THE MOLECULAR UNDERPINNINGS OF NEURONAL CELL IDENTITY IN THE STOMATOGASTRIC GANGLION OF CANCER BOREALIS

A Dissertation

Presented to

The Faculty of the Graduate School

At the University of Missouri

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

By

Northcutt, Adam J. (MU-Student)

Dr. David J. Schulz, Dissertation Supervisor

May 2019

The undersigned, appointed by the dean of the Graduate School,

have examined the dissertation entitled

THE MOLECULAR UNDERPINNINGS OF NEURONAL CELL IDENTITY IN THE

STOMATOGASTRIC GANGLION OF CANCER BOREALIS

Presented by Northcutt, Adam J. (MU-Student)

A candidate for the degree of

Doctor of Philosophy

And hereby certify that, in their opinion, it is worthy of acceptance.

_____

Dr. David J. Schulz

_____

Dr. Andrew D. McClellan

_____

Dr. Michael L. Garcia

_____

Dr. Johannes Schul

_____

Dr. David Kline

Acknowledgements

My Ph.D., and the work herein, is a product of the David Schulz lab at University of Missouri-Columbia.

First and foremost, I must give my sincerest appreciation to my advisor and mentor Dr. David Schulz. When I first met Dr. Schulz during my interviews, I could tell that this man was a leader worth following, and his passion for his work was contagious. In his lab, I found a home where I could thrive intellectually and push myself to find answers to questions that at times seemed daunting. This is due in no small part to the exceptional support and guidance Dr. Schulz provided to me and all students under his purview. He has shaped my way of thinking and problem solving, giving me confidence inside and outside the lab to take on new challenges and responsibilities. He is more than just an incredible mentor, he is also a dear friend who has been there for me as I've gone through some of the most critical moments in my life.

Next I'd like to give thanks to the members of my advisory committee: Drs. David Schulz, Andrew McClellan, Mike Garcia, Johannes Schul, and David Kline. Not only has my committee been exceptional in guiding me and providing feedback throughout my graduate career, but I have also gained much experience publishing collaborative papers with them and their graduate students. The broad expertise and skill sets that these professors have provided me with continues to bolster my progression as a professional research scientist.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Throughout the life of an organism, the nervous system must be able to balance changing in response to environmental stimuli with the need to produce reliable, repeatable activity patterns to create stereotyped behaviors. Understanding the mechanisms responsible for this regulation requires a wealth of knowledge about the neural system, ranging from network connectivity and cell type identification to intrinsic neuronal excitability and transcriptomic expression. To make strides in this area, we have employed the well-described stomatogastric nervous system of the Jonah crab *Cancer borealis* to examine the molecular underpinnings and regulation of neuron cell identity.

Several crustacean circuits, including the stomatogastric nervous system and the cardiac ganglion, continue to provide important new insights into circuit dynamics and modulation (Diehl, White, Stein, & Nusbaum, 2013; Marder, 2012; Marder & Bucher, 2007; Williams et al., 2013), but this work has been partially hampered by the lack of extensive molecular sequence knowledge in crustaceans. Here we generated *de novo* transcriptome assembly from central nervous system tissue for *C. borealis* producing 42,766 contigs, focusing on an initial identification, curation, and comparison of genes that will have the most profound impact on our understanding of circuit function in these species. This included genes for 34 distinct ion channel types, 17 biogenic amine and 5 GABA

receptors, 28 major transmitter receptor subtypes including glutamate and acetylcholine receptors, and 6 gap junction proteins – the Innexins.

With this reference transcriptome and annotated sequences in hand, we sought to determine the strengths and limitations of using the neuronal molecular profile to classify them into cell types. Unambiguous classification of neuronal cell types is a long-standing goal in neuroscience with the aim to understand the functional components of the nervous system that give rise to circuits and, ultimately, behavior. Because ion channels, receptors, gap junctions, and neurotransmitters are some of the most salient features that give rise to unique neuronal properties (Tripathy et al., 2017), we used a multiplexed qRT-PCR approach to assay 65 transcripts related to neuronal function from the wide diversity of cell types found in the STG and CG of the *C. borealis* to attempt to recapitulate known cell identity. Using both a variety of commonly employed statistical analyses and machine learning approaches for identifying novel cell types, we measured the efficacy by which accurate cell identity could be achieved. Our results demonstrate the limitations of unsupervised methods that rely solely on the molecular profile to recapitulate neuron identity by showing how poorly predicted clusters recapitulated known cell identity. However, once we employed supervised methods incorporating known cell identification and expected cluster numbers the results greatly improved to capturing most of the known identities using only a focal set of genes.

Since the resulting activity of a neuron is the product of the expression of ion channel genes, we sought to further probe the expression profile of neurons across a range of cell types to understand how these patterns of mRNA abundance relate to the properties of individual cell types. Distinct activity patterns have been shown to drive the expression of different ensembles of genes (Tyssowski et al., 2018). To drive a consistent activity pattern over the lifetime of a neuron requires homeostatic feedback to maintain excitability and synaptic strength (Turrigiano & Nelson, 2004). Therefore, one possibility of the mechanism by which neurons achieve maintenance of activity is though the differential regulation and expression of genes most directly related to neuronal function. We examined the absolute mRNA abundances and correlation networks of 57 mRNAs related to neuronal function to glean information about potential co-regulation of expression patterns. Our results demonstrate cell-type specific patterns of mRNA abundances that relate to known the electrophysiological and network properties of STG neurons.

Finally, we sought to better understand the molecular underpinnings of how these correlated patterns of mRNA expression are generated and maintained. We used a multifaceted approach to tackle this question by perturbing STG neurons through attempting to disrupt preexisting correlations or forming new mRNA correlation patterns not previously found in the endogenous neuron. Our results demonstrated that the 3' untranslated region of mRNAs alone is insufficient to drive correlations; certain mRNA correlations are

dependent on transcription, while others are not; and finally that introduction of

exogenous mRNA can disrupt endogenous mRNA correlations in a time-

dependent manner.

Chapter 1

Molecular Mechanisms of Homeostatic Plasticity in Central Pattern Generating Networks

Introduction

Central pattern generating networks (CPGs) produce rhythmic neural activity that drives periodically repeating behaviors, such as locomotion, respiration, and digestion (Delcomyn, 1980). Being highly evolutionarily conserved, CPGs are present in both vertebrate and invertebrate nervous systems (Katz, 2016). The number of neurons comprising these networks can vary across orders of magnitude; only a handful of neurons controlling swimming in Tritonia (Getting, 1983b) or digestion in crabs and lobsters (Marder & Bucher, 2007), while the mammalian spinal cord coordinates thousands of neurons for locomotion (Guertin, 2009). While the activity of CPGs is modulated by sensory inputs, the hallmark of a CPG network is the ability to independently produce outputs even in the absence of sensory feedback, often with the network removed and isolated from the rest of the organism (Marder & Bucher, 2001).

Throughout the life of an organism, the nervous system must be able to balance changing in response to environmental stimuli with the need to produce

reliable, repeatable activity patterns to create stereotyped behaviors. While most plasticity research focuses on changes in neuronal activity in processes such as learning, it is as important to understand how neurons respond to changing environmental conditions to produce the same output. This is particularly true of CPGs, in that they must be able to homeostatically regulate their activity patterns to produce consistent outputs, often ones necessary for the survival of the organism, such as respiration. This review will focus on addressing molecular aspects of plasticity that are characteristic of CPG networks in general. Plasticity is characterized by a stimulus-induced, long-term alteration in activity that persists beyond the duration of the stimulus (Mitchell & Johnson, 2003). Homeostatic plasticity refers to changes in neuronal properties that help stabilize and maintain a given activity pattern (Turrigiano & Nelson, 2004).

The principles of homeostasis rely on a set point by which changes from this set point are sensed, integrated, and subsequently result in alterations that return the variable back to the desired set point. A consensus on the molecular basis of a regulatory set point for homeostatic plasticity has not been clearly defined, although evidence suggests the mechanism involves regulating intracellular [Ca^{2+}] as a broad governing signal (O'Leary, Williams, Franci, & Marder, 2014; Turrigiano, 2012). Here we examine what is known about the molecular regulation of homeostatic plasticity in CPG neurons. These include activity-dependent, activity-independent, and neuromodulator-dependent

pathways that result in changes in transcription, translation, epigenetic modifications, and RNA stability to maintain network activity in CPGs.

Activity-Dependent Homeostatic Plasticity in CPGs

An established mechanism by which activity-dependent changes in transcription take place is through calcium influx associated with the phosphorylation of the transcription factor cAMP-response element binding (CREB) protein (West et al., 2001). Further, long-term [Ca^{2+}] changes have been found to regulate the expression of ion channel (Barish, 1998; Mermelstein, Bito, Deisseroth, & Tsien, 2000; O'Leary, van Rossum, & Wyllie, 2010; Turrigiano, Abbott, & Marder, 1994; Wheeler et al., 2012) and receptor (Greenberg, Ziff, & Greene, 1986a) subtypes. Computational models have been generated simulating the dynamics with which neurons of CPG networks can tune their transcription and translation of ion channels to enable maintenance of homeostatic plasticity for conserving activity (O'Leary, Williams, Franci, & Marder, 2014). As shown in Figure 1.1, this ion channel expression model proposes that individual neurons have a set optimum [Ca^{2+}] that maintains calcium-dependent transcription of ion channel mRNAs, which in turn have unique rates of synthesis and stability that will further affect the degree to which they are translated into functional proteins. These ion channel proteins then will influence the intracellular calcium concentrations, acting as a feedback loop for maintaining a set level of calcium that is related to the activity of the neuron.

Since CPG networks are rhythmically active, the unique time scales on which

calcium sensors would measure activity are potentially a fast band corresponding

to transient calcium caused by spiking, a slow band from rhythmic bursting

releasing waves of calcium, and steady-state calcium influx levels (O'Leary,

Williams, Caplan, & Marder, 2013). Biological evidence supporting this model

was generated using the STG of *C. borealis*, showing that correlated ion channel

mRNA abundances required activity to maintain expression patterns (Temporal,

Lett, & Schulz, 2014). This study used the muscarinic agonist pilocarpine in the

presence and absence of TTX in decentralized preparations to decouple activity

from modulation, showing that ion channel mRNA correlation patterns were only

maintained following decentralization if network activity was still present. This

data supports the computational model that calcium influx through voltage-gated

calcium channels is the signal for transcription, as pilocarpine is known to

activated the modulator-evoked current I_{MI}, which is highly selective for K^+ and

Na^+, but not Ca^{2+} (Golowasch & Marder, 1992).

In *Drosophila,* male courtship behavior is considered to be driven by fixed

action patterns produced by CPG networks (Yamamoto & Koganezawa, 2013).

Typical male-female courtship involves a complex series of behaviors, including

following and tapping the female's legs, licking of her genitalia, and singing a

courtship song (Pavlou & Goodwin, 2013). In male flies that have a mutant

fruitless protein (a transcription factor required for motor neuron development),

male-to-male courtship presents, forming chains of courting males (Yamamoto &

Kohatsu, 2017). Recent studies have shown that this behavior is not rigid, but in fact can be influenced by experience-dependent and sensory-guided inputs, leading to plasticity in courtship behaviors, particularly in (Pan & Baker, 2014). The courtship decision-making center resides in the P1 neuron cluster of male fly brain (Hoopfer, Jung, Inagaki, Rubin, & Anderson, 2015). Expression of *fruitless* in P1 is hypothesized as a potential encoder of experience-dependent tuning of courtship behavior. Beyond its role in development, *fruitless* has been shown to fine-tune excitability through regulating transcription. The male isoform of *fruitless,* FruM, has zinc finger motifs and a BTB domain (Ito et al., 1996), interacts with more than 100 loci on polytene chromosomes (Hiroki Ito et al., 2012), and FruM fused with DNA-methylase targets specific genomic enhancers (Neville et al., 2014). Further, *fru*-expressing neurons have been shown to have a marked increase in Ca^{2+}-mediated transcription factor activity upon housing male flies with 10 female flies, as compared with only allowing the odor of females but no direct contact (Masuyama, Zhang, Rao, & Wang, 2012). Together, these results indicate that *fruitless* potentially mediates activity-dependent changes in transcription through epigenetic modifications that result in both behavioral and neural plasticity.

In mammalian systems, CPGs in motor networks of the spinal cord underlie the various rhythmic movement patterns involved in locomotion. Following a spinal cord injury (SCI), these motor networks rework themselves in an activity-dependent manner (Bertrand & Cazalets, 2013), leading to an overall

increase in network excitability that can manifest as pain and spasticity (Bennett et al., 1999; Frigon & Rossignol, 2006; Hultborn, 2003). This hyperexcitability has been found to be in part the result of time-dependent changes in a complex series of discrete gene clusters due to differential expression of transcription factors (Ryge et al., 2010). Specifically, the transcription factors that resulted in the down-regulation of genes include E2F5, GT2D1, and NFIA, while transcription factors that increased expression include PBX3 and NRF2F. Further, genes primarily associated with motor neuron excitability and spasticity as a result of injury were found to be regulated at later time points, and excitatory receptor and channel genes were generally up-regulated, while receptors associated with inhibition are mainly down-regulated. This indicates that the plasticity in these motor CPG networks undergoes is positively correlated with genes enhancing neural excitability and negatively correlated with the inhibitory GABA receptor system (Ryge et al., 2010).

Taken together, these studies paint a picture of plasticity of CPG networks being mediated through activity-dependent feedback with calcium signaling at the heart of changes in transcription (both mRNA and miRNA), translation, epigenetic modification, and RNAi degradation. These changes ultimately result in the long-term alterations in correlated mRNA expression patterns, network activity, and current conductances carried by ion channels. This plasticity can be homeostatic (to maintain a desired set point) or potentiating

(increasing in response to a stimuli or perturbation), which in turn can lead ultimately to behavioral changes at the whole organism level.
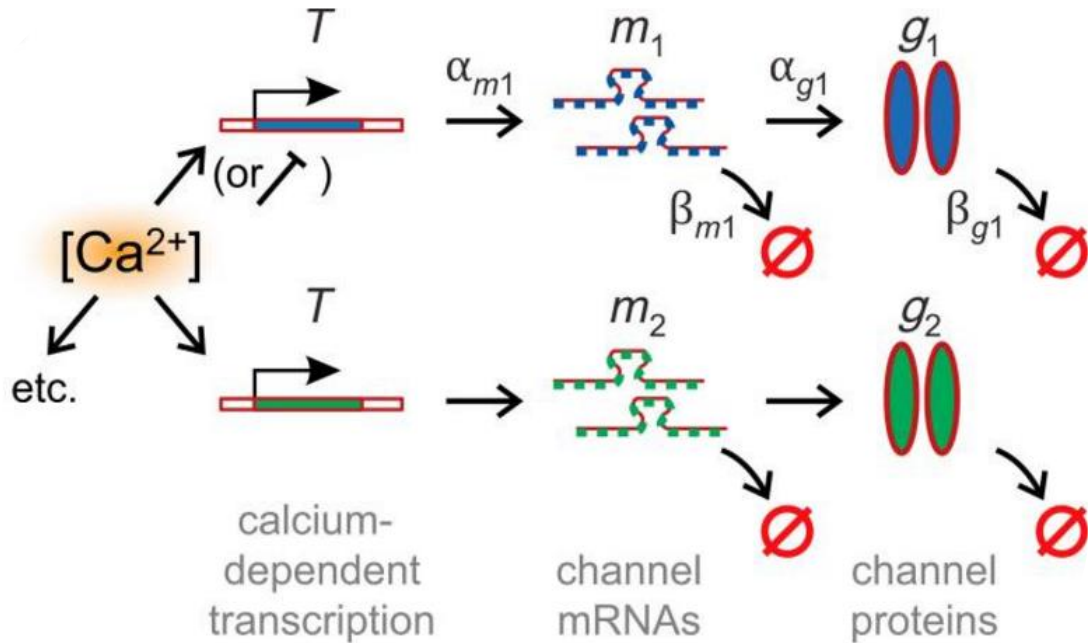
FIGURE 1.1: *Activity-dependent plasticity mediated through calcium signaling.* In this model, neurons maintain a set [Ca2+] used as a signal for changes in activity patterns to influence transcriptional patterns. Channel mRNAs are synthesized at a rate α_m that depends on a Ca^{2+}-activated factor, *T*, and subsequently degraded at rate β_m. Channel proteins are then transcribed at a rate α_g from mRNAs and degraded at a rate β_g. Adapted from O'Leary et al. 2014.

Activity-Independent Homeostatic Plasticity in CPGs

In *Drosophila* motoneurons, it has been shown that a loss of the A-type potassium channel *Shal* (Kv4.2 ortholog) leads to an increase in *Shaker*

expression as a compensation, indicating transcriptional coupling between the two genes to regulate A-type channel expression (Bergquist, Dickman, & Davis, 2010). The molecular mechanism that underlies this compensation is activation of the transcription factor Krüppel that selectively responds to the loss of *Shal,* rather than responding to changes in neuronal activity (Parrish et al., 2014). In the STG of the spiny lobster, *Panulirus interruptus,* pyloric dilator (PD) neurons were injected with *Shal* mRNA, increasing the A-current, but leaving the activity of PD unchanged (MacLean, Zhang, Johnson, & Harris-Warrick, 2003). This was shown to be due to a compensatory increase in the hyperpolarization-activated inward current (I_H), but the precise mechanism for this compensation is still unclear. However, it is unlikely that the increase in I_H is the result of changes in a transcription factor such as Krüppel since blockade of transcription by actinomycin-D did not prevent the compensation, instead indicating that post-transcriptional or translational mechanisms are likely at play (MacLean et al., 2003). Interestingly, the relationship of the compensation between I_A and I_H in *P. interruptus* appears to be unidirectional, in that increase of I_H via *PAIH* mRNA injection into PD neurons does not result in a corresponding increase in I_A (Zhang et al., 2003). The firing activity of the PD neurons is altered by the *PAIH* mRNA injection, and blockade of I_H through application of Cs+ returned the injected neurons back toward control PD activity. Theoretical modeling of this relationship in a two-cell model suggests that the oscillations of the neuron are more sensitive to changes in I_A conductance than that of I_H, and therefore homeostatically controlling overexpression of I_A might be of greater consequence

than expression of I_H (Zhang et al., 2003). Further, the I_H-I_A relationship is bore

out at the channel mRNA level as well with correlated absolute abundances

across multiple neuron types in the STG (Schulz, Goaillard, & Marder, 2007).

