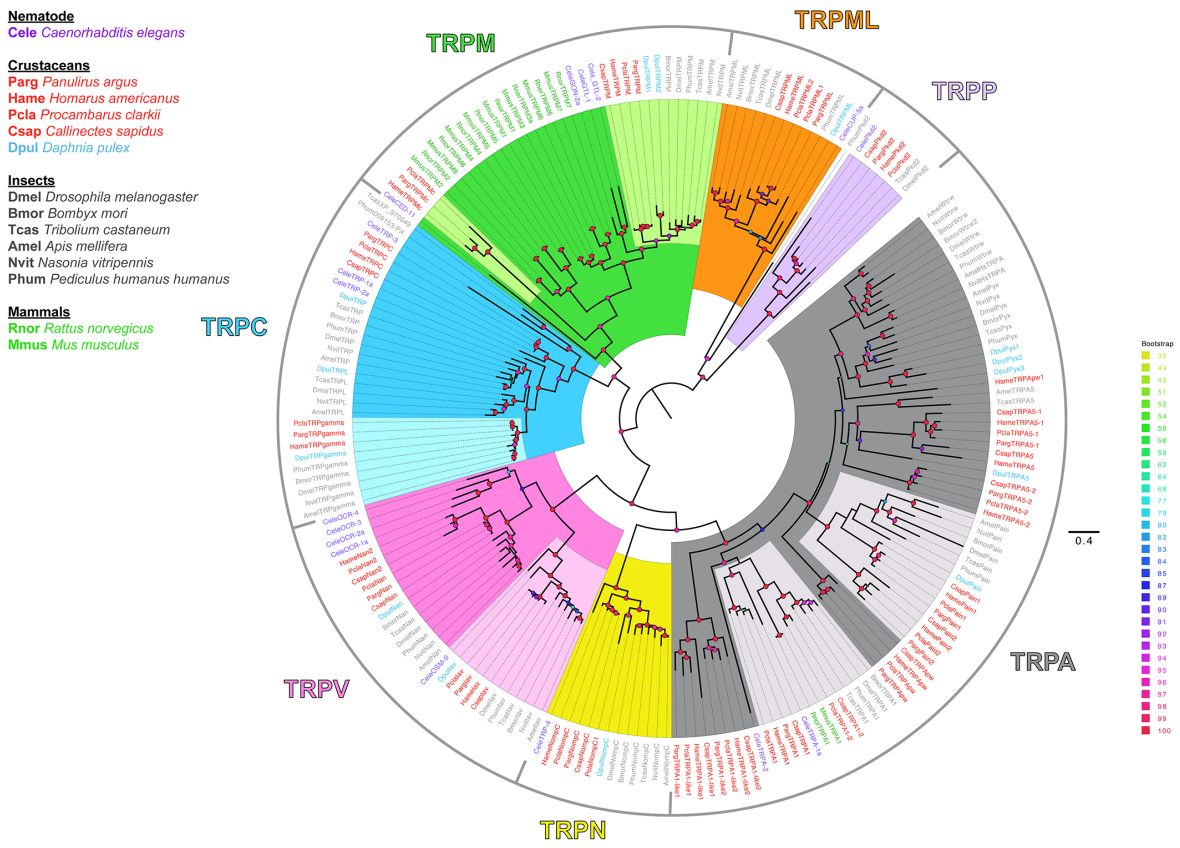

Figure 3.8 Phylogenetic tree of TRP channels across animals.

The maximum likelihood phylogenetic tree shows the conservation of TRP channel sequences from the transcriptomes of four decapod crustaceans with TRP channels from insects, nematodes, and mammals. Among crustaceans, decapods are in red and branchiopod in light blue. All four decapod crustaceans have several homologues to each subfamily of TRP channels. The tree was constructed on IQ-Tree with 1000 UFBoot replications under the LG+F+G4 model of substitution according to BIC, as determined by ModelFinder. Tree was visualized on FigTree v.1.4.4. Tree was unrooted but is drawn with the Group 2 subfamilies, TRPML and TRPP, clades as the root. Support for some inner nodes is low due to incomplete sequences and high divergence. Scale bar represents expected number of substitutions per site.

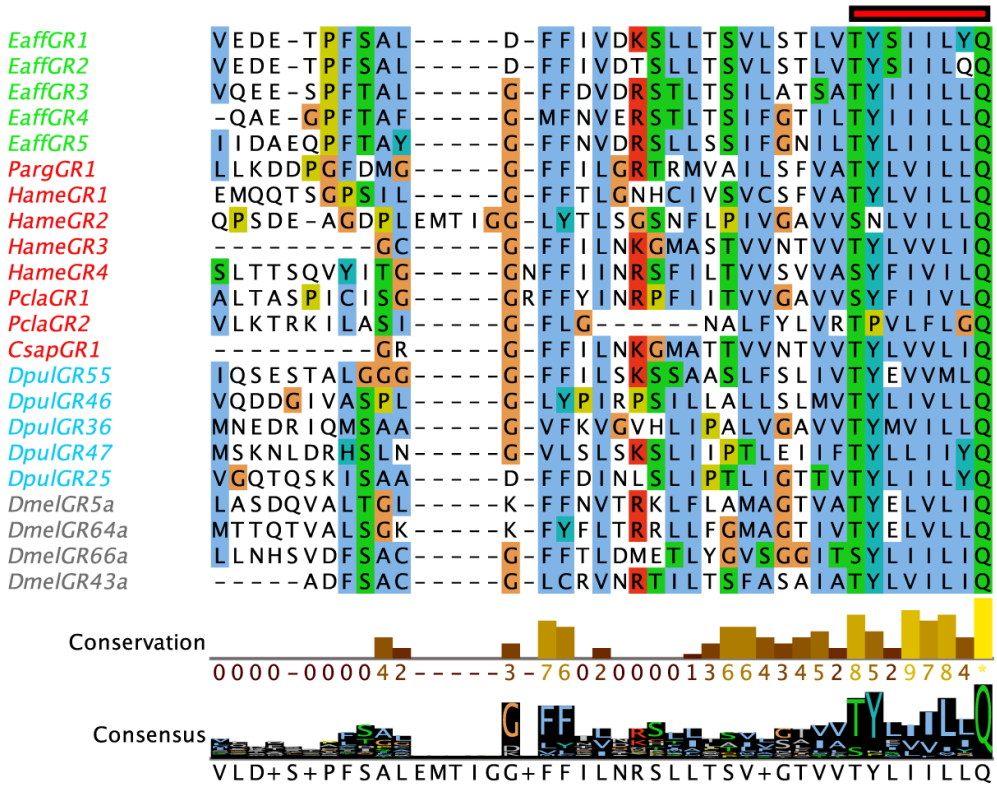

Figure 3.9 Multiple sequence alignment of GRL fragments in decapod crustaceans and GRs in arthropods.

Multiple sequence alignment shows the TM7 region of the sequences that have the highly conserved "TYxxxxxQF" motif (red bar). Sequences were aligned using MAFFT and visualized on Jalview. Decapod crustaceans are in red. Green – Eaff – *E. affinis*; Blue – Dpul – *D. pulex*; Grey – Dmel – *D. melanogaster*. Conservation of amino acids and the consensus histogram were annotated on Jalview. Clustal X color scheme was used to color residues.

Table 3.1 Number of predicted IRs and iGluRs in transcriptomes of four decapod crustacean species, based on either or both PF domains.

Species	PF00060	PF10613	Both
Panulirus argus	463	375	252
Homarus americanus	259	200	183
Procambarus clarkii	181	134	96
Callinectes sapidus*	253	198	184

Number of sequences that have the respective domain region represented in the columns for each of the four decapod crustaceans. "Both" shows the number of sequences that have both PF00060 and PF10613 domain regions. * indicates that transcriptome only has LF and dactyl tissue, while the others have LF, dactyl, and brain.

Table 3.2 Number of predicted IRs and iGluRs in transcriptomes from four species of decapod crustaceans.

Species	iGluRs	IRs			
		Co-receptor	Tuning IRs		
			Conserved	Species-Specific	Total
Panulirus argus	10	4	85	169	254
Homarus americanus	11	4	128	53	181
Procambarus clarkii	9	4	73	19	92
Callinectes sapidus*	6	4	87	99	186

Conservative estimate of number of IRs and iGluRs expressed in the transcriptomes of the four decapod crustaceans. iGluRs include NMDA and non-NMDA receptor sequences. Co-receptor IRs are IR25a, IR8a, IR76b, and IR93a. Conserved tuning IRs are variant tuning IR sequences that have homologues in other species. Most sequences accounted here have both domain regions representative of iGluRs and IRs. Some incomplete sequences that are homologous to conserved IRs in another species are also included (Supplemental S10-S12 Files: https://doi.org/10.1371/journal.pone.0230266). * indicates transcriptome generated from LF and dactyls only, while transcriptomes from the other species also include reads from brain tissue.

3.7 Supplemental Figure and Tables

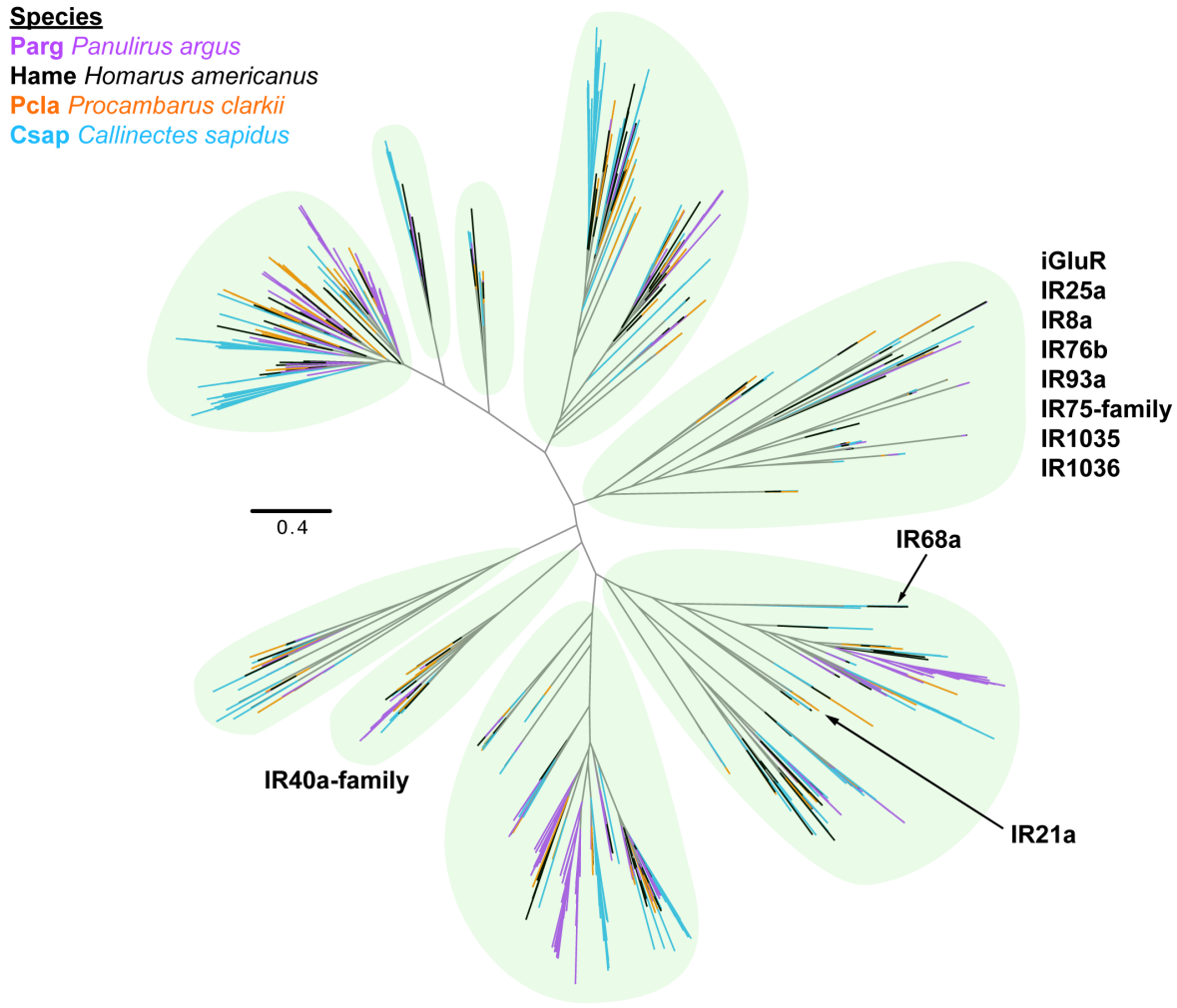

Figure 3.10 Radial tree configuration of phylogenetic analysis of iGluRs and IRs from decapod crustaceans (otherwise represented in Figure 3.4).

Table 3.3 Transcript and protein coding gene counts generated from EVG pipeline.

Species	Transcripts	Protein-Coding Genes	Percent Complete CDS
Panulirus argus	78,173	78,421	61.8%
Homarus americanus	66,199	66,471	60.4%
Procambarus clarkii	90,561	90,860	56.0%
Callinectes sapidus	59,752	60,080	48.7%

Table 3.4 BUSCO output.

BUSCO terms	***P. argus***	***H. americanus***	***P. clarkii***	***C. sapidus***
Complete BUSCOs (C)	978	975	989	1026
Complete and single-copy BUSCOs (S)	950	954	959	915
Complete and duplicated BUSCOs (D)	28	21	30	111
Fragmented BUSCOs (F)	11	11	10	14
Missing BUSCOs (M)	77	80	67	26
Total BUSCO groups searched	1066	1066	1066	1066

Table 3.5 Decapod iGluRs.

Species	NMDA	Non-NMDA
Panulirus argus	NMDAr1 NMDAr2a NMDAr2b NMDAr2d NMDAr3	11268664 11268665 73816 86731 83058
Homarus americanus	NMDAr1 NMDAr1b NMDAr2a NMDAr2c NMDAr2d NMDAr3	83058 86731 12031145 11268664 11268665
Procambarus clarkii	NMDAr1 NMDAr2a NMDAr2c NMDAr2d	83058 86731 11268664 11268665 11774099
Callinectes sapidus	NMDAr1 NMDAr2b NMDAr2c	73816 11268664 11268665

4 GENERAL DISCUSSION

Animals express a diversity of chemoreceptor proteins in their chemosensory systems. Animals express two broad types of chemoreceptor protein classes: 1) classes of chemoreceptor proteins with large families of genes, and 2) classes of receptor proteins that have multimodal sensitivity with smaller expansions in the number of genes (e.g. transient receptor potential channels, epithelial sodium channels). However, classes of chemoreceptor proteins that fall into these two types are not rigid across species and in fact are quite fluid in terms of number of genes expressed and sensory function. This means that classes of chemoreceptor proteins that can be categorized into type 1 in one phylum, may not be type 1 or even be expressed in another phylum. Also, the expansion in the number of genes of a type 1 class of chemoreceptor proteins can vary vastly depending on the organism in question. For example, chemosensory GPCRs are a primary class of proteins expressed across chordates with large expansions spanning to the order of thousands of genes in some species (Churcher and Taylor 2009, Churcher and Taylor 2011), but have smaller expansions and much more sporadic appearance among different phyla of protostomes where large gene family expansions of chemosensory GPCRs have only been detected in nematodes and molluscs so far (Thomas and Robertson 2008, Cummins et al. 2009a, Cummins et al. 2009b, Nordstrom et al. 2011). In the last two decades, extensive characterizations and functional analyses on chemosensory neurons of insects, particularly in the fruit fly *Drosophila melanogaster*, and transcriptome research in chelicerates and crustaceans revealed several classes of ionotropic receptors as primary chemoreceptor proteins in one of the largest phyla of animals – the arthropods (Benton et al. 2009, Peñalva-Arana et al. 2009, Croset et al. 2010, Benton 2015, Robertson 2015, Eyun et al. 2017, Vizueta et al. 2018, Robertson 2019, Kozma et al. 2020).

Classes of chemosensory ionotropic receptors with large family of genes include seven-transmembrane proteins known gustatory receptors (GR) and olfactory receptors (OR), and three-transmembrane proteins known as variant ionotropic receptors (IR). So far, there is limited evidence that GPCRs function as chemoreceptor proteins in arthropods; recently, non-visual opsins have been shown to mediate taste in *Drosophila* (Leung et al. 2020). Among arthropods, all subphyla do not express all chemosensory ionotropic receptor classes mentioned above. GR gene family expansions appear sporadically across chelicerates and crustaceans, whereas most insects have larger number of GR genes ranging from 20 to over 300 in some species of beetles (Benton 2015, Saina et al. 2015, Robertson 2019). Ionotropic ORs are an insect specific class of chemoreceptor proteins that have evolved from GRs and believed to be more recently evolved class of receptors compared to other chemosensory ionotropic receptors (Benton 2015, Robertson 2015). Variant IRs, which evolved from ionotropic glutamate receptors and still share their domain architecture, are the most commonly expressed class of chemoreceptor proteins across arthropods. While one IR (IR25a) is conserved across all protostomes, several other variant IRs have varying degrees of conservation ranging from 'species-specific' to 'arthropod-conserved' (Croset et al. 2010, Eyun et al. 2017, Vizueta et al. 2018, Kozma et al. 2020). Arthropods in general seem to express large gene families of variant IRs ranging from ~ 20 genes in hermit crabs (from transcriptome) to ~ 500 in the Caribbean spiny lobster (from transcriptomes) to ~ 900 in the German cockroach (from genome) (Groh et al. 2013, Groh-Lunow et al. 2014, Robertson et al. 2018). Each arthropod species expresses at least a few IRs that are unique to it, but also expresses many IRs that are conserved across species within an order of arthropods (Croset et al. 2010, Kozma et al. 2018, Kozma et al. 2020).

The second type of chemoreceptor proteins that have smaller number of genes (< 25) in their class and have multimodal sensitivity tend to be more consistent in their numbers and presence across animals. Transient receptor potential (TRP) channels and epithelial sodium channels (ENaC) are prominent examples of type 2 chemoreceptor proteins (Venkatachalam et al. 2014, Elkhatib et al. 2019). These classes of ionotropic receptor proteins can have sensitivity to light, temperature, pressure, protons, and chemical stimuli. They are not dedicated chemical sensors.

Identifying the expression of genes for this diversity of chemoreceptor proteins in decapod crustaceans is an important first step in unraveling the molecular mechanisms of their chemical sensing. The coding of olfaction and distributed chemoreception in decapod crustaceans cannot be discerned without knowing which chemoreceptor proteins are expressed in their chemosensory neurons. Knowing the expression patterns of chemoreceptor proteins in individual OSNs and CSNs would allow us to figure out the wiring logic of sensory input from chemosensory neurons to ganglia of the central nervous system such as the olfactory lobe in decapod crustaceans. My dissertation research lead to the creation of an extensive list of candidate chemoreceptor proteins for four decapod crustaceans. My research revealed the diversity and number of chemoreceptor proteins expressed by two chemosensory organs of these decapods, and through phylogenetic classification I described their similarities and differences from other arthropods (Kozma et al. 2018, Kozma et al. 2020). I found that the decapod crustaceans, Caribbean spiny lobster *Panulirus argus*, American clawed lobster *Homarus americanus*, red swamp crayfish *Procambarus clarkii*, and blue crab *Callinectes sapidus* all express genes for variant IRs, TRP channels, GPCRs, ENaCs and GRLs in two of their main chemosensory organs: lateral flagella of the antennules and dactyls of walking legs.

4.1 Variant IRs as primary chemoreceptors proteins of decapod crustaceans

Variant IRs including co-receptor IRs and tuning-IRs are expressed in chemosensory tissues of four species of decapod crustaceans (Corey et al. 2013, Groh et al. 2013, Groh-Lunow et al. 2014, Kozma et al. 2018, Kozma et al. 2020). IR25a is a variant ionotropic glutamate (IR) receptor protein classified as a co-receptor that consistently appears across all classes of protostomes. IR25a has been detected in the chemosensory tissues and neurons of several arthropod phyla (Hollins et al. 2003, Schmidt et al. 2006, Benton et al. 2009, Tadesse et al. 2011, Corey et al. 2013, Groh-Lunow et al. 2014, Kozma et al. 2018, Sánchez-Alcañiz et al. 2018, Kozma et al. 2020). The expression pattern and function of this co-receptor protein has been most extensively studied in insects, particularly in sensory neurons of fruit flies and mosquitoes. In these insects, with few exceptions, IR25a is essential for the assembly and function of the heterotetrameric variant IRs as chemoreceptor proteins and in a few cases as thermo- and hygroreceptor proteins (van Giesen and Garrity 2017). Other co-receptor IRs include IR8a, IR93a, and IR76b, which are all expressed in the LF and dactyl tissues of four decapod crustaceans. Immunocytochemistry (ICC) in *P. argus* and *H. americanus* revealed that IR25a is expressed in nearly all olfactory sensory neurons (OSNs) and most chemosensory neurons (CSNs) of the lateral flagella of antennules (LF), walking leg dactyls, and 2nd antenna (Kozma et al. 2018, Kozma et al. 2020). However, ICC revealed that *P. argus* and *H. americanus* have a substantial difference of IR25a expression in central nervous system (CNS). Axon terminals of OSNs in the olfactory lobe of *H. americanus* are distinctly and positively immunolabeled for IR25a, whereas in *P. argus* only a set of large orphan cells, found in the lateral division of the antennular nerve and the axon sorting

zone proximal to the olfactory lobe, are exclusively immunolabeled for IR25a expression indicating different co-receptor expression patterns in the CNS of these two decapod crustaceans. In recent work using single transcriptomics, IR25a and IR93a were detected with high expression in 7 single OSNs of P. argus (Kozma et al., 2020 submitted). In fact, these two co-receptor IRs were amongst the top 20 most highly expressed genes in each of the seven single-OSN transcriptomes. While the other two co-receptor IRs IR8a and IR76b are also expressed in LF of *P. argus*, there was no evidence of their presence in single-OSN transcriptomes of *P. argus*. More single cell transcriptome work on OSNs and CSNs is necessary to make broader conclusions, but initial findings suggest that IR25a and IR93a are the co-receptor IRs for OSNs, with IR8a and IR76b likely to be expressed in the CSNs of the distributed chemoreceptor proteins. These findings also show a departure from expression patterns of co-receptor IRs in insects; IR25a is the only variant IR, with a few notable exceptions, that is ubiquitously expressed in all IR-expressing OSNs and gustatory sensory neurons (GSNs) of *Drosophila*, and IR93a is only co-expressed in the antenna and in a small subset of sensory neurons (Benton et al. 2009, Enjin et al. 2016, Knecht et al. 2016, Frank et al. 2017, Knecht et al. 2017, Sánchez-Alcañiz et al. 2018). IR76b is expressed in all chemosensory organs of Drosophila but is most widely co-expressed with IR25a in GSNs. IR8a is broadly expressed in *Drosophila* OSNs along with overlapping IR25a co-expression in some neurons around the sacculus on the antenna (Benton et al. 2009).