Modulator-Dependent Homeostatic Plasticity in CPGs

In CPG networks, the stereotyped and repetitive activity patterns must be

robust to changing environmental conditions, requiring modifications at

transcriptional, translational, and post-translational levels to maintain a

conserved output. These modifications can be driven in part by neuromodulatory

signals transmitted from other neural centers. Directly following removal of

descending input (deafferentation) in the stomatogastric ganglion of the lobster,

there is a loss of triphasic patterning in the pyloric network (Thoby-Brisson &

Simmers, 1998). However, over the next few days the network restructures itself

in an activity-independent manner to resume producing a triphasic activity pattern

in spite of the loss of neuromodulation. Beyond whole network restructuring, it

has also been shown that LP neurons exhibit persistent increases in fast

transient outward K$^+$ current (I_A) and hyperpolarization-activated current (I_H) in

the presence of tonic dopamine (Rodgers, Fu, Krenz, & Baro, 2011; Rodgers,

Krenz, & Baro, 2011), a common neuromodulator often found in CPG networks.

This dopaminergic driven plasticity was determined to be dependent on Protein

Kinase A (PKA), RNA polymerase II transcription, mTORC1-dependent

translation, and the RNA interference (RNAi) pathway (Krenz, Parker, Rodgers,

& Baro, 2014; Krenz, Rodgers, Baro, Selverston, & Liu, 2015). One proposed mechanism for this plasticity is signaling through the dopamine receptor D1R-PKA axis acting, either directly or indirectly, on calcineurin to influence changes in I_H. Tonic dopamine has been shown to increase the release of Ca^{2+} from internal stores (Kadiri, Kwan, Webb, & Harris-Warrick, 2011). Taken together, these experiments suggest CPG plasticity in LP neurons is carried out through calcineurin activation, potentially leading to changes in the production of mRNAs and miRNAs (which are generated by RNA polymerase II along with mRNAs), which in turn could influence the translation rates of ion channel transcripts thought to carry the I_A and I_H currents (*Shal/Shaker* and *HCN*, respectively).

Another, non-mutually exclusive, proposed mechanism for achieving homeostatic plasticity in the stomatogastric ganglion of *C. borealis* is through extrinsic rather than intrinsic pathways. That is, the extracellular matrix composition is a necessary component of the recovery and reworking of network activity following removal of neuromodulation. Removal of chondroitin sulfate proteoglycans (CSPGs) from the extracellular matrix does not disrupt activity of the pyloric network, but prevents the short-term recovery of the network after deafferentation has removed descending neuromodulatory input (Hudson, Gollnick, Gourdine, & Prinz, 2015). CSPGs are implicated in acting as guidance cues for growth cones (Siebert & Osterhout, 2011), neurogenesis and axonal regeneration following spinal cord injury (Siebert, Conta Steencken, & Osterhout, 2014), and postnatal development and plasticity (Galtrey & Fawcett, 2007). In

attempts to improve recovery of the central pattern generating networks in the

mammalian spinal cord, degrading CSPGs during critical stages during

transplantation enhances the functional recovery of locomotion (Kim, Dai,

Lynskey, McAtee, & Bregman, 2006). Taken together, these data suggest a

multifactorial role for the presence of CSPGs in maintaining CPG state, as the

removal of which prevents recovery of activity in the invertebrate STG, while

CSPG removal allows for improved recovery in the mammalian spinal cord.

Conclusions:

    As discussed in this review, CPGs can maintain their activity patterns

through the homeostatic regulation of intrinsic properties (ion channel expression

and balancing conductance), network properties (synaptic reorganization and

neuromodulation), and extrinsic properties (extracellular protein regulation). The

molecular basis for a regulatory set point is still a point of contention, but

intracellular calcium signaling is a strong contender as a signal for activity that

feeds back to molecular changes in CPGs to maintain their rhythmic, reliable

outputs.

CHAPTER 2

Deep Sequencing of Transcriptomes from the Nervous Systems

of Two Decapod Crustaceans to Characterize Genes Important

for Neural Circuit Function and Modulation

Introduction

Crustaceans have been studied extensively as model systems for nervous

system function from single neuron properties to behavior. However, lack of

molecular sequence information and tools have slowed the adoption of these

physiological systems as molecular model systems. In this study, we sequenced

and performed de novo assembly for the nervous system transcriptomes of two

decapod crustaceans: the Jonah crab (*Cancer borealis*) and the American

lobster (*Homarus americanus*). 42,766 and 60,273 contigs were assembled from

C. borealis and *H. americanus* respectively, representing 9,489 and 11,061

unique coding sequences. From these transcripts, genes associated with neural

function were identified and manually curated to produce a characterization of

multiple gene families important for nervous system function. This included

genes for 34 distinct ion channel types, 17 biogenic amine and 5 GABA

receptors, 28 major transmitter receptor subtypes including glutamate and

acetylcholine receptors, and 6 gap junction proteins – the Innexins. With this resource, crustacean model systems are better poised for incorporation of modern genomic and molecular biology technologies to further enhance the interrogation of fundamentals of nervous system function.

Despite their status as important economic species, their important place in understanding the evolution and phylogeny of arthropods, and as models for neurobiology research, crustaceans have been largely overlooked in the rush to apply modern molecular biology and high throughput sequencing approaches to work in "non-genetic" systems. Even among arthropods they are fairly poorly represented, with insects dominating the ranks of those with available genome and transcriptome assemblies. Two indicators of this are seen in the number of relative sequence read archive (SRA) and GEO profile publicly available in NCBI: at the time of this article, there were 2,323 crustacean and 46,866 insect SRAs, and 4,608 crustacean and 1,275,029 insect GEO profiles. To date, only two crustacean genomes have been made publicly available (Colbourne et al., 2011; Zeng et al., 2011), with the first, the water flea *Daphnia pulex*, coming only in 2011 (Colbourne et al., 2011) – a full 11 years after the first arthropod genome was sequenced (Adams et al., 2000). As a microcrustacean, *Daphnia* is still a far cry from the large decapod crustaceans that are common models in neuroscience research, such as crabs, lobsters, crayfish and shrimp. Thus even with some very recent additions to decapod crustacean transcriptome data

(Christie et al., 2015; McGrath, Vollmer, Kaluziak, & Ayers, 2016), there is a strong need to add to our sequence knowledge of these species.

Many fundamental findings in neuroscience were made with crustacean preparations. To mention only a small subset of these, command fibers (Wiersma & Ikeda, 1964), electrical coupling (Furshpan & Potter, 1959) and presynaptic inhibition (Dudel & Kuffler, 1961) were first described using crustacean preparations. Work on crayfish and lobsters established GABA as an inhibitory transmitter (Otsuka, Iversen, Hall, & Kravitz, 1966; Otsuka, Kravitz, & Potter, 1967), and allowed early studies of the relevance of the fast outward current, I_A, for action potential generation and propagation (Connor, 1975; Connor, Walter, & McKown, 1977). The first intracellular fluorescent dye-fills were pioneered with crustaceans (Remler, Selverston, & Kennedy, 1968; Stretton & Kravitz, 1968), and crustacean systems were used early on to understand the organization of circuits in behavior (Hartline, 1967; Heitler, 1978; Maynard, 1972; Mulloney & Selverston, 1974).

Several crustacean circuits, including the stomatogastric nervous system and the cardiac ganglion, continue to provide important new insights into circuit dynamics and modulation (Diehl et al., 2013; Marder, 2012; Marder & Bucher, 2007; Williams et al., 2013), but this work has been partially hampered by the lack of extensive molecular sequence knowledge in crustaceans, in particular in *C. borealis*. In this study, we generated *de novo* transcriptome assemblies from

central nervous system tissue for two commonly used species in neuroscience research: the Jonah crab (*Cancer borealis*) and the American lobster (*Homarus americanus*). We focus on an initial identification, curation and comparison of genes that will have the most profound impact on our understanding of circuit function in these species, namely channels and receptors, with the hopes of fostering new avenues of research for these preparations that continue to be valuable assets in our understanding of nervous system dynamics. Additionally, such sequence information provides a valuable resource for comparative molecular neuroscience approaches across phyla.

Materials and Methods

Tissue collection and RNA preparation

Adult lobsters, *H. americanus*, and crabs, *C. borealis*, were obtained from The Fresh Lobster Company (Gloucester, Massachusetts, USA) and maintained in artificial seawater at 12°C until used. Lobsters and crabs were anesthetized by packing them in ice for 30 minutes before dissection. The brain, abdominal nerve cord, cardiac ganglion and complete stomatogastric nervous system (STNS) (including the commissural, esophageal, and stomatogastric ganglia) was dissected out of two lobsters and pinned out in a Sylgard (Dow Corning)–coated dish containing chilled (12–13 °C) physiological saline. From two crabs we

dissected out brain, complete STNS, and cardiac ganglia. Any adherent connective tissue and muscle was removed to the extent possible, and the tissues were rinsed several times in physiological saline (Lobster saline composition in mM/l: 479.12 NaCl, 12.74 KCl, 13.67 $CaCl_2$, 20.00 $MgSO_4$, 3.91 Na_2SO_4, 11.45 Trizma base, and 4.82 maleic acid [pH = 7.45]; Crab saline composition in mM/l: 440.0 NaCl, 11.0 KCl, 13.0 $CaCl_2$, 20.00 $MgCl_2$, 11.2 Trizma base, and 5.1 maleic acid [pH = 7.45]) made with ultrapure, RNase free water. After dissection, tissues for each species were combined and homogenized in Trizol (Invitrogen). The resulting pool of RNA was therefore derived from mixed nervous system tissue. Insoluble tissues were pelleted by centrifugation, and the supernatant removed and stored at -80⁰ C until RNA extraction. Total RNA was isolated as per the protocol provided by the manufacturer (Invitrogen), and subsequently treated with DNase (Zymo Research) prior to library construction.

Library production, sequencing, and de novo transcriptome assembly

Library construction and RNA-sequencing were performed for a fee by GENEWIZ, Inc. (South Plainfield, New Jersey, USA). Briefly, RNA samples were quantified using Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, California, USA) and RNA integrity checked with RNA6000 Nano Assay using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, California, USA). cDNA library preparation and sequencing reactions were conducted by GENEWIZ, Inc.

Illumina TruSeq RNA library prep, clustering, and sequencing reagents were used throughout the process following the manufacturer's recommendations (Illumina, San Diego, California, USA). Specifically, mRNAs were purified using poly-T oligo-attached magnetic beads and then fragmented. The first and the second strand cDNAs were synthesized and end repaired. Adaptors were ligated after adenylation at the 3'ends. Then cDNA templates were enriched by PCR. cDNA libraries were validated using a High Sensitivity Chip on the Agilent 2100 Bioanalyzer. The cDNA library was quantified using Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, California, USA) and by qPCR. The samples were clustered on a flow cell using the cBOT. After clustering, the samples were loaded on the Illumina HiSeq 2000 instrument for sequencing with a 2x100 paired-end configuration.

Raw sequence data generated from Illumina HiSeq 2000 was converted into fastq files and de-multiplexed using Illumina CASSAVA 1.8.2 program. Fastq files from the sample were imported into CLC Genomics Workbench Server 5.0.1. Sequence reads were trimmed to remove bases with low quality ends. *De novo* assembly was conducted with the trimmed reads utilizing the CLC Genomics Server. The total length of the assembled transcripts was 99,847,148 bp for lobster and 66,058,464 bp for crab (see Table 2.1). To ensure that the CLC Genomics assembly was of high quality, we performed a second round of *de novo* assembly using the SeqMan NGen assembler from the DNAstar software suite (SeqMan NGen®. Version 13.0. DNASTAR. Madison, WI.).

Following assembly, quality of assembled contigs was investigated by comparison with species-specific sequences contained within GenBank that were previously generated largely by Sanger sequencing approaches. We used BLAST+ command line application (Version 2.2.30+) to perform blastn comparisons of the curated GenBank sequence versus transcriptome contigs, and calculate percent nucleotide identity for the top hit of each sequence for both species.

BUSCO transcriptome quality assessment

To analyze the completeness of our transcriptomes, a reference-based alignment was performed using Benchmarking Universal Single-Copy Orthologs (BUSCO) software (Version 1.22). The arthropod BUSCO reference contains 2675 orthologous genes found within >90% of the 38 arthropod species' genomes used to construct the reference (Simão, Waterhouse, Ioannidis, Kriventseva, & Zdobnov, 2015). The four transcriptomes we assembled were aligned against the arthropod reference, resulting in percentages of the reference genes found as complete, fragmented, or missing from our transcriptomes. "Complete" genes are those which align to a reference gene with a mean length within two standard deviations (i.e. 95%) of the reference value. Genes that only partially align are deemed "fragmented", and those present in the reference with no match found in the transcriptome are classified as "missing".

Functional Annotation

For gene ontology (GO) term analysis, the Blast2GO software package (Version 3.1.3) (Conesa et al., 2005) was used for functional annotation of the assembled transcriptomes. A blastx search with an E-value threshold of 10^{-5} was carried out against the NCBI non-redundant (nr) protein sequence database. Assignment of gene names to each contig was based on the highest scoring BLAST hit. Scoring of the annotated sequences utilized a threshold score of ≥ 55. The top 10 significant hits for each query extracted from the blastx search were used for further gene annotation. Query sequences were categorized into three broad ontological classifications: molecular function, cellular component, and biological process. GO annotation filters included: E-value-Hit-Filter of 1.0e-6, Annotation CutOff of 55, and GO Weight of 5.

Whole-transcriptome alignment comparison

The software VennBLAST (Zahavi, Stelzer, Strauss, Salmon, & Salmon-Divon, 2015) was used to compare the whole *C. borealis* and *H. americanus* transcriptomes against the *Daphnia pulex* (GCA_000187875.1) protein sequences from Ensembl Metazoa. Protein sequence database for *D. pulex* was chosen as a common subject to query against the *C. borealis* and *H. americanus*

transcriptomes. Initially, a local blastx of *C. borealis* or *H. americanus* contigs

against *D. pulex* protein sequences was performed with the BLAST+ command

line application (Version 2.2.30+). This output was run through the VennBLAST

Merge tool with the InterGroup Option: Use Subject to quantify the relative

overlap of *C. borealis* and *H. americanus* with *D. pulex*. A second layer of filtering

was performed using the VennBLAST Filter tool with an Identity percent of 70

and an E-value threshold of 1.0e-5, and this output was subsequently merged in

the same manner mentioned previously.

Ion channel and receptor sequence identification and alignment analysis

We identified putative orthologs of channels and receptors from the

transcriptomes of crab and lobster as follows. We created local blast databases

from the assembled contigs of each transcriptome. Because channels and

receptors are fairly well conserved across diverse taxa, and because the mouse

research community has agreed upon a well-curated systematic naming system

for channel and receptor genes, we used mouse reference mRNA sequence for

each gene of interest as the query in a tblastn search of each transcriptome

database. We used initially stringent e-value cutoffs (1e-100 to 1e-50) for our

searches to find very high sequence similarity matches. Top contig matches from

these blast searches were then compared with the results of the remainder of the

blast queries for a given gene family. Often multiple mouse input sequences

resulted in the same top hit from the crustacean transcriptomes, indicating that

fewer members of the gene family were present in our invertebrate sequence than the mammalian gene families. Once a complete gene family search was obtained, all putative orthologs were then blasted against the Non-Redundant Protein (NR) Sequence Database hosted at NCBI via blastx. This allowed us to look for conserved sequence across all taxa and confirm a given gene identification. Once gene families were obtained from the *C. borealis* transcriptome, the process was repeated with *H. americanus* as the subject database. We additionally used the crab sequences as queries to find the direct ortholog for a given gene in *H. americanus*. These sequences were confirmed in the same way via blastx against the NR database, and moved forward into sequence alignment as described below. As is to be expected from manually performed sequence-by-sequence discovery and curation such as this, at times other searches and sequence comparisons were performed on a case-by-case basis with comparator species such as *Daphnia pulex*, *Drosophila melanogaster*, or other insect species to gain insight or clarification as to the best possible identification for a given transcript.

The web-based software tool Biology Workbench (Version 3.2) (Subramaniam, 1998) was used for sequence analysis of putative ion channels and receptors from the assembled transcriptomes. Coding regions were determined based on the longest open reading frame (ORF) from the SIXFRAME tool in Biology Workbench. ClustalW was utilized (default parameters) to perform the multiple sequence alignment (MSA) for ion channel and receptor family

subtypes based on amino acid sequences from predicted coding regions. The rooted phylogenetic trees were constructed from the output of the MSA from ClustalW.

We used blastp to generate percent identities and similarities for predicted amino acid sequences of orthologs between species. We used only sequences that were full-length or those that were near full length. Sequences were assumed to be full length coding sequences if they met three criteria: the sequence began with a start codon, was approximately the same length as similar sequences in the non-redundant database based on a blastx search, and the sequence ended with a stop codon. Sequences were considered to be close to full-length if they were at least 80% the length of similar sequences on the NR database, regardless of the presence of start and stop codons. In addition, the two sequences from crab and lobster had to be at least 80% of the length of one another. This generated 42 pairwise comparisons for orthologous protein sequence between crabs and lobsters.

Results

We note that all nomenclature for transcripts described in this study will conform to one nomenclature convention: transcripts from *Cancer borealis*, will be noted with the species prefix *Cb-*, and the species prefix *Ha-* will be used for

Homarus americanus gene products. All curated gene sequences described below were submitted to GenBank and assigned individual accession numbers as noted in Tables 2.2-2.5.

Mixed nervous system transcriptome sequencing and de novo assembly

Constructing RNA-seq libraries from nervous tissues of adult crustaceans, a total of 414,978,768 and 452,237,240 raw reads were obtained from the paired end sequencing of *C. borealis* and *H. americanus,* respectively. The average read length for both species was approximately 97 bp, as expected for 100 bp paired-end Illumina Sequencing. Following quality checks removing adaptors, contaminating sequences, and low-quality sequences, 391,060,790 (94.2%) clean reads were found for *C. borealis* and 426,712,238 (94.4%) for *H. americanus*. These high-quality cleaned reads were subsequently assembled *de novo* into contigs using two different assemblers: CLC Genomics and Seqman NGen. For *C. borealis,* CLC assembly resulted in 42,766 contigs with an average length of 1544 bp and an N50 length of 2178 bp, while SeqMan assembly resulted in 67,380 contigs with an average length of 1076 bp and N50 of 1239 (Table 2.1; Figure 2.1A). For *H. americanus,* CLC assembly resulted in 60,273 contigs with an average length of 1657 bp and N50 length of 2357 bp, while SeqMan NGen resulted in 45,043 contigs with an average length of 1799 bp and N50 of 2258 (Table 2.1; Figure 2.1A).

To compare the quality of our transcriptomes from multiple assemblies, a reference-based alignment was performed using Benchmarking Universal Single-Copy Orthologs (BUSCO) software (Version 1.22) (Simão et al., 2015). The arthropod BUSCO reference contains 2675 orthologous genes expected within most arthropod species that were compared against the gene content of our transcriptomes. The alignment of our transcriptomes against the arthropod reference resulted in similar percentages of the reference genes found as complete (C:), fragmented (F:), or missing (M:) across our transcriptomes (Figure 2.1B). The *C. borealis* metrics were C: 59.0%, F: 13.5%, and M: 27.5% for the CLC Genomics assembly and C: 58.4%, F: 20.1%, and M: 21.5% for the SeqMan NGen assembly. The *H. americanus* metrics were C: 56.1%, F: 16.4%, and M: 27.5% for the CLC Genomics assembly and C: 66.5%, F: 13.4%, and M: 20.1% for the SeqMan NGen assembly. These results were compared against the arthropod transcriptome reference scores provided in the BUSCO supplementary materials, a recently published *Homarus americanus* nervous system transcriptome assembled using Trinity (McGrath et al., 2016), and a recent transcriptome of the freshwater crayfish *Astacus astacus* (Theissinger et al., 2016). Our results are comparable with the *Astacus* transcriptome in completeness and an apparent improvement upon the published *Homarus americanus* transcriptome (McGrath et al., 2016). One possible explanation for the missing arthropod genes from our transcriptomes can be explained by the

fact that our sequences were derived solely from nervous system tissue, while the references were built from arthropod genomic sequences.

Using the NCBI BLAST+ suite to perform a blastn of 28 *Cancer borealis* sequences already contained within GenBank against our assembled contigs, we found an average sequence identity from the CLC assembly of 99.2%, with the lowest identity score 96%. SeqMan NGen assembly for *C. borealis* transcripts had an average sequence identity of 99.03%, and the lowest percent identity was 95%. For *H. americanus*, 75 GenBank sequences were aligned against our transcriptome, resulting in an average sequence identity to CLC assembled sequences of 99.3% with the lowest being 94.5%. The average transcriptome contig length was 3.0 times as long as that of the GenBank curated sequences. SeqMan NGen assembly for *H. americanus* transcripts had an average sequence identity of 98.97%, and the lowest percent identity was 87%.

Based on the relative similarity in many of the metrics for these two assembly methods, the somewhat better performance of CLC contigs when compared with Sanger sequencing generated orthologs, and the fact that portions of the *H. americanus* transcriptome based on the CLC assembly have previously been published (Christie et al., 2015), we chose to perform the remaining representative analysis of these sequence data based on the CLC assembled contigs. The *H. americanus* Transcriptome Shotgun Assembly (TSA)

project has been deposited at GenBank under the Accession No.

GEBG00000000 (BioProject No. PRJNA300643; BioSample No.

SAMN04230440). The *C. borealis* Transcriptome Shotgun Assembly (TSA)

project has been deposited at GenBank under the Accession No.

GEFB00000000 (BioProject No. PRJNA310325; BioSample No.

SAMN04450329). The versions described in this paper represent the first

versions, GEBG01000000 and GEFB01000000 respectively.

Annotation and gene ontology mapping

Entrez Gene IDs were obtained for both transcriptomes using blastx

against the NCBI non-redundant (nr) protein database. These annotations

consisted of 9,489 unique proteins among *C.borealis* transcripts, and 11,061

among *H, americanus* transcripts. Mapping these gene IDs to Gene Ontology

(GO) categories yielded 9351 (22%) of the *C. borealis* and 6191 (10%) of the *H.*

americanus contigs successfully identified (Figure 2.2A). Similar percentages

have been observed in other *de novo* transcriptome analyses (Elmer et al., 2010;

Gu, Zhang, Kang, Dong, & Zhang, 2015). From the functional annotation,

transcripts were classified into three broad categories: cellular compartment

(CC), molecular function (MF), and biological process (BP) (Conesa et al., 2005).

Within these broad categorizations, the highest abundance GO terms of *H.*

americanus and *C. borealis* were compared against each other, which included

the top 9 CCs, 18 MFs, and 16 BPs for both species (Figure 2.2B). The

arrangement of GO terms was based on the highest abundance *H. americanus* terms, in descending order. The only notable exception to this order was the BP GO term "RNA-dependent DNA Replication" ontology due to its high abundance in *C. borealis* but relatively low abundance in *H. americanus.* These same GO terms were compared between the two species using the relative percentage of each GO term for its broad GO classification (CC, MF, BP) (Figure 2.3). The most striking differences between the GO ontologies of each species include a much higher incidence of "protein binding" terms for *C. borealis* MF than that of *H. americanus,* a much higher incidence of "metabolic process" in *H. americanus* BP, and a prominent difference between the "RNA-dependent DNA Replication" term for BP. The source of these differences could be attributed to factors including, but not limited to, the variation in tissue types (abdominal nerve cord was used in *H. americanus,* but not *C. borealis*), depth of sequencing, or natural variation in transcript abundance.

Species comparisons: distribution and VennBLAST analysis

Using the Blast2GO software suite, the number of species that the *C. borealis* and *H. americanus* neural transcriptomes align with was determined from a blastx against the NCBI non-redundant database. The species distribution for both *C. borealis* and *H. americanus* gave similar top species, such as *Tribolium castaneum, Daphnia pulex,* and *Strongylocentrotus purpuratus* within the top 5 species hits (Figure 2.4). The absence of termite (*Zootermopsis*

nevadensis) from the *C. borealis* species distribution of blast hits is due to the fact that the *Z. nevadensis* protein sequences had yet to be uploaded to the NCBI non-redundant database at the time of the blast analysis of the *C. borealis* transcriptome, while the *H. americanus* analysis was performed after the *Z. nevadensis* reference became available.

       Venn diagrams were generated (Figure 2.5A) using the software VennBLAST (Zahavi et al., 2015) to compare the whole *C. borealis* and *H. americanus* transcriptomes against the *Daphnia pulex* (GCA_000187875.1) protein sequences from Ensembl Metazoa. The protein sequence database for *D. pulex* was chosen as a common subject to query the *C. borealis* and *H. americanus* transcriptomes against due to the mutual high top-hit species distribution (Figure 2.4), as well as the well-annotated genome of the crustacean *D. pulex* (Colbourne et al., 2011). Initially, a local blastx of *C. borealis* or *H. americanus* contigs against *D. pulex* protein sequences resulted in 17,343 and 14,818 hits, respectively. Upon overlapping these hits with the VennBLAST Merge tool, 11,258 hits from *C. borealis* and *H. americanus* were found to have the same top hit for *D. pulex*. A second analysis with increased stringency was performed using the VennBLAST Filter tool to retain only high-quality matches, leaving *C. borealis* with 7,460 and *H. americanus* with 7,268 hits to *D. pulex.* Subsequent merger of these filtered hits resulted in 6,226 common top-hits for *D. pulex*, resulting in an increased percentage (from 54% overlap to 73%) of common top-hits.

For the remainder of our transcriptome analysis, we identified and characterized sequences for 6 different Innexin proteins (gap junctions), 34 distinct ion channel types, 17 biogenic amine receptors, 5 GABA receptors, and 28 major transmitter receptor subtypes including glutamate and acetylcholine receptors. These are described in detail below. These receptor groups consisted of 27 different ligand-gated channel subunits (ionotropic receptors) and 23 metabotropic receptor types. To better quantify the similarity across lobster and crab, we performed analyses of percent amino acid identity and similarity between orthologs of a subset of genes (see Methods). Overall, the sequence similarity is very high between these species, as one might expect for members of the same Order (Figure 2.5B); across 42 genes surveyed, there was a mean ± SD of 85.27% ± 8.46% amino acid identity between genes. However, we also noticed a significant trend across different types of gene products: there was significantly lower amino acid identity and similarity for metabotropic receptors than for the other classes of genes (Figure 2.5B). In particular, the G-protein coupled (GPCRs) receptors that we analyzed were some of the most divergent between crab and lobster. For example, *Octβ-R1, 3,* and *4* shared 72%, 72%, and 74% amino acid identity respectively. Conversely, the most highly conserved genes were in the Shaker family of voltage-gated K$^+$ channels. *Shaker, Shal,* and *Shaw1* were 98%, 98%, and 96% identical between crabs and lobsters. From these results we would predict more conservation in channel function and physiology across species than that of the GPCRs.

Ion Channels

For our initial analysis of these crustacean transcriptomes, we decided to focus on some of the most critical proteins involved in nervous system function. We therefore first conducted an analysis of ion channel subtypes. Putative ion channels were identified based on tblastx searches utilizing the transcriptomes as a reference database and querying with known channel protein largely consisting of sequences from mouse (*Mus musculus*) and *Drosophila melanogaster*. A 100% overlapping set of ion channels were found to be present in both *C. borealis* and *H. americanus* (Table 2.2; Figure 2.6). We specifically hand-curated and annotated these channel sequences, and the full list is available in Table 2.2, including putative current types carried by each channel. We used the multiple sequence alignment (MSA) output from CLUSTALW (Thompson, Higgins, & Gibson, 1994) to develop a fairly comprehensive ion channel tree based on amino acid sequence similarity, allowing us to cluster channels by type to effectively interrogate the nervous system channel content of these crustaceans.

Our analysis of the crab and lobster transcriptomes led us to identify and characterize 34 distinct channel subtypes representing several gene families (Table 2.2; Figure 2.6). Each identified channel transcript was found in both crab

and lobster transcriptomes. Our analysis confirmed the presence of 3 major voltage-dependent calcium channel subtypes (Ransdell, Temporal, West, Leyrer, & Schulz, 2013) corresponding one each to the L-type (*CaV1*), P/Q-type (*CaV2*), and T-type (*CaV3*) families of calcium channels. In addition, we identified a single member of the NaV-type voltage-gated Na$^+$ channel representing the *para* type channel identified in other species. Finally, in both species we identified a non-selective sodium leak channel (*NALCN*) thought to underlie TTX-resistant Na$^+$ conductance important in baseline neuronal excitability (Lu et al., 2007). One other major family of non-selective cation channels we identified was the cyclic-nucleotide-gated channels of the HCN/CNG type. Both species contained one member of the hyperpolarization-activated cyclic nucleotide-gated (*HCN*) channel family, the channels that give rise to I$_H$ type currents in crustacean neurons (Ouyang, Goeritz, & Harris-Warrick, 2007). In addition, we identified 3 α -subunits and 1 β-subunit of the cyclic nucleotide gated ion channel types (*CNG*), which are activated by the binding of cAMP and cGMP to carry a non-selective cation current (Podda & Grassi, 2014).

The pore-forming α-subunits of K$^+$ channels can be sub-divided into voltage-dependent subunits (K$_v$), inward rectifiers (K$_{ir}$), two-pore subunits (K2P), and those activated by intracellular calcium (K$_{Ca}$) or sodium (K$_{Na}$) ions. We identified a diverse array of voltage-dependent potassium channel subtypes in the transcriptomes of both the crab and lobster (Table 2.2; Figure 2.5). The best already characterized of these channels in crustaceans are the Shaker family of

channels, having been identified in crab (Schulz, Goaillard, & Marder, 2006;

Schulz et al., 2007) and spiny lobsters (*Panulirus interruptus*, (Baro, Coniglio, et

al., 1996; Baro, Cole, & Harris-Warrick, 1996)), with the latter having an

extensive characterization via expression in oocyte systems (Baro, Quiones,

Lanning, Harris-Warrick, & Ruiz, 2001; M. Kim et al., 1998). Previously, 4

members of this family were already known from *Cancer borealis*: *shaker, shal,

shab,* and *shaw.* We found orthologs to each of these in *H. americanus* as well.

We further discovered that in both species there actually were 2 distinct *shaw*-

related channel transcripts (Figure 2.6). The newly re-named "*Shaw1*" transcript

from this analysis is a perfect match with the previously identified *shaw* transcript

from *Cancer borealis* (Accession #EF089569), while the newly identified *Shaw*-

like transcript is presented as *Shaw2.* We also identified two members of the

KCNQ family of K^+ channels. *KCNQ* genes encode a family of six

transmembrane domain K^+ channel alpha-subunits that have a wide range of

physiological roles, including likely underlying the slow voltage-gated M-type

currents (Robbins, 2001). Rounding out the voltage-dependent K^+ channel

subtypes are 3 members of the *ether-a-go-go/KCNH* family (Ganetzky et al.,

1999). In addition to voltage-dependent K^+ channels, we also identified one K_{ir}

channel (*IRK*), two members of the K2P family (*KCNK*), one sodium-activated K^+

channel (*KCNT*), and two calcium-activated K^+ channel types. These K_{Ca}

channels had previously been identified in *C. borealis* (Schulz et al., 2007) and

correspond to one BK- and one SK- type channel.

Transient receptor potential (TRP) channels have been implicated as a primary channel for generation of sensations including temperature, taste, pain, pressure, and vision. In our analysis, we found various TPR subfamilies within both *C. borealis* and *H. americanus* (Table 2.2; Figure 2.6). These subfamilies included TRPV (vanilloid), mediating odor and pain sensations; TRPA (ankyrin), associated with mechanical stress receptors; TRPM (melastatin), associated with magnesium reabsorption (Wissenbach, Niemeyer, & Flockerzi, 2004); and TRP pyrexia, a thermal sensing receptor (Lee et al., 2005). In crustaceans, TRP channels have been primarily studied for their role in olfactory reception (Bobkov & Ache, 2005) and stretch reception (Rydqvist & Purali, 1993). Not all orthologs of TRP channels were identified in both species. We did not identify in this data set orthologs of *TRP-V6* and *pyrexia* from *C. borealis* and an ortholog of *TRP-M1* was not identified in *H. americanus.* It is most likely that these "missing" orthologs are due to limitations in the sequence depth, although we cannot rule out the possibility that these two species have distinct complements of TRP channel genes. No sequences were found in either species that represent the TRPC, TRPP, TRPL, or TRPN subfamilies. These results are consistent with found in a previously published transcriptome of *H. americanus* that identified 2 TRPA, one *pyrexia,* and two TRPM type channels (McGrath et al., 2016). We extend these results to include identification of the TRPV family of channels in both *H. americanus* and *C. borealis.*

Biogenic Amine Receptors

Biogenic amine neuromodulators were some of the first modulatory

compounds to be thoroughly studied in the crustacean nervous system (Barker,

Kushner, & Hooper, 1979; Dudel, 1965; Glanzman & Krasne, 1983; Livingstone,

Harris-Warrick, & Kravitz, 1980), and specifically in the stomatogastric nervous

system (Barker et al., 1979; Flamm & Harris-Warrick, 1986a, 1986b; E. Marder &

Eisen, 1984) where some of the most comprehensive understanding of the

multiple targets and modulatory impacts of these compounds on neural circuits

has been described (Ayali, Johnson, & Harris-Warrick, 1998; Beltz et al., 1984;

Cleland & Selverston, 1997; Harris-Warrick & Johnson, 2010; Johnson,

Kloppenburg, & Harris-Warrick, 2003; Johnson, Schneider, Nadim, & Harris-

Warrick, 2005; Kiehn & Harris-Warrick, 1992). Therefore, we decided it would be

valuable to provide a thorough characterization of these receptor subtypes as

well to complement the extensive and elegant physiology work that has been

going on for decades.

Dopamine has been perhaps the most extensively characterized biogenic

amine from a functional and biochemical perspective in crustaceans. Previous

work (Clark & Baro, 2006; Clark, Khan, & Baro, 2008) identified 3 subtypes of

dopamine receptors in the nervous system of the spiny lobster, *Panulirus*

interruptus: $D1_{\alpha Pan}$ (Type 1A DAr), $D1_{\beta Pan}$ (Type 1B DAr), and $D2_{\alpha Pan}$ (Type 2

DAr). We found clear orthologs to all three of these receptor subtypes in both *C.*

borealis and *H. americanus* (Table 2.3; Figure 2.7), and our transcriptome search

protocol did not come up with any other putative DAr subunit transcripts.

Therefore, it is likely that these three receptor subtypes represent the complete

complement of dopamine receptors in these decapod crustaceans. In deference

to the extensive characterization of these receptors in the closely related spiny

lobster, we conform the naming of these channels to match with the *Panulirus*

nomenclature: for example, *Cb-D1αR, Cb-D1βR,* and *Cb-D2αR* (Figure 2.7).