Tuning IRs form the other important part of the variant IR story. Tuning IRs are necessary for conferring chemical specificity to variant IR chemoreceptor proteins by forming ionotropic heterotetrameric receptor channels in combination with co-receptor IRs. 2 – 3 tuning IRs are typically expressed in a chemosensory neuron, and the

combination of which tuning IRs are expressed in combination with co-receptor IRs confers the ability of a functional variant IR chemoreceptor to detect specific molecules. *Drosophila* and indeed insects in general have large expansions of other chemoreceptor proteins besides variant IRs such as ORs that are more broadly tuned to odorants (Hallem and Carlson 2006, Silbering et al. 2011). In *Drosophila*, IR expressing sensory neurons often act as 'specialist' neurons that are fine tuned to respond to a very narrow range of molecules and in some cases even single modalities, for example, ammonia and amines detection by IR92a expressing OSNs (Silbering et al. 2011, Min et al. 2013). Another well-described example of specialization of IR expressing sensory neurons is in the case hygro-sensory and thermo-sensory neurons in antenna of *Drosophila*; co-expression of IR25a, IR93a, and IR40a mediate hygro-sensitivity in dry conditions, whereas the combinatorial co-expression of IR25a, IR93a, and IR68a mediate hygro-sensitivity in moist conditions, and co-expression of IR25a, IR93a, and IR21a mediates sensitivity to cool temperatures (Enjin et al. 2016, Knecht et al. 2016, Frank et al. 2017, Knecht et al. 2017). Therefore, the difference of one tuning IR can change the functionality and sensitivity of IR-expressing sensory neurons to external stimuli.

Hundreds of tuning IRs are expressed by decapod crustaceans in their LF and dactyls. It is possible that these numbers of tuning IRs might be double or more with the inclusion of more transcriptomes of decapod chemosensory tissues such as mouth parts or tail fans. Sequencing their genomes will, of course, provide us with a more accurate assessment of the number of variant IRs genes in decapods. Single-cell transcriptomics analyses also detected expression of tuning IRs in OSNs of *P. argus* (Kozma et al., submitted). With ionotropic ORs and GRs (mostly) missing in decapod crustaceans, variant IRs are the most prominent candidate chemoreceptor proteins in OSNs and CSNs.

While IR-expressing sensory neurons in insects are 'specialists,' and OR-expressing neurons are more broadly tuned to a range of chemical stimuli, it very likely that IR-expressing sensory neurons in decapod crustaceans, with their large IR gene families, are capable of being both specialists and broadly tuned to various chemical stimuli. The presence of several 'decapod-conserved' or 'arthropod-conserved' tuning IRs and 'species-specific' tuning IRs lends support to the idea that some tuning IRs may function as generalist chemoreceptor proteins, while other 'species-specific' IRs may be involved in detecting conspecifics' individual chemical signatures, chemical signals and cues such as sex pheromones, or pheromones that reveal molt-stage or social status.

Single cell transcriptomics and proteomics combined with physiological experiments are needed to answer more refined questions about chemoreceptor protein expression in OSNs and CSNs of decapod crustaceans. Tissue level transcriptome data go only so far, especially when gene expression does not necessarily mean translation to functional proteins. The preliminary single-cell transcriptome work on 7 OSNs in *P. argus* starts to address the question, 'which variant IRs are expressed in OSNs vs. CSNs?' Comparing differential gene expression between the LF and dactyl gives us some clues as to which IRs are expressed in OSNs and CSNs. However, this is not a clean comparison of data as the LF houses both OSNs and CSNs. We could infer that variant IRs that are exclusively expressed in the dactyl tissue may be CSN-specific variant IRs – "Would that it were so simple!" (Coen and Coen 2016). Almost all IRs are expressed in LF and less than 5 IRs are exclusive to the dactyl in each of the four decapod crustaceans studied. There was also no discernible relationship between the phylogeny of IRs and organ-specific expression of variant IRs. Although we only examined and compared LF and dactyl chemosensory organs of decapod crustaceans, such relationships were not

observed within *Drosophila* either (Sánchez-Alcañiz et al. 2018). Any functional role of phylogenetic proximity among IRs of decapod crustaceans or insects is still unknown.

Although there is currently no physiological evidence on whether variant IRs are truly chemoreceptor proteins for decapod crustaceans, high gene and protein expression of variant IRs within OSNs and CSNs makes them the most likely candidates for primary chemoreceptor proteins of decapod crustaceans.

4.2 Other candidate chemoreceptor proteins

Chemosensory roles of 'non-typical' chemoreceptor proteins in arthropods is an emerging avenue of study – although it would be more accurate to say a 'revisited' avenue of study. Canonically, metabotropic G-protein coupled receptors (GPCR) were expected to be chemoreceptor proteins across all animals. In the last two decades, extensive research on arthropods changed this view dramatically. Arthropods are one of the largest phyla of animals where putative ionotropic receptors and not metabotropic receptors are the predominant chemoreceptor proteins. This should not ignore the fact that metabotropic receptor machinery is found in chemosensory neurons of arthropods including in decapod crustaceans (McClintock et al. 2006), hawkmoths (Gawalek and Stengl 2018), and *Drosophila*. Recent findings in GSNs of *Drosophila* showed that non-visual opsins, which are GPCRs of the Class A – rhodopsin family, are involved in bitter taste sensing with the TRP channel TRPA1 involved in the signal transduction pathway (Leung et al. 2020). Over 100 GPCRs were detected in the LF and dactyls of all four decapod crustaceans (Rump, Kozma, Ngo-Vu, Derby – in prep). Many of these GPCRs were detected at the tissue level and phylogenetically classified into classes of receptors that are typically associated with neuromodulation. However, we detected expression of

an opsin and a small number of GPCRs in single OSN transcriptomes of *P. argus*. This small subset of GPCRs would be prime targets for future functional studies to determine their role in OSNs and chemical sensing, or in other sensory roles, such as opsin-expressing OSNs may detect light.

TRP channels while considered as ionotropic receptors are often involved in signal transduction pathways of metabotropic receptors, particularly GPCRs. While TRP channels are receptors to many sensory modalities such as light, temperature, pressure, and chemicals, their exact mechanism of activation by these modalities are still unclear. In *Drosophila*, it is suspected that tissue damage caused by the above-mentioned stimuli prompts the release of reactive oxygen species (ROS) /H_2O_2 and reactive nitrogen species (RNS) which trigger activation of TRP channels (Guntur et al. 2015, Arenas et al. 2017). With decapod crustaceans expressing TRP channels in both LF and dactyl, and in single OSN transcriptomes of *P. argus*, such a mechanism of activation could have been adapted by chemosensory neurons to detect external presence of ROS/RNS/H_2O_2, all of which are commonly found in their environment or in the defensive ink released by their prey, such as *Aplysia californica* (Aggio and Derby 2008, Aggio et al. 2012). Whether they act directly as chemoreceptor proteins or as part of signal transduction pathways, TRP channels present themselves as candidates to be investigated in future studies of chemical sensing in decapod crustaceans. In a similar vein, another class of multimodal sensors, degenerin/epithelial sodium channels (DEG/ENaC), has been shown to act as both ionotropic receptors and as part of signal transduction pathways in some OR- and IR-expressing OSNs of *Drosophila* where they amplify olfactory responses (Ng et al. 2019). ENaCs are not expressed in single OSN transcriptomes from *P. argus* (Kozma et al., submitted) but are detected in LF and dactyl tissues.

Given the presence of GPCRs, TRP channels, and ENaCs in two chemosensory organs of four decapod crustaceans, it is worth studying the cellular expression pattern and function of these receptor classes to assess their role in OSNs, CSNs, and chemosensory pathways of decapod crustaceans.

4.3 Model of chemosensory neurons in decapod crustaceans

My dissertation research combined with recent single cell transcriptomics of olfactory sensory neurons in *P. argus* begins to draw a picture of the molecular machinery used by decapod crustaceans for chemical sensing. High gene expression detected through transcriptomics along with evidence of translated protein through immunocytochemistry of variant IRs in both olfactory sensory neurons of the olfactory system and chemosensory neurons of the distributed chemoreceptor system suggests that variant IRs are the primary chemoreceptor proteins of the decapod crustaceans, *Panulirus argus*, *Homarus americanus*, *Procambarus clarkii*, and *Callinectes sapidus*.

There are still many questions to be answered, as is the case of all studies of natural phenomena. Some of these questions are: How many variant IR genes are in the genome of a given decapod? What is the combinatorial expression pattern of variant IRs in individual OSNs and CSNs? What is the governing principle behind the expression of a particular variant IR in sensory neurons? Given the multimodal sensory nature of variant IRs, are OSNs and CSNs in decapod crustaceans sensitive to modalities beyond chemicals? What is the function of GPCRs, TRP channels, and ENaCs in OSNs and CSNs? Are all chemoreceptor protein classes capable of multimodal sensing across animals? Single cell transcriptomics of chemosensory tissues in combination with proteomics would aid in answering many of these questions. Integrating physiological analyses with

-omics data is the current big challenge in decapod crustaceans. A number of technologies such as RNAi and CRISPR-CAS9 systems are ripe for use. Sequencing the genomes of large decapod crustaceans such as *Panulirus argus*, *Homarus americanus*, *Procambarus clarkii*, and *Callinectes sapidus* for which there is much behavioral and physiological data will enable carcinologists to enter the genomic revolution. To quote Science Officer Spock, "Logical!" (Roddenberry 1966).

5 REFERENCES

Abramova, A., M. Alm Rosenblad, A. Blomberg and T. A. Larsson (2019). "Sensory receptor repertoire in cyprid antennules of the barnacle *Balanus improvisus*." PLoS ONE **14**(5): e0216294.

Abuin, L., B. Bargeton, M. H. Ulbrich, E. Y. Isacoff, S. Kellenberger and R. Benton (2011). "Functional architecture of olfactory ionotropic glutamate receptors." Neuron **69**(1): 44-60.

Abuin, L., L. L. Prieto-Godino, H. Pan, C. Gutierrez, L. Huang, R. Jin and R. Benton (2019). "In vivo assembly and trafficking of olfactory Ionotropic Receptors." BMC Biol **17**(1): 34.

Ache, B. W. (2002). Crustaceans as animal models for olfactory research. Crustacean experimental systems in neurobiology. K. Wiese. Berlin, Heidelberg, Springer**:** 189-199.

Aggio, J. F. and C. D. Derby (2008). "Hydrogen peroxide and other components in the ink of sea hares are chemical defenses against predatory spiny lobsters acting through non-antennular chemoreceptors." J Exp Mar Biol Ecol **363**(1-2): 28-34.

Aggio, J. F., R. Tieu, A. Wei and C. D. Derby (2012). "Oesophageal chemoreceptors of blue crabs, *Callinectes sapidus*, sense chemical deterrents and can block ingestion of food." J Exp Biol **215**(Pt 10): 1700-1710.

Ahyong, S. T., J. K. Lowry, M. Alonso, R. N. Bamber, G. A. Boxshall, P. Castro, S. Gerken, G. S. Karaman, G. W. Goy, D. S. Jones, K. Meland, D. C. Rogers and J. Svavarsson (2011). Animal biodiversity: an outline of higher-level classification and survey of taxonomic richness. Zootaxa. Zhang Z-Q. Auckland, New Zealand, Magnolia Press. **3148:** 165-191.

Albertin, C. B., O. Simakov, T. Mitros, Z. Y. Wang, J. R. Pungor, E. Edsinger-Gonzales, S. Brenner, C. W. Ragsdale and D. S. Rokhsar (2015). "The octopus genome and the evolution of cephalopod neural and morphological novelties." Nature **524**(7564): 220-224.

Arenas, O. M., E. E. Zaharieva, A. Para, C. Vasquez-Doorman, C. P. Petersen and M. Gallio (2017). "Activation of planarian TRPA1 by reactive oxygen species reveals a conserved mechanism for animal nociception." Nat Neurosci **20**(12): 1686-1693.

Atema, J. (2018). "Opening the chemosensory world of the lobster, *Homarus americanus*." Bull Mar Sci **94**(3): 479-516.

Atema, J. and R. Voigt (1995). Behavior and sensory biology. Biology of the lobster: *Homarus americanus*. J. R. Factor. San Diego, CA, Academic Press**:** 313-348.

Bandell, M., L. J. Macpherson and A. Patapoutian (2007). "From chills to chilis: mechanisms for thermosensation and chemesthesis via thermoTRPs." Curr Opin Neurobiol **17**(4): 490-497.

Barnea, G., S. O'Donnell, F. Mancia, X. Sun, A. Nemes, M. Mendelsohn and R. Axel (2004). "Odorant receptors on axon termini in the brain." Science **304**(5676): 1468.

Bauer, R. T. (2010). Chemical communication in decapod shrimps: the influence of mating and social systems on the relative importance of olfactory and contact pheromones. Chemical Communication in Crustaceans. T. Breithaupt and M. Thiel. New York, NY, Springer**:** 277-296.

Belanger, R. M. and P. A. Moore (2006). "The use of the major chelae by reproductive male crayfish (*Orconectes rusticus*) for discrimination of female odours." Behaviour **143**: 713-731.

Beltz, B. S., K. Kordas, M. M. Lee, J. B. Long, J. L. Benton and D. C. Sandeman (2003). "Ecological, evolutionary, and functional correlates of sensilla number and glomerular density in the olfactory system of decapod crustaceans." J Comp Neurol **455**(2): 260-269.

Benton, R. (2015). "Multigene family evolution: perspectives from insect chemoreceptors." Trends Ecol Evol **30**(10): 590-600.

Benton, R., S. Sachse, S. W. Michnick and L. B. Vosshall (2006). "Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors *in vivo*." PLoS Biol **4**(2): e20.

Benton, R., K. S. Vannice, C. Gomez-Diaz and L. B. Vosshall (2009). "Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*." Cell **136**(1): 149-162.

Blaustein, D. N., R. B. Simmons, M. F. Burgess, C. D. Derby, M. Nishikawa and K. S. Olson (1993). "Ultrastructural localization of 5'AMP odorant receptor sites on the dendrites of olfactory receptor neurons of the spiny lobster." J Neurosci **13**(7): 2821-2828.

Bobkov, Y. V. and B. W. Ache (2005). "Pharmacological properties and functional role of a TRP-related ion channel in lobster olfactory receptor neurons." J Neurophysiol **93**(3): 1372-1380.

Bobkov, Y. V., A. Pezier, E. A. Corey and B. W. Ache (2010). "Phosphatidylinositol 4,5-bisphosphate-dependent regulation of the output in lobster olfactory receptor neurons." J Exp Biol **213**(Pt 9): 1417-1424.

Boekhoff, I., W. C. Michel, H. Breer and B. W. Ache (1994). "Single odors differentially stimulate dual second messenger pathways in lobster olfactory receptor cells." J Neurosci **14**(5 Pt 2): 3304-3309.

Bolger, A. M., M. Lohse and B. Usadel (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data." Bioinformatics **30**(15): 2114-2120.

Brand, P., H. M. Robertson, W. Lin, R. Pothula, W. E. Klingeman, J. L. Jurat-Fuentes and B. R. Johnson (2018). "The origin of the odorant receptor gene family in insects." eLife **7**.

Bräunig, P. (1982). "The peripheral and central nervous organization of the locust coxo-trochanteral joint." J Neurobiol **13**(5): 413-433.

Bräunig, P. and R. Hustert (1980). "Proprioceptors with central cell bodies in insects." Nature **283**(5749): 768-770.

Breithaupt, T. (2001). "Fan organs of crayfish enhance chemical information flow." Biol Bull **200**(2): 150-154.

Breithaupt, T. and M. Thiel (2010). Chemical communication in crustaceans. New York, NY, Springer.

Brochtrup, A. and T. Hummel (2011). "Olfactory map formation in the *Drosophila* brain: genetic specificity and neuronal variability." Curr Opin Neurobiol **21**(1): 85-92.

Buck, L. and R. Axel (1991). "A novel multigene family may encode odorant receptors: a molecular basis for odor recognition." Cell **65**(1): 175-187.

Cameron, P., M. Hiroi, J. Ngai and K. Scott (2010). "The molecular basis for water taste in *Drosophila*." Nature **465**(7294): 91-95.

Cate, H. S. and C. D. Derby (2001). "Morphology and distribution of setae on the antennules of the Caribbean spiny lobster *Panulirus argus* reveal new types of bimodal chemo-mechanosensilla." Cell Tissue Res **304**(3): 439-454.

Cate, H. S. and C. D. Derby (2002). "Ultrastructure and physiology of the hooded sensillum, a bimodal chemo-mechanosensillum of lobsters." J Comp Neurol **442**(4): 293-307.

Chen, C., E. Buhl, M. Xu, V. Croset, J. S. Rees, K. S. Lilley, R. Benton, J. J. Hodge and R. Stanewsky (2015). "*Drosophila* Ionotropic Receptor 25a mediates circadian clock resetting by temperature." Nature **527**(7579): 516-520.

Chen, Z., Q. Wang and Z. Wang (2010). "The amiloride-sensitive epithelial Na+ channel PPK28 is essential for *Drosophila* gustatory water reception." J Neurosci **30**(18): 6247-6252.

Chipman, A. D., D. E. Ferrier, C. Brena, J. Qu, D. S. Hughes, R. Schroder, M. Torres-Oliva, N. Znassi, H. Jiang, F. C. Almeida, C. R. Alonso, Z. Apostolou, P. Aqrawi, W. Arthur, J. C. Barna, K. P. Blankenburg, D. Brites, S. Capella-Gutierrez, M. Coyle, P. K. Dearden, L. Du Pasquier, E. J. Duncan, D. Ebert, C. Eibner, G. Erikson, P. D. Evans, C. G. Extavour, L. Francisco, T. Gabaldon, W. J. Gillis, E. A. Goodwin-Horn, J. E. Green, S. Griffiths-Jones, C. J. Grimmelikhuijzen, S. Gubbala, R. Guigo, Y. Han, F. Hauser, P. Havlak, L. Hayden, S. Helbing, M. Holder, J. H. Hui, J. P. Hunn, V. S. Hunnekuhl, L. Jackson, M. Javaid, S. N. Jhangiani, F. M. Jiggins, T. E. Jones, T. S. Kaiser, D. Kalra, N. J. Kenny, V. Korchina,

C. L. Kovar, F. B. Kraus, F. Lapraz, S. L. Lee, J. Lv, C. Mandapat, G. Manning, M. Mariotti, R. Mata, T. Mathew, T. Neumann, I. Newsham, D. N. Ngo, M. Ninova, G. Okwuonu, F. Ongeri, W. J. Palmer, S. Patil, P. Patraquim, C. Pham, L. L. Pu, N. H. Putman, C. Rabouille, O. M. Ramos, A. C. Rhodes, H. E. Robertson, H. M. Robertson, M. Ronshaugen, J. Rozas, N. Saada, A. Sanchez-Gracia, S. E. Scherer, A. M. Schurko, K. W. Siggens, D. Simmons, A. Stief, E. Stolle, M. J. Telford, K. Tessmar-Raible, R. Thornton, M. van der Zee, A. von Haeseler, J. M. Williams, J. H. Willis, Y. Wu, X. Zou, D. Lawson, D. M. Muzny, K. C. Worley, R. A. Gibbs, M. Akam and S. Richards (2014). "The first myriapod genome sequence reveals conservative arthropod gene content and genome organisation in the centipede *Strigamia maritima*." PLoS Biol **12**(11): e1002005.

Churcher, A. M. and J. S. Taylor (2009). "Amphioxus (*Branchiostoma floridae*) has orthologs of vertebrate odorant receptors." BMC Evol Biol **9**: 242.

Churcher, A. M. and J. S. Taylor (2011). "The antiquity of chordate odorant receptors is revealed by the discovery of orthologs in the cnidarian *Nematostella vectensis*." Genome Biol Evol **3**: 36-43.

Coen, J. and E. Coen (2016). Hail, Caesar! [Motion Picture]. United States, Universal Pictures.

Colbert, H. A., T. L. Smith and C. I. Bargmann (1997). "OSM-9, a novel protein with structural similarity to channels, is required for olfaction, mechanosensation, and olfactory adaptation in *Caenorhabditis elegans*." J Neurosci **17**(21): 8259-8269.

Corey, E. A., Y. Bobkov, K. Ukhanov and B. W. Ache (2013). "Ionotropic crustacean olfactory receptors." PLoS ONE **8**(4): e60551.