Serotonin receptors are less described in crustaceans than the dopamine

receptors. Previous reports describe two distinct subtypes of serotonin receptors

in the *P. interruptus* (Clark et al., 2004) as well as the crayfish *Procambarus*

clarkii (Sosa, Spitzer, Edwards, & Baro, 2004): one type-1 and one type-2

serotonin receptor. Our analysis of the transcriptomes of *C. borealis* and *H.*

americanus found clear orthologs to both of these receptor subtypes, and based

on homology with mouse and *Drosophila* sequences we identified two novel

putative serotonin receptor subtypes as well. Figure 2.6 uses the existing *P.*

interruptus and *P. clarkii* sequences to root the new sequences in a tree

representing these crustacean serotonin receptors. Our analysis suggests that

the previous type-1 5HTr subtypes identified are most similar to mammalian

$5HT_{1A}$ (*HTR1A*) receptors, while the crustacean type-2 5HTr subunit is most

similar to $5HT_{2B}$ (*HTR2B*) receptors. We also identified a putative type-1B

receptor (*HTR1B*) as well as a putative type-7 receptor (*HTR7*). These identities

are assigned based on mouse query sequences used in our tblastx protocols that

generated the strongest hits (i.e. lowest e-values) when queried against the

crustacean transcriptomes. We follow the mammalian classification and nomenclature guidelines for these 5HT receptors in assigning gene names (Table 2.3; Figure 2.7), as these are well defined and organized relative to the invertebrate nomenclature: e.g. *Cb-HTR1A, Cb-HTR1B, Cb-HTR2B,* and *Cb-HTR7.*

Octopamine receptors are virtually undescribed in crustaceans, with the sole decapod receptor described as a tyramine/octopamine receptor from the freshwater prawn, *Macrobrachium rosenbergii* (Reyes-Colón et al., 2010). Thorough work with crustacean octopamine receptors is found in the barnacle, *Balanus improvisus*, where one alpha- and four beta-like receptor subtypes have been very nicely characterized (Lind et al., 2010). Our analysis identified the same distribution of receptor types in *C. borealis* and *H. americanus* as was described in the barnacle – one alpha- and four beta-like subunits (Table 2.3; Figure 2.7). However, there were no particularly conserved motifs that resulted in a clustering of decapod and barnacle receptor subtypes to converge on a common nomenclature for these receptors; the four beta-like receptors in barnacle most closely related one another rather than subtypes across species. As a result, we have simply named these β-like octopamine receptors subtypes with ascending numbers and in the style of the descriptions given to those identified in *B. improvisus* (*Bi*): e.g. *Cb-Oct-αR, Cb-Octβ-R1, Cb-Octβ-R2, Cb-Octβ-R3,* and *Cb-Octβ-R4.* However, we do not mean to imply direct orthology between *BiOctβ-R1* and *Cb-Octβ-R1*, for example. Finally, we also identified a

single putative tyramine receptor in both of our species (*Cb-Tyr-R* and *Ha-Tyr-R*), and these sequences are most similar to the tyramine/octopamine receptor reported from *Macrobrachium rosenbergii*, which is most effectively activated by tyramine (Jezzini, Reyes-Colón, & Sosa, 2014; Reyes-Colón et al., 2010).

Histamine is also a major neurotransmitter in invertebrates, but sequence information for histamine receptors in crustaceans has yet to be described. In vertebrates, these receptors belong to a family of G-protein-coupled histamine receptors. In invertebrates, the actions of histamine are transmitted exclusively through ionotropic histamine receptors, as no metabotropic histamine receptors appear to be present in invertebrates (Roeder, 2003). We discovered putative orthologs for 3 distinct ionotropic histamine receptors in both *C. borealis* as well as *H. americanus* (*HisR1-3*) and a fourth sequence was found in *H. americanus* that did not have an obvious match from *C. borealis* (Figure 2.7). In sequence comparisons with *Drosophila melanogaster*, the closest orthologs are described histamine-gated chloride channels. This is consistent with the known physiological characterization of crustacean histamine gating of chloride channels (Claiborne & Selverston, 1984; McClintock & Ache, 1989). Because naming conventions are lacking for these receptors in crustaceans, we have simply assigned these transcripts ascending numbers (1-4) to distinguish between different putative histamine receptor subtypes (Figure 2.7; Table 2.3).

Metabotropic Glutamate Receptors

37

Metabotropic glutamate receptors (mGluRs) are seven-transmembrane domain proteins coupled to G proteins capable of controlling many cellular processes through signaling cascades. mGluRs can be classified into three primary classes (Ferraguti & Shigemoto, 2006): Class I consists of mGluR1 and mGluR5 and are coupled to phospholipase-C and utilize intracellular calcium signaling cascades; Class II consists of mGluR2 and mGluR3; and Class III consists of mGluR4, mGluR6, mGluR7, and mGluR8 and are negatively coupled to adenylyl cyclase activity. These receptors exist as either homo- or heterodimers on the cell surface, and it is the associated G protein alpha-subunit that determines which class the mGluR falls under (e.g. Class I associates with G_q and G_{11}). In crustaceans, mGluRs have a relatively short history of study, mostly owing to the fact that the metabotropic form of glutamate receptors had not been characterized until the late 1980s (Hirono et al., 1988). In crustacean preparations, it has been found that mGluRs play a role in rhythm generation in the stomatogastric ganglion (Krenz, Nguyen, Pérez-Acevedo, & Selverston, 2000; Pérez-Acevedo & Krenz, 2005). In our analysis of these transcriptomes, we found six mGluR sequences for each species, which covered all three primary classes of mGluRs (Table 2.4; Figure 2.8). We did not find mGluR6 and mGluR8 orthologs in either species. It should also be noted that *Ha-mGluR2* and *Ha-mGluR4* aligned more closely to one another than to their *C. borealis* counterparts. This discrepancy could be due to the relatively short partial sequence found for *Cb-mGluR2*; that is, only the first 200 amino acids were

found for *Cb-mGluR2*, while the *Ha-mGluR2* sequence found is 1027 amino acids long.

Ionotropic Glutamate Receptors

The most common excitatory neurotransmitter found in crustaceans at the neuromuscular synapse is glutamate (Takeuchi and Takeuchi, 1964). The three primary ionotropic glutamate receptors are NMDA, AMPA, and kainate receptors, named respectively after the agonists *N*-methyl-D-aspartate, α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid, and kainic acid that activate them. One of the initial characterizations of NMDA receptors from crustaceans was performed in the crayfish optic lobe (Pfeiffer-Linn and Glantz, 1991). Since then, NMDA receptors have been studied in crustaceans for their role in memory (Troncoso & Maldonado, 2002), axon-to-glial signaling (Gafurov, Urazaev, Grossfeld, & Lieberman, 2002), and central pattern generation (Cleland & Selverston, 1995). We were able to identify 4 separate NMDA receptors from both *C. borealis* and *H. americanus*, which fell into two primary categories: 1-like and 2-like (Table 2.4; Figure 2.8). We further identified these receptors into 1A, 1B, 2A, and 2B based on their pairing, but this is not meant to imply orthology with other naming schemes in other species. We also found one NMDA receptor in both *C. borealis* and *H. americanus* that we did not find a highly similar pair for, which we have included as *Cb-NMDA-2-like* and *Ha-NMDA-2-like*.

AMPA receptors have been virtually undescribed in crustacean preparations, which coincides with our results. In our analysis, we did not find any receptors that most closely resembled AMPA receptors. Known AMPA receptors blasted against our transcriptomes aligned best against the putative kainite-type receptors, which is unsurprising considering both AMPA and kainate are considered non-NMDA receptors. Kainate receptors have been implicated as modulators of synaptic transmission and excitability (Contractor, Mulle, & Swanson, 2011). In crustacean systems, kainic acid historically has been shown to elicit depolarizations at the crab neuromuscular junction (King & Wheal, 1984), as well as the crayfish neuromuscular junction (Shinozaki & Ishida, 1981). Five pairs of kainite-like receptors were found between the *C. borealis* and *H. americanus* transcriptomes (Table 2.4; Figure 2.8), falling into two discrete categories: kainate type-1-like and 2-like. Sequences were further subdivided into A, B, and C on the simple basis of pairing between the two species and not based on any specific orthology to other species.

Beyond the excitatory glutamate-gated cation channels (NMDA- and kainate-like), an inhibitory glutamate-gated chloride channel (GluCl) was also found for both *C. borealis* and *H. americanus* (Table 2.4). Soon after their discovery as extrajunctional receptors in locust muscles (Cull-Candy, 1976), GluCls were described as postsynaptic receptors in the crustacean stomatogastric ganglion (Marder & Eisen, 1984; Marder & Paupardin-Tritsch,

1978). In our analysis, we found a single GluCl transcript for each species, named *Cb-GluCl* and *Ha-GluCl.* The finding of a single channel is consistent with some other invertebrates, with a single GluCl also found in most insects (Semenov & Pak, 1999). Because of their distinct characteristics, the GluCl channels we identified were not placed on any of the trees shown in the Figures.

GABA Receptors

The neurotransmitter γ-amino butyric acid (GABA) has been studied in crustacean species for decades for its role in synaptic transmission and neural inhibition (Dudel, Gryder, Kaji, Kuffler, & Potter, 1963; Florey & Chapman, 1961; Iversen, Kravitz, & Otsuka, 1967; E. Kravitz, Potter, & Van Gelder, 1962). Interestingly, in invertebrates several different neural responses to GABA have been found that have distinct profiles from that of vertebrate GABA receptors (Beg & Jorgensen, 2003; Fuchs & Getting, 1980; Marder & Paupardin-Tritsch, 1978; Swensen et al., 2000). Further, it was originally in crustacean species in which the role of GABA as a neurotransmitter was first demonstrated (Dudel et al., 1963; Dudel & Kuffler, 1961). GABA receptors are classified into two major groups: $GABA_A$ type receptors, comprising receptor complexes that are part of a ligand-gated ion channels, or $GABA_B$ type receptors, G protein-coupled receptors that act via metabotropic signaling systems. $GABA_A$ receptors are pentameric transmembrane receptors responsible for fast, usually inhibitory synaptic currents, and heteromultimers of the individual subunit types can form distinct

channel properties in invertebrates (Gisselmann, Plonka, Pusch, & Hatt, 2004).

We identified orthologs of three GABA$_A$ type receptor subunits from both *H. americanus* and *C. borealis* (Figure 2.9; Table 2.3), including orthologs of *Drosophila* LCCH3-, RDL-, and GRD-like receptor subunits – and we have preserved naming conventions for these subtypes. Two GABA receptor subunits previously have been cloned from *H. americanus* (Xu, Hollins, Landers, & McClintock, 1998), and very recently a GABA$_A$ type receptor was identified in the crayfish, *Procambarus clarkii* (Jiménez-Vázquez, Díaz-Velásquez, Uribe, Arias, & García, 2016). Sequence comparison reveals these previously described sequences to be orthologous to the RDL-like receptor from our data. The metabotropic GABA$_B$ type receptors are GPCRs responsible largely for slower inhibitory synaptic effects, and functional GABA$_B$ receptors are heterodimers formed by GABA$_{B1}$ and GABA$_{B2}$ subunits (Padgett & Slesinger, 2010). We identified orthologs of both GABA$_{B1}$ and GABA$_{B2}$ subunits in both *C. borealis* and *H. americanus* (Table 2.3; Figure 2.9).

Acetylcholine Receptors

Acetylcholine receptors are classified into two family subtypes: nicotinic receptors (nAChRs), which are ligand-gated ion channels that are activated by nicotine; and muscarinic receptors (mAChRs), which are metabotropic GPCRs that respond to the agonist muscarine. Muscarinic acetylcholine receptors are further classified into subtypes based on the specific G-protein associated with

the receptor (Ren, Folke, Hauser, Li, & Grimmelikhuijzen, 2015). In our analysis, we discovered two discrete subtypes of mAChRs (Table 2.4; Figure 2.9) in both *C. borealis* and *H. americanus,* which is consistent with other arthropods (Collin et al., 2013). The A- and B-type mAChR are defined based on their differential sensitivity to muscarine (A-type is 1000x more sensitive than B-type), as well as the antagonist binding properties (atropine, scopolamine, and QNB block A-type but not B-type) that each receptor exhibits (Collin et al., 2013).

Nicotinic acetylcholine receptors are common throughout the invertebrate central nervous system, mediating largely fast excitatory neurotransmission (Florey, 1963; Hildebrand, Townsel, & Kravitz, 1974; Marder, 1976; Marder & Paupardin-Tritsch, 1980). For nAChRs in nervous systems of *C. borealis* and *H. americanus*, each species was found to have 1 β-subtype and 8 α-subtypes (Figure 2.9). A thorough characterization of invertebrate nAChRs has been performed for the snail, *Lymnaea stagnalis* (Nierop et al., 2006)*.* Our analysis revealed that few direct orthologous sequences occurred from snail (Mollusca) to crustaceans that would allow us to adopt the nomenclature put forward in *Lymnaea.* Therefore, the crustacean receptor subtypes were named based on their most similar mammalian counterpart, with the exception of the α16 subunit found, named so due to its similarity to *acr-16* found in *C. elegans* (Touroutine et al., 2005)*.* This subunit is most comparable to α7 in humans, but we found other sequences nAChRs more similar to human α7 than that of the putative α16.

Gap Junction Proteins (Innexins)

The proteins responsible for gap junctions in invertebrates are the Innexins (Phelan, 2005). A family of Innexins have previously been described for both *C. borealis* and *H. americanus* (Shruti, Schulz, Lett, & Marder, 2014). We include them here as characterized through the transcriptome analysis for completeness. Three full-length sequences were named Innexins 1-3 (Figure 2.10; Table 2.5) based on significant sequence similarity to coding sequences of Innexins from multiple other organisms, including other decapod crustaceans (Ducret et al., 2006). Innexins 4-6 (Figure 2.10; Table 2.5) were subsequently identified from our transcriptome analysis of *C. borealis* nervous system. We identified clear orthologs from lobster and crab for Innexins 1-4 and 6; but two Innexins showed enough dissimilarity to be classified separately, and these are classified as Innexin 5 in *C. borealis* and Innexin 7 in *H. americanus* (Figure 2.10). All of these identified Innexin sequences have the signature motif YYQWV in the second TM domain as well as a series of other conserved amino acid residues considered hallmarks of Innexins (Shruti et al., 2014).

Discussion

The era of modern genomics and high-throughput sequencing has revolutionized the study of neuroscience, and provided an opportunity for classic

physiology systems in the study of neural circuit properties to experience a

renewed level of impact. In particular, invertebrate model systems that historically

have been invaluable to our understanding of basic circuit properties, dynamics,

systems neuroscience, and neuromodulation now present themselves as novel

contributors to molecular neuroscience. In particular, classic preparations such

as the *Tritonia* swim system (Getting, 1983a), *Aplysia* feeding circuits as well as

the classic gill and siphon withdrawal reflex (Baxter & Byrne, 2006; Castellucci &

Kandel, 1976), and crustacean stomatogastric systems (Marder & Bucher, 2007)

revolutionized our understanding of neural circuitry. Each of these systems is the

renewed focus of genomic and transcriptomic approaches (Christie et al., 2015;

McGrath et al., 2016; Moroz et al., 2006; Senatore, Edirisinghe, & Katz, 2015) –

including this study – that promise to merge the unparalleled experimental

accessibility on the neurophysiology end of the spectrum with new molecular

tools to understand and manipulate these circuits. Decapod crustacean systems

have also been foundational in the understanding of modulation of behavioral

states. The earliest work implicating serotonin broadly in aggression can be

traced to seminal work in lobster behavioral studies (Kravitz, 2000; Livingstone et

al., 1980). GABA was first identified as an inhibitory transmitter in these same

decapod lobster species (Grundfest, Reuben, & Rickles, 1959; Otsuka et al.,

1966). Finally, the crayfish escape behavioral response has been a paradigm for

the true integration of neuroethological work across single neurons, neural

circuitry, behavior, modulation, and social status (Krasne & Edwards, 2002).

Therefore, the accessibility of a molecular perspective and toolset will allow

researchers to revisit these seminal works with greater potential to understand integrated nervous system function.

It can be challenging to stay up to date with all of the sequence data being published. To the best of our knowledge, there are a relatively small number of published transcriptome projects with decapod crustaceans as model systems. Sequence discovery projects in decapod crustaceans began with expressed sequence tag (EST) analyses over a decade ago (Stepanyan et al., 2006; Towle & Smith, 2006), and these have been used to examine olfactory receptor expression in the lobster system (McClintock, Ache, & Derby, 2006). Since then, a mixed tissue transcriptome sequencing projects have been performed from the spider crab, *Hyas araneus* (Harms et al., 2013, 2014) and spiny lobster *Sagmariasus verreauxi* (Ventura et al., 2015). Some of the most proactive work in crustacean transcriptomics lies in the area of neuropeptidomics for the discovery of neuroactive peptides and their receptors (Buckley, Fitzgibbon, Smith, & Ventura, 2016; Christie, 2016; Christie & Chi, 2015; Ventura, Cummins, Fitzgibbon, Battaglene, & Elizur, 2014), including a study that makes use of the *H. americanus* sequence data contained reported here (Christie et al., 2015). However, none of these studies provides analysis of ion channel or receptor sequence as contained within our study.

The most directly comparable transcriptome study was carried out in *H. americanus*, and used RNA-seq to examine differences in gene expression

46

across tissues, including nervous system, as well as responses to hormonal treatment and temperature perturbations (McGrath et al., 2016). While this study uses GO and other annotation approaches to identify channel and receptor groups that are differentially expressed, identification of these gene products is at the level of gene family. For example, McGrath et al. identify glutamate receptors as a broad group that are differentially expressed in tissues, but make no further distinction as to receptor subtype –reporting that they have 126 transcripts that map to a global glutamate receptor subtype group. This high number is likely due to identification of multiple short contigs of the same receptor sequence, as opposed to distinct receptor sequences, and could be explained by the lower N_{50} value of their transcriptome (1,289). In contrast, the N_{50} of our *H. americanus* transcriptome is 2,357, leading to longer contigs and often the identification of full-length coding sequence that greatly aided our ability to characterize and identify distinct channel and receptor subtypes. Nevertheless, to the extent these studies can be compared, we saw results largely consistent with the overall identifications provided. For example, both studies identified 5 octopamine receptor subunits, including at least 3 beta-type receptors. Further, all reported differentially expressed voltage-dependent ion channel and transmitter receptor subtypes were present in our data set. While there is no direct overlap in the goals of these studies, the characterization provided in our study will provide a reference by which subsequent expression analyses can be combined with thorough annotation to provide more insight into gene-specific changes related to neural function.

In this study we sequenced the nervous system transcriptomes for two highly utilized species in invertebrate neuroscience research: the Jonah crab (*Cancer borealis*) and the American lobster (*Homarus americanus*). Our sequencing, assembly, and annotation efforts have yielded an extensive set of sequence information from which we can begin to mine gene products critical to fundamental nervous system output: channels and receptors. This study represents the first attempt to characterize to this extent these critical building blocks of circuit function from these model systems. In doing so we have identified for the first time in these species previously undescribed channel and receptor families, as well as added to the incomplete characterization of amine receptors known to modulate both circuit function and behavior in these animals. This sequence information opens up these target proteins for use in gene manipulation techniques such as overexpression (MacLean et al., 2003) or RNA-interference mediate knockdown (Krenz, Parker, Rodgers, & Baro, 2015) to deeply interrogate circuit function. Finally, the stomatogastric system has been used extensively in computational studies that have revolutionized our understanding of circuit fundamentals, dynamics, and the role of variability in neuronal parameters in circuit function (Hamood & Marder, 2014; Marder, 2011; Marder, O'Leary, & Shruti, 2014; Temporal, Lett, & Schulz, 2014). These models have relied on biological data for identification of likely membrane conductances present in the networks. Molecular screening and quantitative assays of channel expression can effectively be used in concert with computational modeling (Ball,

Franklin, Tobin, Schulz, & Nair, 2010; Tobin, Cruz-Bermúdez, Marder, & Schulz,

2009) to generate better and more biologically realistic models with which to

uncover fundamental aspects of neural circuit dynamics.

A.

Cancer borealis

Homarus americanus
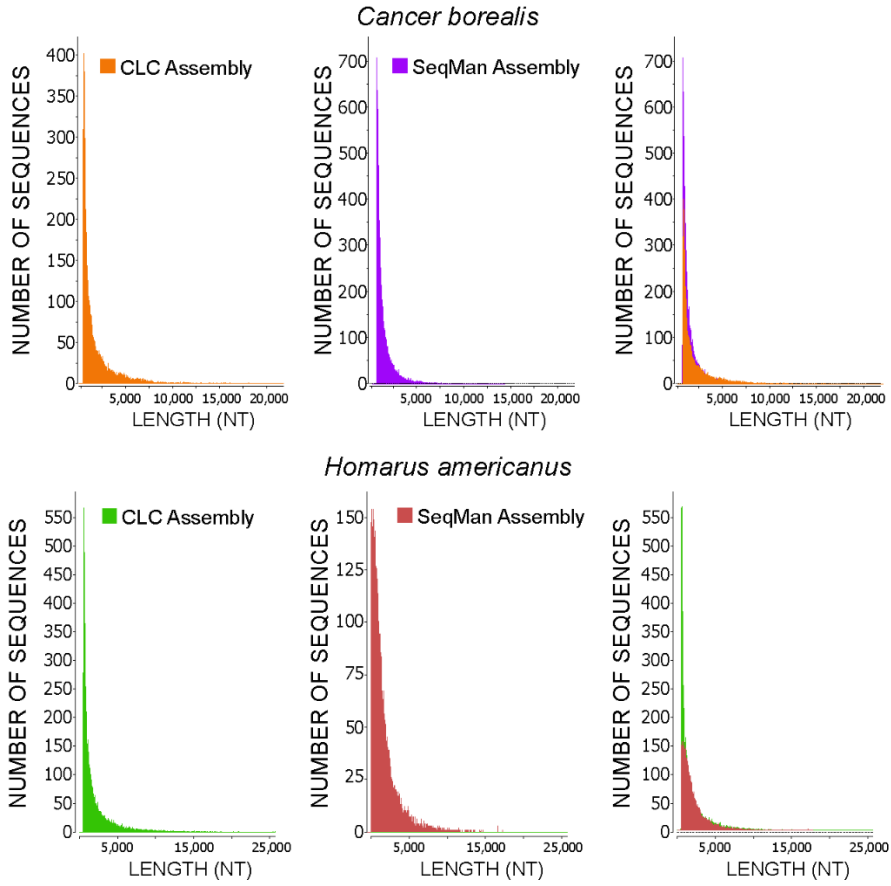
B.
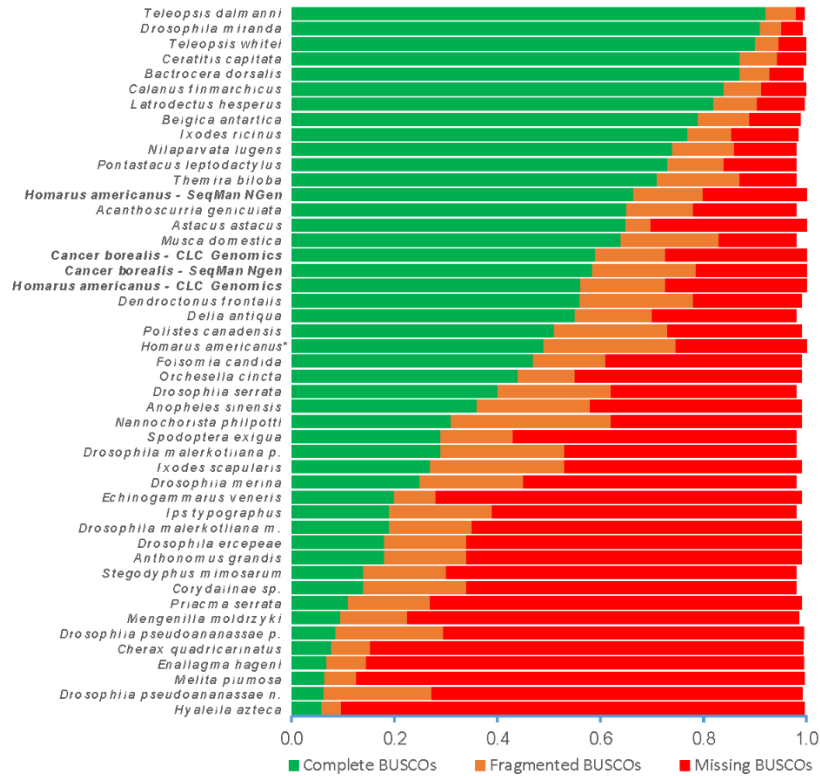
FIGURE 2.1: *Length distribution of the de novo assemblies and annotation coverage of the C. borealis and H. americanus transcriptomes.* **A.** Size distribution of contigs shown for two different *de novo* assemblies of the *C. borealis* and *H. americanus* nervous system transcriptomes. Each assembly is shown individually, and overlaid contig lengths are shown in the right panels. Assembly statistics are shown in Table 2.1. **B.** Horizontal stacked bar plots showing proportions of gene sets in BUSCO quality categories for the 4 different assemblies shown in panel A (CLC and SeqMan NGen, noted in **bold**). A previously published nervous system transcriptome from *H. americanus* (denoted by *) is also provided for comparison. Quality categories are as follows: i) Complete BUSCOs: genes that match a single gene in the BUSCO reference group; ii) Fragmented BUSCOs: genes only partially recovered with gene length exceeding alignment length cut-off; iii) Missing BUSCOs: non-recovered genes.
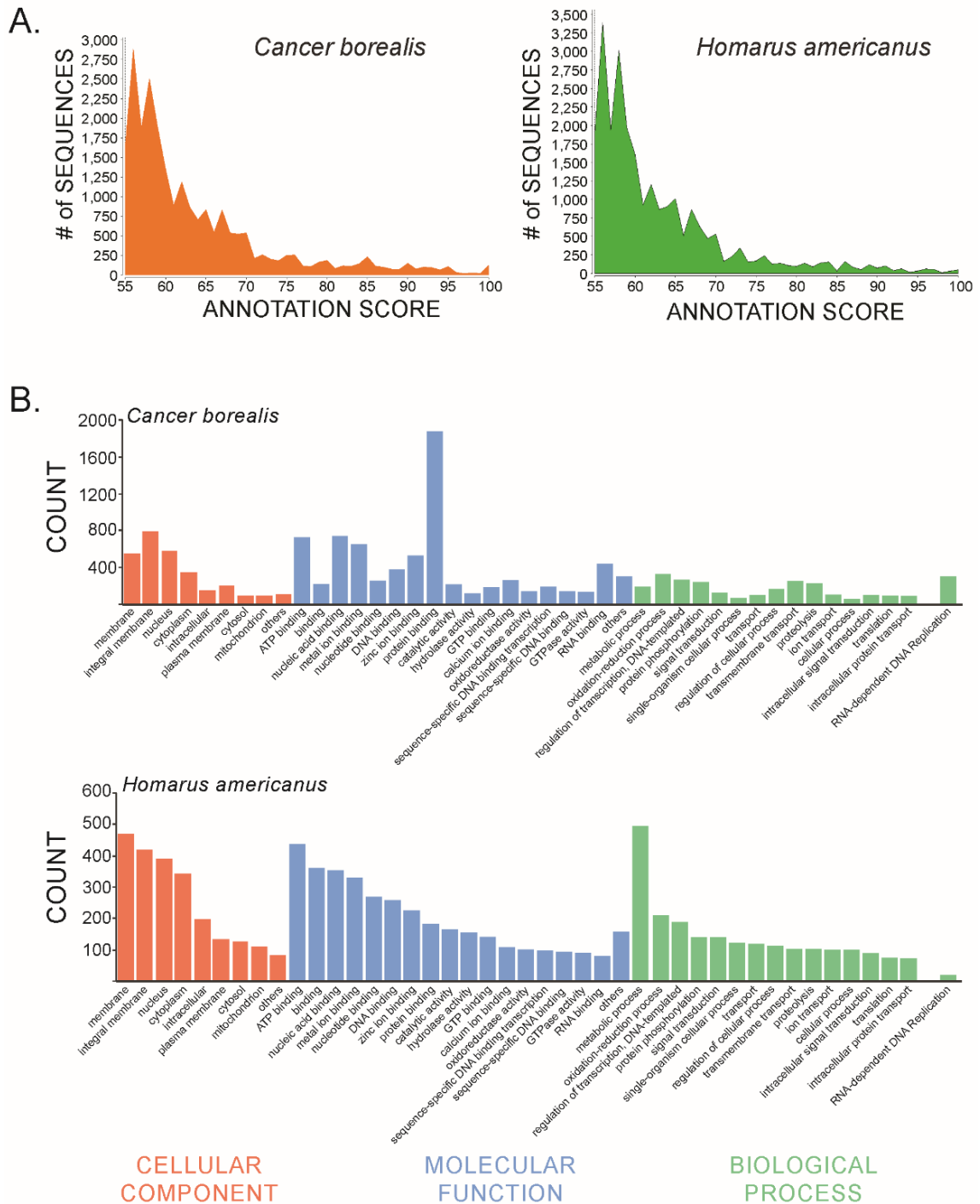
FIGURE 2.2: *Annotation-score distribution of C. borealis and H. americanus transcripts.* **A.** Distribution of annotation scores for all Gene Ontology (GO) terms assigned during the Blast2GO annotation process of the CLC assembled contigs.

B. Distribution of GO terms for *C. borealis* and *H. americanus*. Absolute values of GO annotation for the ontology categories of Cellular Component, Molecular Function, and Biological Process. Order was based on top GO counts for *H. americanus,* except for one case (RNA-dependent DNA Replication) due to high incidence in *C. borealis* biological process.
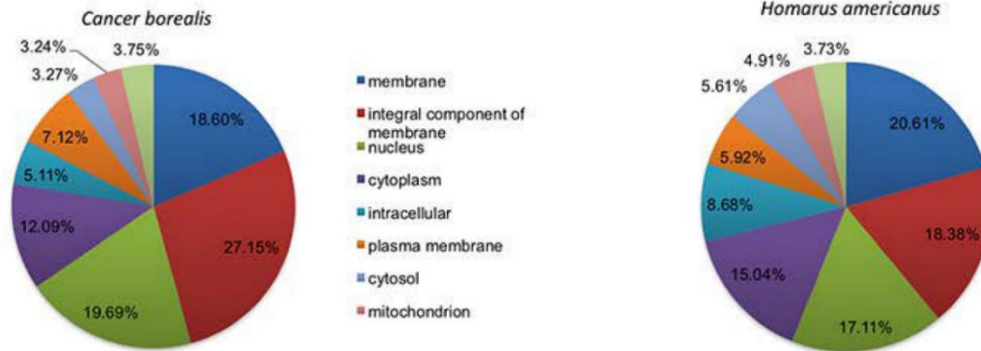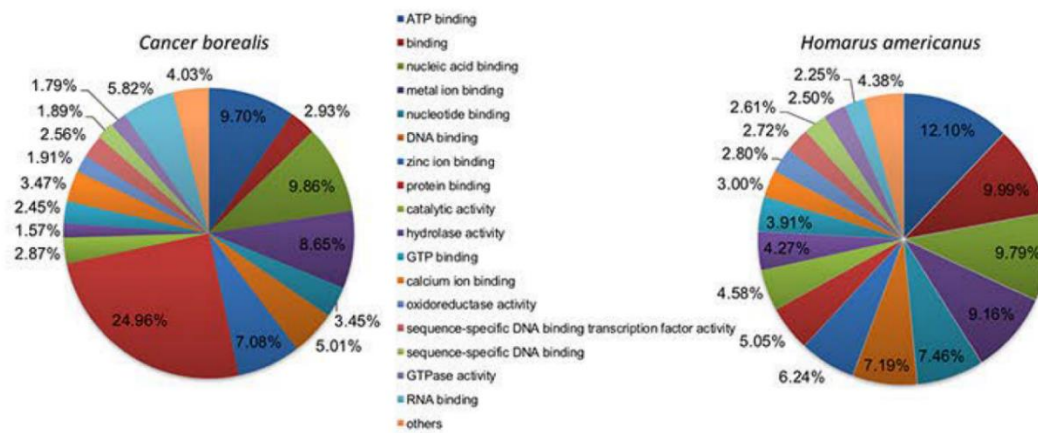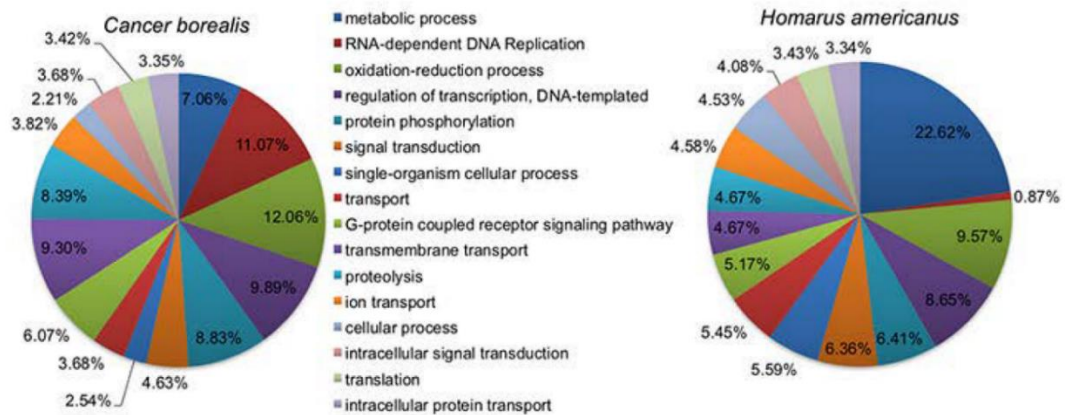
FIGURE 2.3: *Gene ontology (GO) comparison between C. borealis and H. americanus neural transcriptomes.* GO annotation categories Cellular Component, Molecular Function, and Biological Process were plotted as a percentage of their total annotation counts for each category.
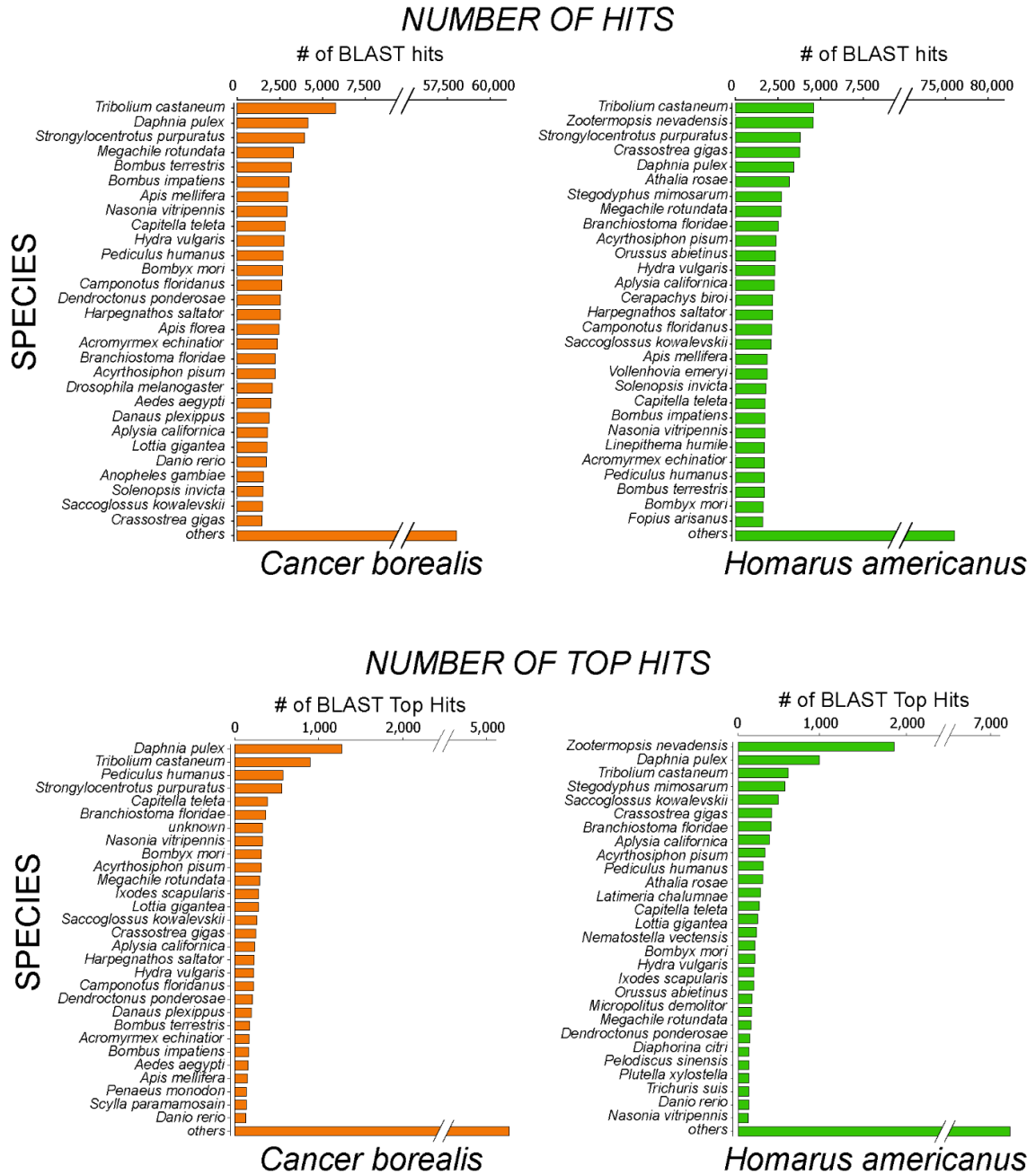
FIGURE 2.4: *Species distribution of blast hits of C. borealis and H. americanus neural transcriptomes.* Total hits and top-hit numbers for a given species from *C. borealis* and *H. americanus* transcriptomes.