Croset, V., R. Rytz, S. F. Cummins, A. Budd, D. Brawand, H. Kaessmann, T. J. Gibson and R. Benton (2010). "Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction." PLoS Genet **6**(8): e1001064.

Croset, V., M. Schleyer, J. R. Arguello, B. Gerber and R. Benton (2016). "A molecular and neuronal basis for amino acid sensing in the *Drosophila* larva." Sci Rep **6**(1): 34871.

Cummins, S. F., D. Erpenbeck, Z. Zou, C. Claudianos, L. L. Moroz, G. T. Nagle and B. M. Degnan (2009a). "Candidate chemoreceptor subfamilies differentially expressed in the chemosensory organs of the mollusc *Aplysia*." BMC Biol **7**: 28.

Cummins, S. F., L. Leblanc, B. M. Degnan and G. T. Nagle (2009b). "Molecular identification of candidate chemoreceptor genes and signal transduction components in the sensory epithelium of *Aplysia*." J Exp Biol **212**(Pt 13): 2037-2044.

Derby, C. D. (2000). "Learning from spiny lobsters about chemosensory coding of mixtures." Physiol Behav **69**(1-2): 203-209.

Derby, C. D. and D. N. Blaustein (1988). "Morphological and physiological characterization of individual olfactory interneurons connecting the brain and eyestalk ganglia of the crayfish." J Comp Physiol A **163**(6): 777-794.

Derby, C. D., H. S. Cate and L. R. Gentilcore (1997). "Perireception in olfaction: molecular mass sieving by aesthetasc sensillar cuticle determines odorant access to receptor sites in the Caribbean spiny lobster *Panulirus argus*." J Exp Biol **200**(Pt 15): 2073-2081.

Derby, C. D., M. T. Kozma, A. Senatore and M. Schmidt (2016). "Molecular mechanisms of reception and perireception in crustacean chemoreception: a comparative review." Chem Senses **41**(5): 381-398.

Derby, C. D. and P. W. Sorensen (2008). "Neural processing, perception, and behavioral responses to natural chemical stimuli by fish and crustaceans." Journal of chemical ecology **34**(7): 898-914.

Derby, C. D. and M. J. Weissburg (2014). The chemical senses and chemosensory ecology of crustaceans. The natural history of the Crustacea. C. Derby and M. Thiel. New York, NY, Oxford University Press**:** 263-292.

Egekwu, N., D. E. Sonenshine, B. W. Bissinger and R. M. Roe (2014). "Transcriptome of the female synganglion of the black-legged tick *Ixodes scapularis* (Acari: Ixodidae) with comparison between Illumina and 454 systems." PLoS ONE **9**(7): e102667.

Elkhatib, W., C. L. Smith and A. Senatore (2019). "A Na(+) leak channel cloned from Trichoplax adhaerens extends extracellular pH and Ca(2+) sensing for the DEG/ENaC family close to the base of Metazoa." J Biol Chem **294**(44): 16320-16336.

Elmore, T. and D. P. Smith (2001). "Putative *Drosophila* odor receptor OR43b localizes to dendrites of olfactory neurons." Insect Biochem Mol Biol **31**(8): 791-798.

Enjin, A., E. E. Zaharieva, D. D. Frank, S. Mansourian, G. S. Suh, M. Gallio and M. C. Stensmyr (2016). "Humidity sensing in *Drosophila*." Curr Biol **26**(10): 1352-1358.

Evans, P. D., B. R. Talamo and E. A. Kravitz (1975). "Octopamine neurons: morphology, release of octopamine and possible physiological role." Brain Res **90**(2): 340-347.

Eyun, S. I., H. Y. Soh, M. Posavi, J. B. Munro, D. S. T. Hughes, S. C. Murali, J. Qu, S. Dugan, S. L. Lee, H. Chao, H. Dinh, Y. Han, H. Doddapaneni, K. C. Worley, D. M. Muzny, E. O. Park, J. C. Silva, R. A. Gibbs, S. Richards and C. E. Lee (2017). "Evolutionary history of chemosensory-related gene families across the Arthropoda." Mol Biol Evol **34**(8): 1838-1862.

Fedotov, V. P. (2009). "Systems of chemoperception in decapod crayfish." J Evol Biochem Physiol **45**(1): 3-24.

Finn, R. D., T. K. Attwood, P. C. Babbitt, A. Bateman, P. Bork, A. J. Bridge, H. Y. Chang, Z. Dosztanyi, S. El-Gebali, M. Fraser, J. Gough, D. Haft, G. L. Holliday, H. Huang, X.

Huang, I. Letunic, R. Lopez, S. Lu, A. Marchler-Bauer, H. Mi, J. Mistry, D. A. Natale, M. Necci, G. Nuka, C. A. Orengo, Y. Park, S. Pesseat, D. Piovesan, S. C. Potter, N. D. Rawlings, N. Redaschi, L. Richardson, C. Rivoire, A. Sangrador-Vegas, C. Sigrist, I. Sillitoe, B. Smithers, S. Squizzato, G. Sutton, N. Thanki, P. D. Thomas, S. C. Tosatto, C. H. Wu, I. Xenarios, L. S. Yeh, S. Y. Young and A. L. Mitchell (2017). "InterPro in 2017-beyond protein family and domain annotations." Nucleic Acids Res **45**(D1): D190-D199.

Finn, R. D., P. Coggill, R. Y. Eberhardt, S. R. Eddy, J. Mistry, A. L. Mitchell, S. C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, G. A. Salazar, J. Tate and A. Bateman (2016). "The Pfam protein families database: towards a more sustainable future." Nucleic Acids Res **44**(D1): D279-285.

Fox, A. N., R. J. Pitts, H. M. Robertson, J. R. Carlson and L. J. Zwiebel (2001). "Candidate odorant receptors from the malaria vector mosquito *Anopheles gambiae* and evidence of down-regulation in response to blood feeding." Proc Natl Acad Sci U S A **98**(25): 14693-14697.

Frank, D. D., A. Enjin, G. C. Jouandet, E. E. Zaharieva, A. Para, M. C. Stensmyr and M. Gallio (2017). "Early integration of temperature and humidity stimuli in the *Drosophila* brain." Curr Biol **27**(15): 2381-2388 e2384.

Freeman, E. G. and A. Dahanukar (2015). "Molecular neurobiology of *Drosophila* taste." Curr Opin Neurobiol **34**: 140-148.

Fukushima, N., D. Furuta, Y. Hidaka, R. Moriyama and T. Tsujiuchi (2009). "Post-translational modifications of tubulin in the nervous system." J Neurochem **109**(3): 683-693.

Galizia, C. G. and S. Sachse (2010). Odor coding in insects. The neurobiology of olfaction. A. Menini. Boca Raton, FL, CRC Press/Taylor & Francis**:** 35-70.

Ganguly, A., L. Pang, V. K. Duong, A. Lee, H. Schoniger, E. Varady and A. Dahanukar (2017). "A molecular and cellular context-dependent role for Ir76b in detection of amino acid taste." Cell Rep **18**(3): 737-750.

Garm, A. and J. T. Høeg (2006). "Ultrastructure and functional organization of mouthpart sensory setae of the spiny lobster *Panulirus argus*: new features of putative mechanoreceptors." J Morphol **267**(4): 464-476.

Gawalek, P. and M. Stengl (2018). "The diacylglycerol analogs OAG and DOG differentially affect primary events of pheromone transduction in the hawkmoth *Manduca sexta* in a zeitgebertime-dependent manner apparently targeting TRP channels." Front Cell Neurosci **12**: 218.

Giribet, G. and G. D. Edgecombe (2012). "Reevaluating the arthropod tree of life." Annu Rev Entomol **57**: 167-186.

Gleeson, R. A. (1991). Intrinsic factors mediating pheromone communication in the blue crab, *Callinectes sapidus*. Crustacean Sexual Biology. R. T. Bauer and J. W. Martin. New York, Columbia University Press**:** 17-32.

Gleeson, R. A., H. G. Trapido-Rosenthal, J. T. Littleton and W. E. S. Carr (1989). "Purinergic receptors and dephosphorylating enzymes occur in both the gustatory and olfactory systems of the spiny lobster." Comp Biochem Physiol Part C **92**(2): 413-417.

Gnatzy, W., M. Schmidt and J. Römbke (1984). "Are the funnel-canal organs the 'campaniform sensilla' of the shore crab *Carcinus maenas* (Crustacea, Decapoda)? I. Topography, external structure and basic organization. ." Zoomorphology **104**(1): 11-20.

Gohl, T. and J. Krieger (2006). "Immunolocalization of a candidate pheromone receptor in the antenna of the male moth, *Heliothis virescens*." Invert Neurosci **6**(1): 13-21.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman and A. Regev (2011). "Full-length transcriptome assembly from RNA-Seq data without a reference genome." Nat Biotechnol **29**(7): 644-652.

Greer, P. L., D. M. Bear, J. M. Lassance, M. L. Bloom, T. Tsukahara, S. L. Pashkovski, F. K. Masuda, A. C. Nowlan, R. Kirchner, H. E. Hoekstra and S. R. Datta (2016). "A family of non-GPCR chemosensors defines an alternative logic for mammalian olfaction." Cell **165**(7): 1734-1748.

Groh-Lunow, K. C., M. N. Getahun, E. Groβe-Wilde and B. S. Hansson (2014). "Expression of ionotropic receptors in terrestrial hermit crab's olfactory sensory neurons." Front Cell Neurosci **8**: 448.

Groh, K. C., H. Vogel, M. C. Stensmyr, E. Groβe-Wilde and B. S. Hansson (2013). "The hermit crab's nose-antennal transcriptomics." Front Neurosci **7**: 266.

Grünert, U. and B. W. Ache (1988). "Ultrastructure of the aesthetasc (olfactory) sensilla of the spiny lobster, *Panulirus argus*." Cell Tissue Res **251**(1): 95-103.

Guntur, A. R., P. Gu, K. Takle, J. Chen, Y. Xiang and C. H. Yang (2015). "*Drosophila* TRPA1 isoforms detect UV light via photochemical production of H_2O_2." Proc Natl Acad Sci U S A **112**(42): E5753-5761.

Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. LeDuc, N. Friedman and A. Regev (2013). "*De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis." Nat Protoc **8**(8): 1494-1512.

Hallem, E. A. and J. R. Carlson (2006). "Coding of odors by a receptor repertoire." Cell **125**(1): 143-160.

Hatt, H. (1984). "Structural requirements of amino acids and related compounds for stimulation of receptors in crayfish walking leg." J Comp Physiol A **155**(2): 219-231.

Hatt, H. and B. W. Ache (1994). "Cyclic nucleotide- and inositol phosphate-gated ion channels in lobster olfactory receptor neurons." Proc Natl Acad Sci U S A **91**(14): 6264-6268.

Hazlett, B. A. (1990). "Source and nature of disturbance-chemical system in crayfish." J Chem Ecol **16**(7): 2263-2275.