A.
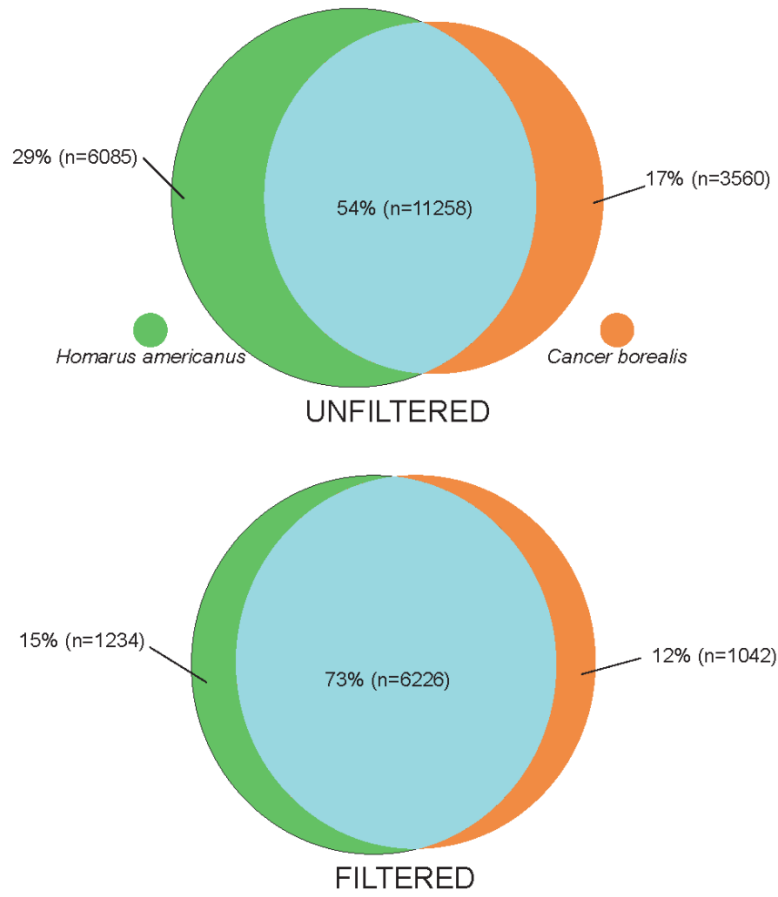
B.
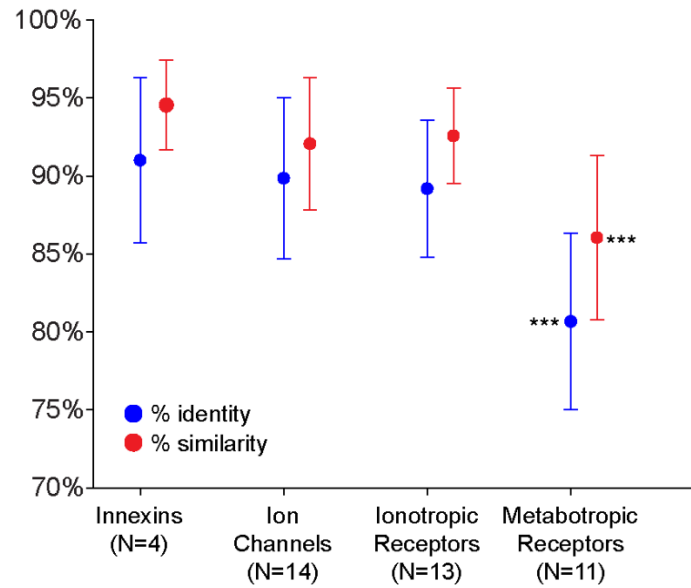
FIGURE 2.5: *Comparison of overlap of C. borealis and H. americanus neural transcriptomes.* **A.** VennBLAST comparison of *C. borealis* and *H. americanus* neural transcriptomes. Alignment of top hit sequence comparison was performed with a tblastx of both *C. borealis* and *H. americanus* against a common top hit species, *D. pulex*, allowing for a highly annotated crustacean database for reference. Filtering added another further stringency on top of that from the tblastx by requiring an amino acid identity percent of 70% and E-value threshold of 1.0e-5. **B.** Percent amino acid sequence identity (*blue points*) and similarity (*red points*) for selected neural function related gene products. For the specifically curated gene products described in the remainder of the study, we found very high (>90%) amino acid sequence identity and similarity between *C. borealis* Innexins (gap junction proteins), ion channels, and ionotropic receptors and the corresponding sequence in *H. americanus*. We saw a significant drop (one-way ANOVA with post-hoc *t*-tests) in similarity in sequences for metabotropic receptor subtypes. This indicates that channel proteins (including gap junction, voltage-gated, and ligand gated) show more highly conserved amino acid sequence than receptors that work via intracellular signal transduction cascades. *** indicates significant difference (P < 0.001, *t*-test) between metabotropic receptors and each of the other three groups. None of the other groups were significantly different from one another.

Voltage-gated
Na⁺/Ca²⁺ Channels

Ha-CaV3
Cb-CaV3
Ha-NaV
Cb-NaV
Ha-CaV2
Cb-CaV2
Ha-CaV1
Cb-CaV1

Ha-NALCN
Cb-NALCN
Na⁺ Leak Channels

Ha-KCNK2
Cb-KCNK2
Ha-KCNK1
Cb-KCNK1
Two-Pore-Domain
K⁺ Channels

Ha-KCNH2
Cb-KCNH2
Ha-KCNH1/EAG
Cb-KCNH1/EAG
Ha-KCNH3
Cb-KCNH3
EAG/ERG related
K⁺ channels

Ha-CNG-β1
Cb-CNG-β1
Ha-CNG-α1
Cb-CNG-α1
Ha-CNG-α2
Cb-CNG-α2
Ha-CNG-α3
Cb-CNG-α3
Ha-HCN/IH
Cb-HCN/IH
Cyclic-Nucleotide-Gated/
HCN channels

Cb-IRK
Ha-IRK
Inward Rectifier
K⁺ Channels

Ha-KCNT1
Cb-KCNT1
Ha-BKKCa
Cb-BKKCa
Na⁺/Ca²⁺ activated
K⁺ Channels

Cb-SKKCa
Ha-SKKCa
Small Conductance Ca²⁺
activated K⁺ Channels

Ha-KCNQ2
Cb-KCNQ2
Ha-KCNQ1
Cb-KCNQ1
KCNQ/Kv7 Type
K⁺ Channels

Ha-Shal
Cb-Shal
Ha-Shaker
Cb-Shaker
Cb-Shab
Ha-Shab
Ha-Shaw1
Cb-Shaw1
Ha-Shaw2
Cb-Shaw2
Shaker Family
Voltage-gated
K⁺ Channels

Ha-TRP-M3
Cb-TRP-M3
Ha-TRP-A-like
Cb-TRP-A-like
Ha-TRP-A1
Cb-TRP-A1
Ha-TRP-pyrexia
Cb-TRP-M1
Cb-TRP-M-like
Ha-TRP-M-like
Cb-TRP-V5
Ha-TRP-V5
Ha-TRP-V6
Transient
Receptor
Potential
Channels

FIGURE 2.6: *Ion channel subtypes and families identified in C. borealis and H. americanus transcriptome assemblies.* Amino acid sequence alignment was carried out using ClustalW, and subsequent trees were generated using ClustalW2_Phylogeny. We were not attempting to generate true phylogenetic relationships, but rather simply used tree-based analysis to identify putative gene families. Hence no bootstrap values are calculated. There was a 100% overlap in identified genes from both species, with the exception of TRP channels. A comprehensive list of channel types identified, their putative membrane currents, and accession numbers are provided in Table 2.2. Transcripts with the prefix "*Cb*" were identified from *C. borealis,* while those with "*Ha*" were identified from *H. americanus.*
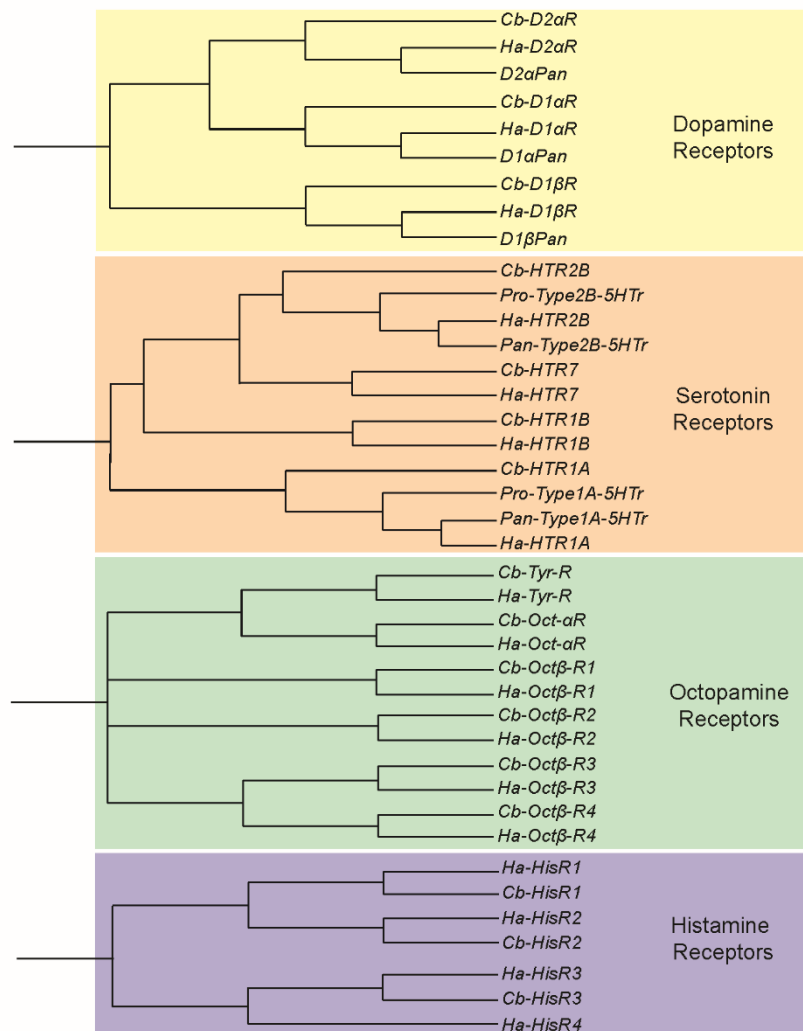
FIGURE 2.7: *Biogenic amine receptor subtypes identified in C. borealis and H. americanus transcriptome assemblies.* Trees were generated as described in Figure 6. Once again a 100% overlap in transcript types was found between the two species, with one exception – a histamine receptor (*Ha-HisR4*) was identified in lobsters that had no corresponding sequence from the crab transcriptome. In the case of serotonin (5HT) and dopamine receptor types, existing crustacean sequence from two different decapod species (*Panulirus interruptus* and

Procambarus clarkii) were used to help identify orthologs from *C. borealis* and *H. americanus.* These are also included in their respective trees as points of reference. *Pan-* represents *P. interruptus* while *Pro-* represent *P. clarkii.* A comprehensive list of amine receptor subtypes, including accession numbers, is provided in Table 2.3.
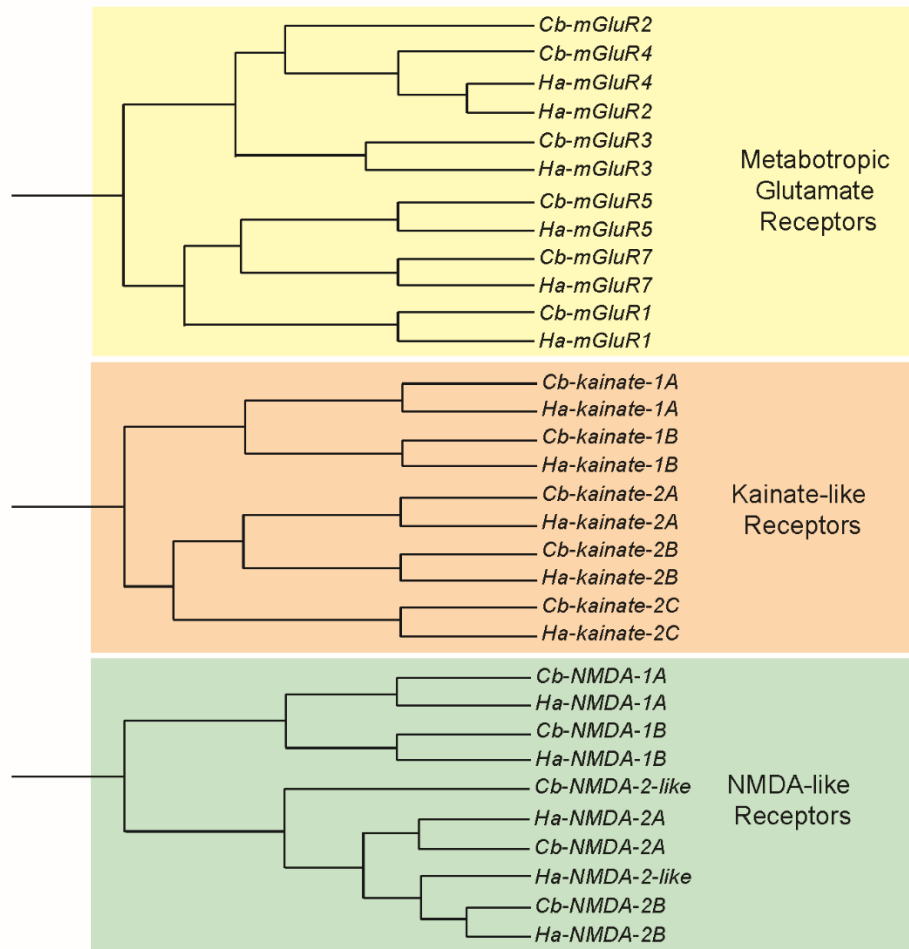
FIGURE 2.8: *Glutamate receptor subtypes identified in C. borealis and H.*

americanus transcriptome assemblies. Trees were generated as described in

Figure 6. We separated glutamate receptor subtypes into metabotropic (G-

protein coupled), and the ionotropic Kainate- and NMDA-like. A single glutamate-

gated chloride channel (*GluCl*) sequence was identified in both *C. borealis* and

H. americanus, and is not included as a member of a given receptor subclass in

the figure. A comprehensive list of glutamate receptor subtypes, including

accession numbers, is provided in Table 2.4.
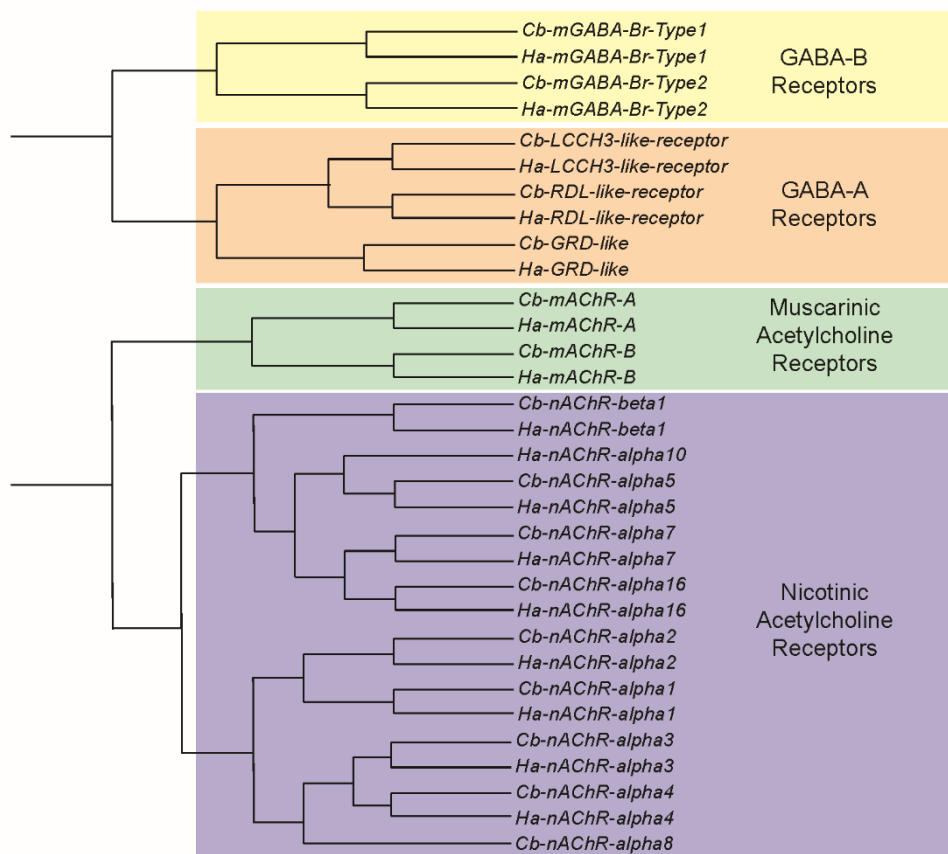
FIGURE 2.9: *GABA and acetylcholine receptor subtypes identified in C. borealis and H. americanus transcriptome assemblies.* Trees were generated as described in Figure 6. GABA and acetylcholine are both small molecule transmitters in crabs and lobsters. Both transmitters act through ionotropic and metabotropic receptor subtypes. Metabotropic GABA receptors (GABA$_B$-type) and ionotropic GABA subunits (GABA$_A$-type) were identified in both species. A comprehensive list of GABA receptor subtypes, including accession numbers, is provided in Table 2.3. Both nicotinic (ionotropic) and muscarinic (metabotropic) acetylcholine receptors were identified from both species, including one nicotinic beta-subunit and 8 alpha-subunit types. A comprehensive list of acetylcholine receptor subtypes, including accession numbers, is provided in Table 2.4.
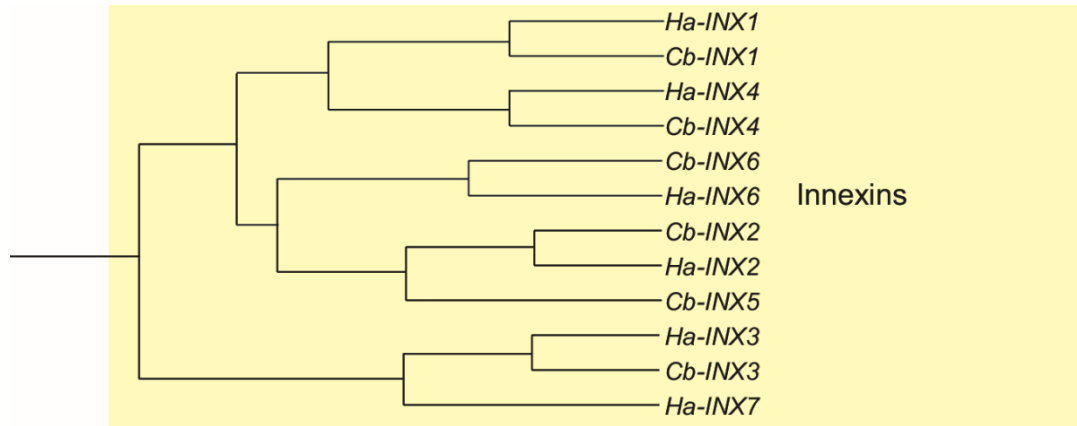
FIGURE 2.10: *Innexin subtypes identified in C. borealis and H. americanus transcriptome assemblies.* Trees were generated as described in Figure 6. Innexins are proteins responsible for gap junctions in invertebrates. Six distinct Innexin subtypes were identified in both *C. borealis* and *H. americanus* (see also [103]). Of these six, one from each species did not contain enough sequence homology to classify as the same type across species (*Cb-INX5* and *Ha-INX7*) hence are named as distinct subtypes. A comprehensive list of Innexin subtypes, including accession numbers, is provided in Table 2.5.