Herbst, C. (1916). "Ueber die Regeneration von Antennenahnlichen Organen an Stelle von Augen.
VII. Die Anatomie der Gehirnerven und des Gehirns bei Krebsen mit Antennulis an Stelle von Augen." Arch Entw Mech Org **42**: 407-489.

Hinman, A., H. H. Chuang, D. M. Bautista and D. Julius (2006). "TRP channel activation by reversible covalent modification." Proc Natl Acad Sci U S A **103**(51): 19564-19568.

Hollins, B., D. Hardin, A. A. Gimelbrant and T. S. McClintock (2003). "Olfactory-enriched transcripts are cell-specific markers in the lobster olfactory organ." J Comp Neurol **455**(1): 125-138.

Hussain, A., M. Zhang, H. K. Ucpunar, T. Svensson, E. Quillery, N. Gompel, R. Ignell and I. C. Grunwald Kadow (2016). "Ionotropic chemosensory receptors mediate the taste and smell of polyamines." PLoS Biol **14**(5): e1002454.

Jones, P., D. Binns, H. Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, S. Pesseat, A. F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S. Y. Yong, R. Lopez and S. Hunter (2014). "InterProScan 5: genome-scale protein function classification." Bioinformatics **30**(9): 1236-1240.

Jones, W. D., P. Cayirlioglu, I. G. Kadow and L. B. Vosshall (2007). "Two chemosensory receptors together mediate carbon dioxide detection in *Drosophila*." Nature **445**(7123): 86-90.

Joseph, R. M. and J. R. Carlson (2015). "*Drosophila* chemoreceptors: a molecular interface between the chemical world and the brain." Trends Genet **31**(12): 683-695.

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler and L. S. Jermiin (2017). "ModelFinder: fast model selection for accurate phylogenetic estimates." Nat Methods **14**(6): 587-589.

Kamio, M., M. A. Reidenbach and C. D. Derby (2008). "To paddle or not: context dependent courtship display by male blue crabs, *Callinectes sapidus*." J Exp Biol **211**(Pt 8): 1243-1248.

Kamio, M., M. Schmidt, M. W. Germann, J. Kubanek and C. D. Derby (2014). "The smell of moulting: *N*-acetylglucosamino-1,5-lactone is a premoult biomarker and candidate component of the courtship pheromone in the urine of the blue crab, *Callinectes sapidus*." J Exp Biol **217**(Pt 8): 1286-1296.

Kang, K., S. R. Pulver, V. C. Panzano, E. C. Chang, L. C. Griffith, D. L. Theobald and P. A. Garrity (2010). "Analysis of *Drosophila* TRPA1 reveals an ancient origin for human chemical nociception." Nature **464**(7288): 597-600.

Katoh, K., K. Misawa, K. Kuma and T. Miyata (2002). "MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform." Nucleic Acids Res **30**(14): 3059-3066.

Katoh, K. and D. M. Standley (2013). "MAFFT multiple sequence alignment software version 7: improvements in performance and usability." Mol Biol Evol **30**(4): 772-780.

Keller, T. A. and M. J. Weissburg (2004). "Effects of odor flux and pulse rate on chemosensory tracking in turbulent odor plumes by the blue crab, Callinectes sapidus." Biol Bull **207**(1): 44-55.

Kem, W. R. and F. Soti (2001). "*Amphiporus* alkaloid multiplicity implies functional diversity: initial studies on crustacean pyridyl receptors." Hydrobiologia **456**(1): 221-231.

Kenny, N. J., X. Shen, T. T. Chan, N. W. Wong, T. F. Chan, K. H. Chu, H. M. Lam and J. H. Hui (2015). "Genome of the rusty millipede, *Trigoniulus corallinus*, illuminates diplopod, myriapod, and arthropod evolution." Genome Biol Evol **7**(5): 1280-1295.

Knecht, Z. A., A. F. Silbering, J. Cruz, L. Yang, V. Croset, R. Benton and P. A. Garrity (2017). "Ionotropic receptor-dependent moist and dry cells control hygrosensation in *Drosophila*." eLife **6:e26654**.

Knecht, Z. A., A. F. Silbering, L. Ni, M. Klein, G. Budelli, R. Bell, L. Abuin, A. J. Ferrer, A. D. Samuel, R. Benton and P. A. Garrity (2016). "Distinct combinations of variant ionotropic glutamate receptors mediate thermosensation and hygrosensation in *Drosophila*." eLife **5:e17879**.

Kozma, M. T., H. Ngo-Vu, Y. Y. Wong, N. S. Shukla, S. D. Pawar, A. Senatore, M. Schmidt and C. D. Derby (2020). "Comparison of transcriptomes from two chemosensory organs in four decapod crustaceans reveals hundreds of candidate chemoreceptor proteins." PLoS ONE **15**(3): e0230266.

Kozma, M. T., M. Schmidt, H. Ngo-Vu, S. D. Sparks, A. Senatore and C. D. Derby (2018). "Chemoreceptor proteins in the Caribbean spiny lobster, *Panulirus argus*: Expression of Ionotropic Receptors, Gustatory Receptors, and TRP channels in two chemosensory organs and brain." PLoS ONE **13**(9): e0203935.

Krishnan, A., M. S. Almen, R. Fredriksson and H. B. Schioth (2014). "Insights into the origin of nematode chemosensory GPCRs: putative orthologs of the *Srw* family are found across several phyla of protostomes." PLoS ONE **9**(3): e93048.

Krogh, A., B. Larsson, G. von Heijne and E. L. Sonnhammer (2001). "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes." J Mol Biol **305**(3): 567-580.

Kwon, J. Y., A. Dahanukar, L. A. Weiss and J. R. Carlson (2007). "The molecular basis of CO2 reception in Drosophila." Proc Natl Acad Sci U S A **104**(9): 3574-3578.

Kwon, Y., S. H. Kim, D. S. Ronderos, Y. Lee, B. Akitake, O. M. Woodward, W. B. Guggino, D. P. Smith and C. Montell (2010). "*Drosophila* TRPA1 channel is required to avoid the naturally occurring insect repellent citronellal." Curr Biol **20**(18): 1672-1678.

Lenz, P. H., V. Roncalli, R. P. Hassett, L. S. Wu, M. C. Cieslak, D. K. Hartline and A. E. Christie (2014). "*De novo* assembly of a transcriptome for *Calanus finmarchicus* (Crustacea, Copepoda) - the dominant zooplankter of the North Atlantic Ocean." PLoS ONE **9**(2): e88589.

Leung, N. Y., D. P. Thakur, A. S. Gurav, S. H. Kim, A. Di Pizio, M. Y. Niv and C. Montell (2020). "Functions of opsins in *Drosophila* taste." Curr Biol **30**(8): 1367-1379 e1366.

Li, B. and C. N. Dewey (2011). "RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome." BMC Bioinformatics **12**: 323.

Li, W., L. Jaroszewski and A. Godzik (2001). "Clustering of highly homologous sequences to reduce the size of large protein databases." Bioinformatics **17**(3): 282-283.

Ling, F., A. Dahanukar, L. A. Weiss, J. Y. Kwon and J. R. Carlson (2014). "The molecular and cellular basis of taste coding in the legs of *Drosophila*." J Neurosci **34**(21): 7148-7164.

Long, S. M. (2018). "A novel protocol for generating intact, whole-head spider cephalothorax tissue sections." Biotechniques **64**(4): 163-169.

Love, M. I., W. Huber and S. Anders (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." Genome Biol **15**(12): 550.

Lu, B., A. LaMora, Y. Sun, M. J. Welsh and Y. Ben-Shahar (2012). "ppk23-Dependent chemosensory functions contribute to courtship behavior in *Drosophila melanogaster*." PLoS Genet **8**(3): e1002587.

Macpherson, L. J., A. E. Dubin, M. J. Evans, F. Marr, P. G. Schultz, B. F. Cravatt and A. Patapoutian (2007). "Noxious compounds activate TRPA1 ion channels through covalent modification of cysteines." Nature **445**(7127): 541-545.

Matsuura, H., T. Sokabe, K. Kohno, M. Tominaga and T. Kadowaki (2009). "Evolutionary conservation and changes in insect TRP channels." BMC Evol Biol **9**(1): 228.

Mayer, M. L. (2011). "Emerging models of glutamate receptor ion channel structure and function." Structure **19**(10): 1370-1380.

Mayer, M. L. and N. Armstrong (2004). "Structure and function of glutamate receptor ion channels." Annu Rev Physiol **66**: 161-181.

Mayer, M. L., A. Ghosal, N. P. Dolman and D. E. Jane (2006). "Crystal structures of the kainate receptor GluR5 ligand binding core dimer with novel GluR5-selective antagonists." J Neurosci **26**(11): 2852-2861.

McClintock, T. S., B. W. Ache and C. D. Derby (2006). "Lobster olfactory genomics." Integr Comp Biol **46**(6): 940-947.

McGrath, L. L., S. V. Vollmer, S. T. Kaluziak and J. Ayers (2016). "*De novo* transcriptome assembly for the lobster *Homarus americanus* and characterization of differential gene expression across nervous system tissues." BMC Genomics **17**(1): 63.

Mellon, D., Jr. (1990). Evidence for non-topographic afferent projection and growth-related central reorganization in the crayfish olfactory system. Frontiers in crustacean neurobiology. Advance in life sciences. K. Wiese, W. D. Krenz, J. Tautz, H. Reichert and B. Mulloney, Birkhäuser, Basel**:** 49-57.

Mellon, D., Jr. (2012). "Smelling, feeling, tasting and touching: behavioral and neural integration of antennular chemosensory and mechanosensory inputs in the crayfish." J Exp Biol **215**(Pt 13): 2163-2172.

Mellon, D., Jr. and S. D. Munger (1990). "Nontopographic projection of olfactory sensory neurons in the crayfish brain." J Comp Neurol **296**(2): 253-262.

Meusemann, K., B. M. von Reumont, S. Simon, F. Roeding, S. Strauss, P. Kuck, I. Ebersberger, M. Walzl, G. Pass, S. Breuers, V. Achter, A. von Haeseler, T. Burmester, H. Hadrys, J. W. Wagele and B. Misof (2010). "A phylogenomic approach to resolve the arthropod tree of life." Mol Biol Evol **27**(11): 2451-2464.

Min, S., M. Ai, S. A. Shin and G. S. Suh (2013). "Dedicated olfactory neurons mediating attraction behavior to ammonia and amines in *Drosophila*." Proc Natl Acad Sci U S A **110**(14): E1321-1329.

Minh, B. Q., M. A. Nguyen and A. von Haeseler (2013). "Ultrafast approximation for phylogenetic bootstrap." Mol Biol Evol **30**(5): 1188-1195.

Missbach, C., H. K. Dweck, H. Vogel, A. Vilcinskas, M. C. Stensmyr, B. S. Hansson and E. Grosse-Wilde (2014). "Evolution of insect olfactory receptors." eLife **3**(0): e02115.

Missbach, C., H. Vogel, B. S. Hansson and E. Groβe-Wilde (2015). "Identification of odorant binding proteins and chemosensory proteins in antennal transcriptomes of the jumping bristletail *Lepismachilis y-signata* and the firebrat *Thermobia domestica*: evidence for an independent OBP-OR origin." Chem Senses **40**(9): 615-626.

Miyamoto, T., J. Slone, X. Song and H. Amrein (2012). "A fructose receptor functions as a nutrient sensor in the *Drosophila* brain." Cell **151**(5): 1113-1125.

Mombaerts, P. (2006). "Axonal wiring in the mouse olfactory system." Annu Rev Cell Dev Biol **22**: 713-737.

Montell, C. (2013). "Gustatory receptors: not just for good taste." Curr Biol **23**(20): R929-932.

Moore, P. A. and D. A. Bergman (2005). "The smell of success and failure: the role of intrinsic and extrinsic chemical signals on the social behavior of crayfish." Integr Comp Biol **45**(4): 650-657.

Murthy, V. N. (2011). "Olfactory maps in the brain." Annu Rev Neurosci **34**: 233-258.

Ng, R., S. S. Salem, S. T. Wu, M. Wu, H. H. Lin, A. K. Shepherd, W. J. Joiner, J. W. Wang and C. Y. Su (2019). "Amplification of *Drosophila* olfactory responses by a DEG/ENaC channel." Neuron **104**(5): 947-959 e945.

Nguyen, L. T., H. A. Schmidt, A. von Haeseler and B. Q. Minh (2015). "IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies." Mol Biol Evol **32**(1): 268-274.

Ni, L., P. Bronk, E. C. Chang, A. M. Lowell, J. O. Flam, V. C. Panzano, D. L. Theobald, L. C. Griffith and P. A. Garrity (2013). "A gustatory receptor paralogue controls rapid warmth avoidance in *Drosophila*." Nature **500**(7464): 580-584.

Ni, L., M. Klein, K. V. Svec, G. Budelli, E. C. Chang, A. J. Ferrer, R. Benton, A. D. Samuel and P. A. Garrity (2016). "The ionotropic receptors IR21a and IR25a mediate cool sensing in *Drosophila*." eLife **5:e13254**.

Nordstrom, K. J., M. Sallman Almen, M. M. Edstam, R. Fredriksson and H. B. Schioth (2011). "Independent HHsearch, Needleman--Wunsch-based, and motif analyses reveal the overall hierarchy for most of the G protein-coupled receptor families." Mol Biol Evol **28**(9): 2471-2480.

Northcutt, A. J., K. M. Lett, V. B. Garcia, C. M. Diester, B. J. Lane, E. Marder and D. J. Schulz (2016). "Deep sequencing of transcriptomes from the nervous systems of two decapod crustaceans to characterize genes important for neural circuit function and modulation." BMC Genomics **17**(1): 868.

Núñez-Acuña, G., V. Valenzuela-Muñoz, J. P. Marambio, S. Wadsworth and C. Gallardo-Escaráte (2014). "Insights into the olfactory system of the ectoparasite *Caligus rogercresseyi*: molecular characterization and gene transcription analysis of novel ionotropic receptors." Exp Parasitol **145**: 99-109.

Núñez-Acuña, G., F. Vera-Bizama, S. Boltaña, C. Hawes, J. P. Marambio, S. Wadsworth and C. Gallardo-Escárate (2016). "In-feed additives modulate ionotropic receptor genes

from the sea louse *Caligus rogercresseyi*: A comparative analysis in two host salmonid species." Aquaculture **451**(Supplement C): 99-105.

Peñalva-Arana, D. C., M. Lynch and H. M. Robertson (2009). "The chemoreceptor genes of the waterflea *Daphnia pulex*: many Grs but no Ors." BMC Evol Biol **9**(1): 79.

Peng, G., X. Shi and T. Kadowaki (2015). "Evolution of TRP channels inferred by their classification in diverse animal species." Mol Phylogenet Evol **84**: 145-157.

Pezier, A., Y. V. Bobkov and B. W. Ache (2009). "The Na+/Ca2+ exchanger inhibitor, KB-R7943, blocks a nonselective cation channel implicated in chemosensory transduction." J Neurophysiol **101**(3): 1151-1159.

Pikielny, C. W. (2012). "Sexy DEG/ENaC channels involved in gustatory detection of fruit fly pheromones." Sci Signal **5**(249): pe48.

Pitts, R. J., S. L. Derryberry, Z. Zhang and L. J. Zwiebel (2017). "Variant ionotropic receptors in the malaria vector mosquito *Anopheles gambiae* tuned to amines and carboxylic acids." Sci Rep **7**: 40297.

Poulin, R. X., S. Lavoie, K. Siegel, D. A. Gaul, M. J. Weissburg and J. Kubanek (2018). "Chemical encoding of risk perception and predator detection among estuarine invertebrates." Proc Natl Acad Sci U S A **115**(4): 662-667.

Poynton, H. C., S. Hasenbein, J. B. Benoit, M. S. Sepulveda, M. F. Poelchau, D. S. T. Hughes, S. C. Murali, S. Chen, K. M. Glastad, M. A. D. Goodisman, J. H. Werren, J. H. Vineis, J. L. Bowen, M. Friedrich, J. Jones, H. M. Robertson, R. Feyereisen, A. Mechler-Hickson, N. Mathers, C. E. Lee, J. K. Colbourne, A. Biales, J. S. Johnston, G. A. Wellborn, A. J. Rosendale, A. G. Cridge, M. C. Munoz-Torres, P. A. Bain, A. R. Manny, K. M. Major, F. N. Lambert, C. D. Vulpe, P. Tuck, B. J. Blalock, Y. Y. Lin, M. E. Smith, H. Ochoa-Acuna, M. M. Chen, C. P. Childers, J. Qu, S. Dugan, S. L. Lee, H. Chao, H. Dinh, Y. Han, H. Doddapaneni, K. C. Worley, D. M. Muzny, R. A. Gibbs and S. Richards (2018). "The toxicogenome of *Hyalella azteca*: A model for sediment ecotoxicology and evolutionary toxicology." Environ Sci Technol **52**(10): 6009-6022.

Regier, J. C., J. W. Shultz, A. Zwick, A. Hussey, B. Ball, R. Wetzer, J. W. Martin and C. W. Cunningham (2010). "Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences." Nature **463**(7284): 1079-U1098.

Roberts, R. E., C. A. Motti, K. W. Baughman, N. Satoh, M. R. Hall and S. F. Cummins (2017). "Identification of putative olfactory G-protein coupled receptors in Crown-of-Thorns starfish, *Acanthaster planci*." BMC Genomics **18**(1): 400.

Roberts, R. E., D. Powell, T. Wang, M. H. Hall, C. A. Motti and S. F. Cummins (2018). "Putative chemosensory receptors are differentially expressed in the sensory organs of male and female crown-of-thorns starfish, Acanthaster planci." BMC Genomics **19**(1): 853.

Robertson, G., J. Schein, R. Chiu, R. Corbett, M. Field, S. D. Jackman, K. Mungall, S. Lee, H. M. Okada, J. Q. Qian, M. Griffith, A. Raymond, N. Thiessen, T. Cezard, Y. S. Butterfield, R. Newsome, S. K. Chan, R. She, R. Varhol, B. Kamoh, A. L. Prabhu, A. Tam, Y. Zhao, R. A. Moore, M. Hirst, M. A. Marra, S. J. Jones, P. A. Hoodless and I. Birol (2010). "De novo assembly and analysis of RNA-seq data." Nat Methods **7**(11): 909-912.

Robertson, H. M. (2015). "The insect chemoreceptor superfamily is ancient in animals." Chem Senses **40**(9): 609-614.

Robertson, H. M. (2019). "Molecular evolution of the major arthropod chemoreceptor gene families." Annu Rev Entomol **64**: 227-242.

Robertson, H. M., R. L. Baits, K. K. O. Walden, A. Wada-Katsumata and C. Schal (2018). "Enormous expansion of the chemosensory gene repertoire in the omnivorous German cockroach *Blattella germanica*." J Exp Zool B Mol Dev Evol **330**(5): 265-278.

Robertson, H. M., C. G. Warr and J. R. Carlson (2003). "Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*." Proc Natl Acad Sci U S A **100 Suppl 2**(suppl 2): 14537-14542.

Roddenberry, G. (1966). "Star Trek [TV series]." on National Broadcasting Company.

Rytz, R., V. Croset and R. Benton (2013). "Ionotropic receptors (IRs): chemosensory ionotropic glutamate receptors in *Drosophila* and beyond." Insect Biochem Mol Biol **43**(9): 888-897.

Saina, M., H. Busengdal, C. Sinigaglia, L. Petrone, P. Oliveri, F. Rentzsch and R. Benton (2015). "A cnidarian homologue of an insect gustatory receptor functions in developmental body patterning." Nat Commun **6**: 6243.

Sánchez-Alcañiz, J. A., A. F. Silbering, V. Croset, G. Zappia, A. K. Sivasubramaniam, L. Abuin, S. Y. Sahai, D. Munch, K. Steck, T. O. Auer, S. Cruchet, G. L. Neagu-Maier, S. G. Sprecher, C. Ribeiro, N. Yapici and R. Benton (2018). "An expression atlas of variant ionotropic glutamate receptors identifies a molecular basis of carbonation sensing." Nat Commun **9**(1): 4252.

Sandeman, D., R. Sandeman, C. Derby and M. Schmidt (1992). "Morphology of the brain of crayfish, crabs, and spiny lobsters: a common nomenclature for homologous structures." Biol Bull **183**(2): 304-326.

Sandeman, R. and D. Sandeman (2003). "Development, growth, and plasticity in the crayfish olfactory system." Microsc Res Tech **60**(3): 266-277.

Sarajlic, S., N. Edirisinghe, Y. Lukinov, M. Walters, B. Davis and G. Faroux (2016). "Orion: discovery environment for HPC research and bridging XSEDE resources." Proceedings of Xsede16: Diversity, Big Data, and Science at Scale: 54.

Sarajlic, S., N. Edirisinghe, Y. Wu, Y. Jiang and G. Faroux (2017). Training-based workforce development in Advanced Computing for Research and Education (ACoRE). Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact. New Orleans, LA, USA, Association for Computing Machinery**:** Article 71.

Schachtner, J., M. Schmidt and U. Homberg (2005). "Organization and evolutionary trends of primary olfactory brain centers in Tetraconata (Crustacea + Hexapoda)." Arthropod Struct Devel **34**(3): 257-299.

Schmidt, M. (1990). "Ultrastructure of a possible new type of crustacean cuticular strain receptor in *Carcinus maenas* (Crustacea, Decapoda)." J Morphol **204**(3): 335-344.

Schmidt, M. (2001). "Neuronal differentiation and long-term survival of newly generated cells in the olfactory midbrain of the adult spiny lobster, *Panulirus argus*." J Neurobiol **48**(3): 181-203.

Schmidt, M. (2014). Adult Neurogenesis in Crustaceans. The natural history of the Crustacea. C. Derby and M. Thiel. New York, NY, Oxford University Press.