TABLE 2.1: *Overview of transcriptome assembly statistics for C. borealis and H. americanus.*

	C. borealis		**H. americanus**	
Raw reads	414,978,768		452,237,240	
Clean reads	391,060,790		426,712,238	
% Q Scores ≥ 30	92.96		92.72	
% GC	43.4		39.4	
Average clean read length (bp)	97.05		97.16	
Assemblers	CLC Genomics	SeqMan NGen	CLC Genomics	SeqMan NGen
Number of Contigs	42,766	67,380	60,273	45,043
N_{50} (bp)	2,178	1,239	2,357	2,258
N_{75} (bp)	1,058	763	1,169	1,085
Mean contig length (bp)	1,544	1,076	1,657	1,799
Longest contig (bp)	21,761	14,125	25,723	17,700
Shortest contig (bp)	454	451	453	450
Total assembled bases	66,058,464	72,508,321	99,847,148	81,065,797

TABLE 2.2: *Accession numbers for ion channels identified from transcriptome assemblies of C. borealis and H. americanus.*

Channel Family	Gene Name	Current/Channel Type	*C. borealis*	*H. americanus*
Voltage-dependent K+ Channels	*shaker*	Voltage-gated A-type potassium (I_A or K_v1)	FJ263946	KU702655
	shab	Voltage-gated delayed rectifier (I_{Kd} or K_v2)	DQ103255	KU702656
	shaw1	Voltage-gated delayed rectifier (I_{Kd} or K_v3)	KU681456	KU681443
	shaw2	Voltage-gated delayed rectifier (I_{Kd} or K_v3)	KU681455	KU681444
	shal	Voltage-gated A-type potassium (I_A or K_v4)	DQ103254	KU702654
	KCNQ1	Voltage-gated slow delayed rectifier (M-type or K_v7)	KU681453	KU681441
	KCNQ2	Voltage-gated slow delayed rectifier (M-type or K_v7)	KU681452	KU681440
	KCNH1/EAG	Ether-a-go-go type potassium (K_v10)	KU681458	KU681446
	KCNH2	Ether-a-go-go-related potassium (elk or K_v12)	KU681459	KU681447
	KCNH3	Ether-a-go-go-related potassium (erg of K_v11)	KU681460	KU681448
Other K+ channels	*BKKCa*	Large conductance (BK) voltage/Ca^{2+}-activated potassium	DQ103256	KU712072
	SKKCa	Small conductance (SK) Ca^{2+}-activated potassium	KU710383	KU712071
	KCNT1	Sodium-activated potassium	KU681454	KU681442
	IRK	Inward-rectifier potassium (IRK)	KU681451	KU681439
	KCNK1	Two-pore domain leak potassium (K_{2p})	KU681438	KU681450
	KCNK2	Two-pore domain leak potassium (K_{2p})	KU681437	KU681449
Ca2+ Channels	*CaV1*	L-type high-voltage-activated (HVA) calcium	JN809809	KU702651
	CaV2	P/Q-N high-voltage-activated (HVA) calcium	JN809808	KU702650
	CaV3	T-type low-voltage-activated (LVA) calcium	JN809810	KU702652
Na+ Channels	*NaV*	Voltage-gated fast sodium *para* type (Na_v)	EF089568	KU702653
	NALCN	non-selective sodium leak	KU681457	KU681445
Hyperpolarization-Activated/Cyclic Nucleotide Gated Channels	*HCN/IH*	Hyperpolarization-activated cyclic nucleotide-gated	DQ103257	KU712077
	CNG-Alpha1	Cyclic nucleotide-gated channel alpha 1	KU716097	KU712074
	CNG-Alpha2	Cyclic nucleotide-gated channel alpha 2	KU716098	KU712075
	CNG-Alpha3	Cyclic nucleotide-gated channel alpha 3	KU716099	KU712076
	CNG-Beta1	Cyclic nucleotide-gated channel beta 1	KU716096	KU712073
Transient Receptor Potential (TRP) Channels	*TRP-A1*	Transient receptor potential cation channel, subfamily A, member 1	KX037435	KX037441
	TRP-A-like	Transient receptor potential cation channel, subfamily A, member	KX037434	KX037440
	TRP-M1	Transient receptor potential cation channel, subfamily M, member 1	KX037436	-
	TRP-M3	Transient receptor potential cation channel, subfamily M, member 3	KX037433	KX037439
	TRP-M-like	Transient receptor potential cation channel, subfamily M, member	KX037437	KX037444
	TRP-V5	Transient receptor potential cation channel, subfamily V, member 5	KX037438	KX037445
	TRP-V6	Transient receptor potential cation channel, subfamily V, member 6	-	KX037443
	TRP-pyrexia	Pyrexia transient receptor potential channel	-	KX037442

TABLE 2.3: *Accession numbers for biogenic amine and GABA receptor*

subtypes from transcriptome assemblies of C. borealis and H. americanus.

Receptor Family	Gene Name	*C. borealis*	*H. americanus*
Octopamine/ Tyramine Receptors	*Tyr-R*	KU710373	KU712061
	Oct-αR	KU710375	KU712062
	Octβ-R1	KU710372	KU712063
	Octβ-R2	KU710374	KU712064
	Octβ-R3	KU710370	KU712065
	Octβ-R4	KU710371	KU712066
Dopamine Receptors	*D1αR*	KU710377	KU712059
	D1βR	KU710376	KU712060
	D2αR	KU710378	KU712058
Serotonin Receptors	*HTR1A*	KU710381	KU712070
	HTR1B	KU710382	KU712069
	HTR2B	KU710380	KU712067
	HTR7	KU710379	KU712068
Histamine Receptors	*HisR1*	KU716100	KU716104
	HisR2	KU716101	KU716106
	HisR3	KU716102	KU716103
	HisR4	-	KU716105
GABA Receptors	*mGABAr-1*	KU986868	KU986874
	mGABAr-2	KU986869	KU986875
	LCCH3-like	KU986871	KU986878
	RDL-like	KU986872	KU986876
	GRD-like	KU986873	KU986877

TABLE 2.4: *Accession numbers for glutamate and acetylcholine receptor subtypes from transcriptome assemblies of C. borealis and H. americanus.*

Receptor Family	Gene Name	*C. borealis*	*H. americanus*
Metabotropic Glutamate Receptors	*mGluR1*	KU986879	KU986885
	mGluR2	KU986880	KU986887
	mGluR3	KU986881	KU986888
	mGluR4	KU986882	KU986890
	mGluR5	KU986883	KU986886
	mGluR7	KU986884	KU986889
Kainate-Like Receptors	*Kainate-1A*	KX016772	KX016777
	Kainate-1B	KX016773	KX016778
	Kainate-2A	KX016774	KX016779
	Kainate-2B	KX016775	KX016780
	Kainate-2C	KX016776	KX016781
NMDA-like Receptors	*NMDA-1A*	KX016782	KX016787
	NMDA-1B	KX016783	KX016788
	NMDA-2A	KX016785	KX016789
	NMDA-2B	KX016786	KX016791
	NMDA-2-like	KX016784	KX016790
Glutamate-Gated Chloride Channel	*Glu-Cl*	KX059698	KX059699
Acetylcholine Receptors	*mAChR-A*	KX021822	KX021833
	mAChR-B	KX021821	KX021832
	nAChR-alpha1	KX021828	KX021840
	nAChR-alpha2	KX021827	KX021839
	nAChR-alpha3	KX021829	KX021841
	nAChR-alpha4	KX021830	KX021842
	nAChR-alpha5	KX021824	KX021836
	nAChR-alpha7	KX021825	KX021837
	nAChR-alpha8	KX021831	-
	nAChR-alpha10	-	KX021835
	nAChR-alpha16	KX021826	KX021838
	nAChR-beta1	KX021823	KX021834

TABLE 2.5: *Accession numbers for Innexin subtypes from transcriptome assemblies of C. borealis and H. americanus.*

	Gene Name	*C. borealis*	*H. americanus*
Innexins	*INX1*	JQ994479	KM984498
	INX2	JQ994480	KM984499
	INX3	JQ994481	KM984500
	INX4	KJ642222	KM984501
	INX5	KJ817410	-
	INX6	KJ817411	KM984502
	INX7	-	KM984503

CHAPTER 3

Molecular Profiling to Recapitulate Neuronal Cell Identity:

Lessons from small ganglia of the Crab *Cancer borealis*

Introduction

Since the earliest drawings by Santiago Ramòn y Cajal revealed a

glimpse of the breadth and diversity of neuronal phenotypes (Glickstein, 1991),

herculean efforts have been made to classify neurons into discrete, functional

types. These efforts are exemplified by the numerous and expansive enterprises

fueling cell type discovery and identification, including the NIH BRAIN initiative

(Ecker et al., 2017), the Neuron Registry Initiative from the International

Neuroinformatics Coordinating Facility (Hamilton, Shepherd, Martone, & Ascoli,

2012), the Human Cell Atlas Project from the Chan Zuckerberg Initiative (Regev

et al., 2017), the European Human Brain Project (Amunts et al., 2016), and the

Allen Brain Atlases from the Allen Institute for Brain Science (Sunkin et al., 2012;

Tasic et al., 2018). Unambiguous classification of neuronal cell types is a long-

standing goal in neuroscience with the aim to understand the functional

components of the nervous system that give rise to circuits and, ultimately,

behavior (Cuevas-Diaz Duran, Wei, & Wu, 2017; Luo, Callaway, & Svoboda,

2018; Masland, 2004; Stevens, 1998; Tasic, 2018; Zeng & Sanes, 2017). Beyond

that, agreement upon neuronal cell types provides the opportunity for crucial

technical improvements in reproducibility across investigations, allows for

evolutionary comparisons across species (Tessmar-Raible et al., 2007; Tomer,

Denes, Tessmar-Raible, & Arendt, 2010), and facilitates functional access to and

tracking of neuron types through developmental stages (Mehta et al., 2019). To

this end, attempts at defining neuronal identity have been carried out using

morphology, electrophysiology, gene expression, spatial patterning, and

neurotransmitter phenotypes (Boldog et al., 2018; Cadwell et al., 2015; Chung et

al., 2017; Ho et al., 2018; Parmhans, Sajgo, Niu, Luo, & Badea, 2018; Shrestha

et al., 2018; Södersten et al., 2018; Whitaker et al., 2011; Zeisel et al., 2015).

From the earliest experiments measuring gene expression in single neurons

(Eberwine et al., 1992), single-cell RNA-sequencing (scRNA-seq) experiments

using microarray (Esumi et al., 2008; Tietjen et al., 2003) or next-generation

sequencing (Tang et al., 2009) to capture neural transcriptomes has become the

method of choice for many investigations into neuron cell type. Advances in

library preparation and sequencing technologies have propelled an explosion of

molecular profiling studies seeking to use unique gene expression patterns to

discriminate neuronal types from one another, whether for discovery of new

types or further classification of existing ones (Cembrowski, Wang, Sugino,

Shields, & Spruston, 2016; Crocker, Guan, Murphy, & Murthy, 2016; Davie et al.,

2018; J. P. Doyle et al., 2008; Gokce et al., 2016; Haas et al., 2013; Macosko et

al., 2015; Poulin, Tasic, Hjerling-Leffler, Trimarchi, & Awatramani, 2016; Shin,

Ming, & Song, 2014; Tasic et al., 2016; Trapnell, 2015; Usoskin et al., 2015;

Wagner, Regev, & Yosef, 2016; Zhou & Matsunami, 2018).

Molecular profiling approaches to tackling the problem of neuronal cell

identity have many advantages: first, single-cell transcriptomic data contain

thousands of measurements in the form of gene products that can be used both

in a qualitative (in the form of marker genes) and quantitative (in the form of

absolute transcript counts) manner (Luo et al., 2018). Second, single-cell

sequencing allows for very high-throughput processing of samples with

hundreds, if not thousands, of single cell transcripts simultaneously using

barcoding techniques (Rosenberg et al., 2018). Third, these techniques can be

applied to species that lack well-annotated transcriptomic information, as the cost

to generate *de novo* reference transcriptomes has decreased dramatically in

recent years (Reuter, Spacek, & Snyder, 2015). However, scRNA-seq is not

without its limitations: low abundance transcripts are often missed or inaccurately

classified as differentially expressed (Sha, Phan, & Wang, 2015), and methods to

dissociate and isolate cells can alter their transcriptomic profiles (van den Brink

et al., 2017).

Classifying neurons into different major categories (such as excitatory vs

inhibitory, parvalbumin[+] vs parvalbumin[-], etc.) is an easier feat to accomplish

72

than separating neurons into smaller subclasses that are essential to understand circuit function. Even the sequencing of heterogeneous tissues from the central nervous system (CNS) can be used in conjunction with predictive modeling to reconstruct markers for major classes of CNS cell types, as has been done with oligodendrocytes, astrocytes, microglia, and neurons, in both humans and mice (Kelley, Nakao-Inoue, Molofsky, & Oldham, 2018). Minor classes of neurons may require greater depth of sequencing to separate them from one another within major classes. While neurons can sometimes be divided based on individual marker genes, neuronal gene expression is broad, and neurons are more likely to be defined by the expression of multiple genes rather than unique markers (Hobert, Carrera, & Stefanakis, 2010).

Many studies seek to put a number on the number of transcriptomically-defined cell types present in a given part of the brain. For instance, an initial study of the cell type diversity of the mouse primary visual cortex revealed 42 neuronal and 7 non-neuronal cell types (Tasic et al., 2016). More recent work from the same group updated the dataset to contain 133 transcriptomic cell types (Tasic et al., 2018). This raises the question of where is it appropriate to stop separating neurons into more and more discrete cell types? Work in the retina has lead the way as an example of generating a cell type consensus with an unknown endpoint. Multimodal information of retinal ganglion cell properties, including morphology, physiology, gene expression, and spatial patterning, has converged on over 65 cell types in the macaque fovea and peripheral retina

(Peng et al., 2019). However, not all systems have the same technical advantages as the retinal ganglion cells (such as uniform spatial patterning) that can be indicative of cell type, and multimodal information can be more difficult to obtain than high-throughput transcriptomic profiling methods. Therefore, the reliability of transcriptomic profiling for generating neuronal identity requires additional evaluation.

Is there a point at which the molecular profile of neurons is insufficient to delimit neurons into subtypes because neurons with variable underlying ion channel mRNA abundances can achieve the same functional output with striking reliability (Schulz et al., 2006)? In this study, we use systems in which neuronal identity has been unambiguously established neuronal cell types are already clearly defined and unambiguously identifiable from one animal to another using the stomatogastric ganglion (STG) and cardiac ganglion (CG) of the crab, *Cancer borealis*. Invertebrate systems have long benefited from the ability to unambiguously identify the same cell types from animal to animal allowing for circuits to be more easily teased apart. The neurons of the STG are identified according to the muscles they innervate, their well-characterized electrophysiological patterns, and synaptic connectivity (Selverston & Moulins, 1987), and the 5 large cells of the CG are identified according to their location. Accurately and reliably capturing single cells quickly without disrupting the molecular profile remains one of the greatest challenges in single-cell sequencing (Kolodziejczyk, Kim, Svensson, Marioni, & Teichmann, 2015). The

large size of STG and CG neuronal somata allows for quick, precise acquisition of identified neurons.

Beyond whole transcriptome profiling, studies also employ targeted qRT-PCR approaches to classify neuron types using molecular targets related to the output and function of neurons (Fuccillo et al., 2015; Kodama et al., 2012). Because ion channels, receptors, gap junctions, and neurotransmitters are some of the most salient features that give rise to unique neuronal properties (Tripathy et al., 2017), we used a multiplexed qRT-PCR approach to assay 65 transcripts related to neuronal function from the wide diversity of cell types found in the STG and CG of the crab *Cancer borealis* to attempt to recapitulate known cell identity. Machine learning algorithms are becoming increasingly useful for the identification of neuron types (Armañanzas & Ascoli, 2015). Using both a variety of commonly employed statistical analyses and machine learning approaches for identifying novel cell types, we measured the efficacy by which accurate cell identity could be achieved. We further expanded this into single cell whole transcriptome RNA-sequencing to ensure that we were not failing to capture unique molecular signatures that could be missed by the targeted qRT-PCR approach. Finally, we compare the efficacy of both approaches and contrast their strengths and weaknesses.