Schmidt, M. (2016). Malacostraca. Structure and evolution of invertebrate nervous systems. A. Schmidt-Rhaesa, S. Harzsch and G. Purschke. Oxford, UK, Oxford University Press**:** 529-582.

Schmidt, M. and B. W. Ache (1992). "Antennular projections to the midbrain of the spiny lobster. II. Sensory innervation of the olfactory lobe." J Comp Neurol **318**(3): 291-303.

Schmidt, M. and B. W. Ache (1996). "Processing of antennular input in the brain of the spiny lobster, *Panulirus argus*. I. Non-olfactory chemosensory and mechanosensory pathway of the lateral and median antennular neuropils." J Comp Physiol A **178**(5): 579-604.

Schmidt, M. and B. W. Ache (1996). "Processing of antennular input in the brain of the spiny lobster, *Panulirus argus*. II. The olfactory pathway." J Comp Physiol A **178**(5): 605-628.

Schmidt, M., H. Chien, T. Tadesse, M. E. Johns and C. D. Derby (2006). "Rosette-type tegumental glands associated with aesthetasc sensilla in the olfactory organ of the Caribbean spiny lobster, *Panulirus argus*." Cell Tissue Res **325**(2): 369-395.

Schmidt, M. and C. D. Derby (2005). "Non-olfactory chemoreceptors in asymmetric setae activate antennular grooming behavior in the Caribbean spiny lobster *Panulirus argus*." J Exp Biol **208**(Pt 2): 233-248.

Schmidt, M. and C. D. Derby (2011). "Cytoarchitecture and ultrastructure of neural stem cell niches and neurogenic complexes maintaining adult neurogenesis in the olfactory midbrain of spiny lobsters, *Panulirus argus*." J Comp Neurol **519**(12): 2283-2319.

Schmidt, M. and W. Gnatzy (1984). "Are the funnel-canal organs the 'campaniform sensilla' of the shore crab, *Carcinus maenas* (Decapoda, Crustacea)? II. Ultrastructure." Cell Tissue Res. **237**(1).

Schmidt, M. and W. Gnatzy (1989). "Specificity and response characteristics of gustatory sensilla (funnel-canal organs) on the dactyls of the shore crab, *Carcinus maenas* (Crustacea, Decapoda)." J Comp Physiol A **166**(2).

Schmidt, M. and D. Mellon, Jr. (2011). Neuronal processing of chemical information in crustaceans. Chemical communication in crustaceans. T. Breithaupt and M. Thiel. New York, NY, Springer**:** 123-147.

Schmidt, M., L. Van Ekeris and B. W. Ache (1992). "Antennular projections to the midbrain of the spiny lobster. I. Sensory innervation of the lateral and medial antennular neuropils." J Comp Neurol **318**(3): 277-290.

Schram, F. R. (2012). Comments on crustacean biodiversity and disparity of body forms. The natural history of the Crustacea. L. Watling and M. Thiel. New York, NY, Oxford University Press**:** 1-33.

Scott, K., R. Brady, Jr., A. Cravchik, P. Morozov, A. Rzhetsky, C. Zuker and R. Axel (2001). "A chemosensory gene family encoding candidate gustatory and olfactory receptors in *Drosophila*." Cell **104**(5): 661-673.

Silbering, A. F. and R. Benton (2010). "Ionotropic and metabotropic mechanisms in chemoreception: 'chance or design'?" EMBO Rep **11**(3): 173-179.

Silbering, A. F., R. Rytz, Y. Grosjean, L. Abuin, P. Ramdya, G. S. Jefferis and R. Benton (2011). "Complementary function and integrated wiring of the evolutionarily distinct *Drosophila* olfactory subsystems." J Neurosci **31**(38): 13357-13375.

Simao, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva and E. M. Zdobnov (2015). "BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs." Bioinformatics **31**(19): 3210-3212.

Spencer, M. and K. A. Linberg (1986). "Ultrastructure of aesthetasc innervation and external morphology of the lateral antennule setae of the spiny lobster *Panulirus interruptus* (Randall)." Cell Tissue Res **245**(1): 69-80.

Stensmyr, M. C., S. Erland, E. Hallberg, R. Wallen, P. Greenaway and B. S. Hansson (2005). "Insect-like olfactory adaptations in the terrestrial giant robber crab." Curr Biol **15**(2): 116-121.

Stepanyan, R., B. Hollins, S. E. Brock and T. S. McClintock (2004). "Primary culture of lobster (*Homarus americanus*) olfactory sensory neurons." Chem Senses **29**(3): 179-187.

Steullet, P., H. S. Cate and C. D. Derby (2000). "A spatiotemporal wave of turnover and functional maturation of olfactory receptor neurons in the spiny lobster *Panulirus argus*." J Neurosci **20**(9): 3282-3294.

Strotmann, J., O. Levai, J. Fleischer, K. Schwarzenbacher and H. Breer (2004). "Olfactory receptor proteins in axonal processes of chemosensory neurons." J Neurosci **24**(35): 7754-7761.

Tadesse, T., M. Schmidt, W. W. Walthall, P. C. Tai and C. D. Derby (2011). "Distribution and function of *splash*, an *achaete-scute* homolog in the adult olfactory organ of the Caribbean spiny lobster *Panulirus argus*." Dev Neurobiol **71**(4): 316-335.

Thistle, R., P. Cameron, A. Ghorayshi, L. Dennison and K. Scott (2012). "Contact chemoreceptors mediate male-male repulsion and male-female attraction during *Drosophila* courtship." Cell **149**(5): 1140-1151.

Thomas, J. H. and H. M. Robertson (2008). "The *Caenorhabditis* chemoreceptor gene families." BMC Biol **6**: 42.

Thorne, N. and H. Amrein (2008). "Atypical expression of *Drosophila* gustatory receptor genes in sensory and central neurons." J Comp Neurol **506**(4): 548-568.

Tran, N. M., D. L. Mykles, A. Elizur and T. Ventura (2019). "Characterization of G-protein coupled receptors from the blackback land crab *Gecarcinus lateralis* Y organ transcriptome over the molt cycle." BMC Genomics **20**(1): 74.

Trifinopoulos, J., L. T. Nguyen, A. von Haeseler and B. Q. Minh (2016). "W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis." Nucleic Acids Res **44**(W1): W232-235.

Tsuji, K., K. Kobayashi, E. Hasegawa and J. Yoshimura (2020). "Dimorphic flowers modify the visitation order of pollinators from male to female flowers." Sci Rep **10**(1): 9965.

Upadhyay, A., A. Pisupati, T. Jegla, M. Crook, K. J. Mickolajczyk, M. Shorey, L. E. Rohan, K. A. Billings, M. M. Rolls, W. O. Hancock and W. Hanna-Rose (2016). "Nicotinamide is an endogenous agonist for a *C. elegans* TRPV OSM-9 and OCR-4 channel." Nat Commun **7**: 13135.

van Giesen, L. and P. A. Garrity (2017). "More than meets the IR: the expanding roles of variant Ionotropic Glutamate Receptors in sensing odor, taste, temperature and moisture." F1000Res **6**: 1753.

Venkatachalam, K., J. Luo and C. Montell (2014). "Evolutionarily conserved, multitasking TRP channels: lessons from worms and flies." Handb Exp Pharmacol **223**: 937-962.

Venkatachalam, K. and C. Montell (2007). "TRP channels." Annu Rev Biochem **76**(1): 387-417.

Vizueta, J., J. Rozas and A. Sanchez-Gracia (2018). "Comparative genomics reveals thousands of novel chemosensory genes and massive changes in chemoreceptor repertories across chelicerates." Genome Biol Evol **10**(5): 1221-1236.

Vosshall, L. B. (2001). "The molecular logic of olfaction in *Drosophila*." Chem Senses **26**(2): 207-213.

Wallace, B. G., B. R. Talamo, P. D. Evans and E. A. Kravitz (1974). "Octopamine: selective association with specific neurons in the lobster nervous system." Brain Res **74**(2): 349-355.

Wang, K., Y. Guo, F. Wang and Z. Wang (2011). "*Drosophila* TRPA channel painless inhibits male-male courtship behavior through modulating olfactory sensation." PLoS ONE **6**(11): e25890.

Waterhouse, A. M., J. B. Procter, D. M. Martin, M. Clamp and G. J. Barton (2009). "Jalview Version 2 - a multiple sequence alignment editor and analysis workbench." Bioinformatics **25**(9): 1189-1191.

Waterhouse, R. M., M. Seppey, F. A. Simao, M. Manni, P. Ioannidis, G. Klioutchnikov, E. V. Kriventseva and E. M. Zdobnov (2017). "BUSCO applications from quality assessments to gene prediction and phylogenomics." Mol Biol Evol.

Waterhouse, R. M., M. Seppey, F. A. Simao and E. M. Zdobnov (2019). Using BUSCO to assess insect genomic resources. Insect Genomics. Methods in Molecular Biology. S. Brown and M. Pfrender, Humana Press, New York, NY. **1858:** 59-74.

Wicher, D. and E. Groβe-Wilde (2017). Chemoreceptors in evolution. Evolution of nervous systems. J. H. Kaas. London, UK, Academic Press. **Volume 1: The evolution of the nervous systems of the nonmammalian vertebrates:** 245-255.

Wolfe, J. M., J. W. Breinholt, K. A. Crandall, A. R. Lemmon, E. M. Lemmon, L. E. Timm, M. E. Siddall and H. D. Bracken-Grissom (2019). "A phylogenomic framework, evolutionary timeline and genomic resources for comparative studies of decapod crustaceans." Proc Biol Sci **286**(1901): 20190079.

Xiang, Y., Q. Yuan, N. Vogt, L. L. Looger, L. Y. Jan and Y. N. Jan (2010). "Light-avoidance-mediating photoreceptors tile the *Drosophila* larval body wall." Nature **468**(7326): 921-926.

Xu, H., X. Luo, J. Qian, X. Pang, J. Song, G. Qian, J. Chen and S. Chen (2012). "FastUniq: a fast de novo duplicates removal tool for paired short reads." PLoS ONE **7**(12): e52249.

Yoshida, T., R. Inoue, T. Morii, N. Takahashi, S. Yamamoto, Y. Hara, M. Tominaga, S. Shimizu, Y. Sato and Y. Mori (2006). "Nitric oxide activates TRP channels by cysteine S-nitrosylation." Nat Chem Biol **2**(11): 596-607.

Zbinden, M., C. Berthod, N. Montagne, J. Machon, N. Leger, T. Chertemps, N. Rabet, B. Shillito and J. Ravaux (2017). "Comparative study of chemosensory organs of shrimp from hydrothermal vent and coastal environments." Chem Senses **42**(4): 319-331.

Zhang, Y. V., J. Ni and C. Montell (2013). "The molecular basis for attractive salt-taste coding in *Drosophila*." Science **340**(6138): 1334-1338.

Zufall, F. and S. Munger (2016). The detection of odors, tastes, and other chemostimuli. Chemosensory transduction. . F. Zufall and S. Munger. Cambridge, MA, Academic Press.