Materials and Methods

Tissue collection and RNA preparation

Adult Jonah Crabs were purchased from the Fresh Lobster Company (Gloucester, Massachusetts, USA) and Commercial Lobster (Boston, MA) and housed in filtered artificial seawater tanks chilled at 10–13°C on a 12/12 light:dark cycle until experimentation. Prior to dissection, crabs were put on ice for 30 minutes to induce anesthetization. The complete stomatogastric nervous system (STNS) was dissected and pinned out in a dish coated in Sylgard (Dow Corning) with chilled (12°C) physiological saline (composition in mM/l: 440.0 NaCl, 20.0 MgCl2, 13.0 $CaCl_2$, 11.0 KCl, 11.2 Trizma base, and 5.1 maleic acid pH = 7.4 at 23 °C. in RNase-free water. Following desheathing of the stomatogastric ganglion (STG), individual neurons were identified through simultaneous intra- and extracellular recordings, as is standard procedure for *C. borealis* (Hooper et al., 1986; Weimann, Meyrand, & Marder, 1991). Ten neuron types identified in the STG of *C borealis* were targeted for this study: PD (pyloric dilator), LPG (lateral posterior gastric), LP (lateral pyloric), IC (inferior cardiac), LG (lateral gastric), MG (medial gastric), GM, (gastric mill), PY (pyloric), VD (ventricular dilator), and DG (dorsal gastric). Identified neurons were extracted as previously described (Schulz et al., 2006). Briefly, a Vaseline well was constructed around the ganglion, in which protease (Sigma, St. Louis, MO) was added to disrupt connective tissue during a 10 minute incubation. The well was then thoroughly washed with fresh physiological saline to halt further enzymatic

76

activity, and a 70% solution of chilled ethylene glycol in saline was added to the well. The saline outside the well was replaced with distilled water, and the entire dish was frozen at -20°C for 30 minutes. This kept the STG neurons cold during the removal of identified neurons. Due to the large size of *C. borealis* STG neural somata (50-150 μM in diameter) (Ransdell, Faust, & Schulz, 2010), fine forceps were used to manually remove each neuron. Identified neurons (see Figure 3.1) were immediately placed in a cryogenic microcentrifuge tube containing 400 μL lysis buffer (Zymo Research) and stored at -80°C until RNA extraction. Total RNA was extracted using the Quick-RNA MicroPrep kit (Zymo Research) per the manufacturer's protocol.

Library Preparation and Single-Cell RNA-Seq

Library construction and RNA-sequencing services were carried out by the University of Texas at Austin Genomic Sequencing and Analysis Facility (Austin, TX, USA). Extracted single cell RNA from identified neurons from the STG were used to generate cDNA libraries using TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, California, USA). Libraries were sequenced in a paired-end 150 bp (2x150bp) configuration on the NextSeq 500 Illumina platform (Illumina, San Diego, CA, USA). Following sequencing, 40 identified neurons passed quality checks of at least 4 million uniquely mapped reads per sample comprising 11 PD, 11 GM, 8 LP, and 8 VD cell types.

Mapping and Differential Expression

The software package Kallisto (Bray, Pimentel, Melsted, & Pachter, 2016) (v0.43.1) was used in the quantification of RNA-seq abundances through the generation of pseudo-alignments of paired-end fastq files to the *C. borealis* annotated nervous system transcriptome (Northcutt et al., 2016). Bootstrapping of the quantification was performed iteratively for 100 rounds. Resulting counts were normalized through the transcripts per kilobase million (TPM) method. Differential expression analysis was carried out using the software package DESeq2 (Love, Huber, & Anders, 2014) using TPM normalized counts for each cell type.

Multiplex Primer and Probe Design

Multiplex primer and probe sequences targeting *C. borealis* genes of interest were generated using the RealTimeDesign™ qPCR Assay Design Software from LGC Biosearch Technologies (Teddington, UK) for custom assays. Multiplex cassettes were designed as a unit to ensure minimal interference in simultaneous qPCR reactions. Probe fluorophore/quencher pairs used in this study are as follows: FAM-BHQ1, CAL Fluor Gold 540-BHQ1, CAL Fluor Red 610-BHQ2, Quasar 670-BHQ2 and Quasar 705-BHQ2.

cDNA synthesis and pre-amplification

Following RNA extraction, individual neuron RNA samples were reverse transcribed into cDNA using qScript cDNA SuperMix (QuantaBio, Beverly, MA, USA) primed with random hexamers and oligo-dT per the manufacturer's protocol in 20 µL reactions. Half of each resulting cDNA pool (10 µL) was pre-amplified using PerfeCTa PreAmp Supermix (QuantaBio) with a 14-cycle RT-PCR reaction primed with a pool of target-specific primers in a 20 µL reaction per the manufacturer's protocol to allow for enough product to carry out 15 multiplex qPCR reactions per individual neuron sample. Amplified and unamplified target abundances were compared to ensure minimal amplification bias in the pre-amplification of samples (Figure S3.1).

Quantitative single-cell RT-PCR

Following preamplification of cDNA, samples were diluted 7.5x in nuclease-free water (150 µL final volume) to allow for the quantification of 73 unique gene products across 15 multiplex assays, each able to measure 4-5 different transcripts. Reactions were carried out in triplicate on 96-well plates with 10 µL reactions per well using a CFX96 Touch™ Real-Time PCR Detection System from Bio-Rad (Hercules, CA, USA). Cycling conditions for qPCR

reactions were as follows: 95°C for 3 min; 40 cycles of 95°C for 15 sec and 58°C

for 1 min. Fluorescent measurements were taken at the end of each cycle. The

final concentration of primers in each multiplex qPCR reaction was 2.5 µM and

0.3125 µM for each probe.

To quantify absolute mRNA abundances, standard curves were developed

for each RT-qPCR multiplex assay using custom gBlock gene fragments

(Integrated DNA Technologies, Skokie, IL, USA). Standard curves were

generated using a serial dilution of gBlock gene fragments from 1×10^6 to $1 \times$

10^1 copies for each reaction assay and were shown to be linear and

reproducible. Copy numbers were calculated using the efficiency and slope

generated from the standard curves and accounting for the 14-cycle

preamplification and subsequent cDNA dilution described above.

Statistical Analysis

For the qRT-PCR dataset, we preprocessed the data by applying median

imputation and removing zero variance transcripts. We tested centering and

scaling the data based either by cell or by transcript. We estimated the expected

number of clusters (presumably reflecting distinguishable cell types) using a

screeplot and the findKneePoint function from the sscClust library which provides

the knee point as the point furthest from a line defined by the first and last point.

We tested several clustering algorithms – k-means clustering, hierarchical clustering (using a variety of distance metrics, (euclidean, maximum, manhattan, canberra, binary, minkowski, correlation, uncentered) and clustering methods (ward.D, ward.D2, single, complete, average, mcquitty, median, centroid, ward.D2)), and SNN clustering (C. Xu & Su, 2015). We selected the best preforming clustering method based on the Jaccard index calculated against cell identity. These were selected for visualization.

We applied several classification methods to evaluate predictive power of mRNA abundances in ideal circumstances. Specifically, we tested elastic regression, k-nearest neighbors, linear discriminant analysis, neural network, multinomial neural network, random forest, support vector machine with a radial kernel, and support vector machine with a linear kernel. For each of these models we tested a variety of tuning parameters and selected the most effective parameter set before comparison with other methods. Methods were evaluated by using cross validation (with five folds) to produce the expected accuracy on new data.

The same approaches were applied to the single cell RNA-seq dataset. Due to improved performance in preliminary analysis we centered and scaled the data by cell. We then reduced the dimensionality either by filtering (retaining only those contigs that were differentially expressed at an alpha of 0.2 or 0.05) or by transformation with PCA and retention of the first 37 principal components were

retained (accounting for >99% of the variance). These three transformed

datasets were used to estimate the expected number of clusters, clustered with

Ward's (using correlation as the distance metric), and the same set of supervised

techniques described above.

Results

With the unambiguously identifiable neurons of the STG of *C. borealis*, we

sought to determine the extent to which the molecular profile of each neuronal

cell type recapitulated its known identity. The network connectivity of the STG

neurons that comprise the pyloric and gastric mill rhythms has been extensively

worked out (Figure 3.1A). Neuron types are identified based on simultaneous

intracellular (Figure 3.1B) and extracellular recordings, matching the timing of the

events of the stereotyped intracellular waveforms to the extracellular activity of

nerves that innervate muscles of the crab stomach. Of the 12 cell types in the

STG of *C. borealis*, we investigated 10 cell types (not AB or INT1), as well as

Large Cell (LC) motor neurons of the cardiac ganglion as an outgroup to

compare against neurons from a different network (Figure 3.1B).

Molecular Profiling of Single Identified STG and CG Neurons Using

Candidate Genes

To determine the molecular profile of individual identified neurons, we first

targeted transcripts that are among the most directly relevant to neuronal

function: ion channels, receptors, gap junction innexins, and neurotransmitter-

related transcripts. These cellular components are responsible for giving neurons

much of their unique physiological outputs, and as such we predicted that

correspondingly unique expression patterns for this gene set would be present in

each neuron type. Using multiplex qRT-PCR, we measured the absolute copy

number of 65 genes of interest from 124 individual STG neurons of 11 different

types (10 STG neuron types: PD, LPG, VD, GM, LP, PY, IC, LG, MG, DG and

motor neurons from the cardiac ganglion: LC, N = 10-15 per type). Expression

profiles for each neuron are represented as an unclustered heatmap of

abundances (Figure 3.2). Cells were arbitrarily segregated into two groups by

known transmitter phenotype (glutamatergic and cholinergic; (Marder, 1976;

Marder and Eisen, 1984)). This phenotype is well reflected in the expression of

transmitter-specific transcripts such acetylcholinesterase (*ACHE*), vesicular

acetylcholine transported (*vAChT*), choline acetyltransferase (*ChAT*), and

vesicular glutamate transporter (*vGlutT*). All of the cholinergic neurons show

correspondingly high levels of *ChAT* and *vAChT* as well as low levels of *vGluT*,

while glutamatergic neurons show the inverse expression pattern (Figure 3.2).

The innexins (*INX1-3*), sodium channel (*NaV*), and neurotransmitter-associated

transcripts were typically the highest abundance mRNAs across all cell types.

Additionally, individual neuron types can be seen to have distinct genes that are

more highly expressed relative to other cells. For example, both PD neurons

have higher expression of a glutamate-gated chloride (*GluCl*) and GABA-B receptor (*GABAB_R1*) relative to most neurons, while the delayed rectifier K+ channel *Shaw2* is almost uniquely expressed in LPG cells (Fig3. 2). We then used various methods of unsupervised hierarchical clustering to generate the "best" clustering of these cells based on *a priori* known cell type. This included substituting any missing values in the qRT-PCR data set via median interpolation within known cell type. Our best performing clustering analysis was generated with ward.D clustering and correlation as the distance metric (Figure S3.2). This analysis suggests that even with the most complete *a priori* knowledge of cell type in the analysis, there are limitations to determining cell identity solely through expression profiles of these channels and receptors.

Principal Component Analysis (PCA) is often used to determine if the variance seen among transcript abundances can be used to separate cells into discrete types. The first two principal components (PC1 and PC2) generated from the qRT-PCR data accounted for 31.8% and 16.7% of the variance, respectively (Figure 3.3). PC3 accounted for 10% of the variance across samples. We generated pairwise plots of all three PCs in attempts to visualize separation of samples into distinct cell types. The most consistent result across all comparisons was that LC neurons from the cardiac ganglion formed a cluster that had less overlap with STG neurons than STG neurons did with each other. Visualization of PC1 vs. PC3 and PC2 vs. PC3 also indicated that PD cells can be reasonably well distinguished from other cell types. However, overall it is

difficult to see how PCA would be effective in separating these 11 cell types based on the expression data at hand.

We then used k-means, unsupervised hierarchical, and SNN-Cliq clustering to generate unbiased clustering analyses based on expression of these focal genes of interest. Initial interrogation focused on data transformations with a fixed hierarchical clustering scheme (Ward's D2, Correlation dissimilarity matrix). Unscaled data as well as centered and scaled data (by cell or by gene) resulted in different hierarchical clustering patterns. Using unscaled data, hierarchical clustering performed rather poorly in terms of generating distinct clusters that match known cell identity (Figure S3.3A). Scaling by cell also resulted in poor hierarchical clustering performance (Figure S3.3B). The best performance – as assessed by Jaccard Index – was achieved by scaling data across genes, generating 8 distinct nodes with high bootstrap support in hierarchical clustering that capture some of the features of known cell identity (LC, IC, LG, LPG, VD, GM, LP, PD; Figure 3.4A). However, multiple cell types fall into clusters that either do not show any separation by neuron identity (DG, MG, PY) or show no bootstrap support based on hierarchical clustering (AU p-value = 0).

We sought to determine the upper bound for clustering performance with our dataset. If the molecular profile of STG neurons is capable of recapitulating the known anatomical and physiological cell identity, we would predict that 11

distinct and homogenous clusters would be generated. To determine the feasibility of clustering to sort cell types we tested 107 clusterings (varying clustering methods, distance metrics, and neighbors considered) for each data set. Each clustering was compared against the known cell identities with a Jaccard Index which ranges from 0 to 1 where 1 is perfect correspondence between clusterings. The best performing combination was data scaled by target and processed using Ward's D2 Hierarchical clustering with a correlation distance matrix (Jaccard = 0.636). By contrast the next best clusterings, Ward's D on raw counts using Canberra distance and Ward's D on data scaled by cell using Manhattan distance only achieved Jaccard indices of 0.495 and 0.487 respectively. The three least performant methods were Median hierarchical clustering with Canberra distance (0.084), hierarchical centroid clustering with Manhattan distance (0.859), and SNN-Cliq clustering with Binary distance and 9 neighbors (0.859). Examining the best performing clustering reveals that LP, PD, LG, IC, DG, LC, PY, GM, LPG, and VD separate fairly well. Nonetheless, a targeted approach of measuring electrophysiologically relevant transcripts is not sufficient to fully recapitulate cell identity, and all of these analyses failed to accurately recapitulate cell identity for all 11 cell types.

Given that an *a priori* known number of cell types represented in a sample is rare, we tested whether we would have arrived at the correct number of cell types in our sample had we been blind to their identity. We used the best performing transformations from the clustering analysis, i.e. data center and

scaled by gene and a correlation dissimilarity matrix, and 22 common cluster determination indices provided by the NbClust package. We allowed a minimum of 2 and a maximum of 20 clusters for this and later cluster determination analyses. The mode of 17 different methods of cluster determination estimated that the number of clusters was 8 (4 indices), followed by 17 (3 indices) (see Figure 3.4B). Only one of our tested indices, the Calinski-Harabasz Index, suggested that there were 11 clusters in this dataset. If a researcher were using any one of these, or a majority vote of several, the chance they would conclude the correct number of 11 clusters are present would be vanishingly low. Despite this, it is important to note that although we know these cells are physiologically distinct and make different anatomical projections, we do not know they are distinct at the level of transcript abundances.

To further address this point, we tested the ability of 7 SVML classification algorithms (generalized linear model (GLM), k-Nearest Neighbors (kNN), Neural Network (NN), Multinomial Neural Network (MNN), Random Forest (RF), Support Vector Machine with a linear kernel (SVML), Support Vector Machine with a radial kernel (SVMR)) to sort cells based on their transformed or untransformed mRNA abundances. Each model was tuned, and accuracy on new data was estimated using 5-fold cross validation. Performance was affected based on preprocessing (i.e. whether and how the data was centered and scaled) and varied substantially between algorithms, particularly when data was untransformed or centered and scaled by cell (see Figure S3.4). Across these

transformations the upper accuracy of assigning a cell to its known identity (96%) was achieved by at least one algorithm. The methods performed consistently well for the three transformations were GLM (min = 78%, max = 96%), RF (min = 70%, max = 96%), SVML (min = 74%, max = 96%), and SVMR (min = 67%, max = 96%). Overall, the top performing accuracy methods involved centering and scaling the data across genes, and yielded similar efficacies across algorithms (Figure 3.4C).

The overall result of the qRT-PCR analysis of ion channels, innexins, receptors, and neurotransmitter-related transcripts was that this approach only partially recapitulates known neuron cell identity, particularly if one uses methods where *a priori* knowledge of cell type (or even number of cell types) is unknown. However, it is possible that 65 genes of interest – even those directly related to the physiological output of the cells such as ion channels and receptors – are not sufficient to faithfully recapitulate neuronal identity. For example, expression profiles of Purkinje neurons have been shown to be distinct not based on their expression of genes directly related to neuronal output, but by carbonic anhydrases Car7 and Car8, the glucosyltransferase b3gnt5, centrosomal proteins Cep72 and Cep76, as well as an array of other transcripts of unknown function (Doyle et al., 2008; Heiman et al., 2008) (reviewed in (Fishell & Heintz, 2013)). To this end, we carried out single cell RNA-sequencing to capture the entire transcriptome of each neuron.

Single cell RNA-sequencing generated transcriptomes for PD (N=11), GM (N=11), LP (N=8), and VD (N=8) neurons that were subsequently mapped to the *C. borealis* nervous system transcriptome (Northcutt et al., 2016). We applied parallel clustering analyses carried over from the qRT-PCR data to the RNA-seq (data centered and scaled across genes/contigs, Ward.D2 and correlation). The lowest performance was when the entirety of the RNA-Seq data was used in these hierarchical clustering analyses. After removing transcripts for which there was no expression in any cell type, we were left with a data set that contained 28,459 distinct contigs in the complete RNA-seq data set. Using this complete data set in hierarchical clustering resulted in effectively 5 distinct clusters (Figure. 3.5A), with little pattern with respect to cell type whatsoever with PD cells being the best performers. All but two PD cells fell within one distinct cluster, albeit with a GM cell also identified in this cluster (Figure. 3.5A; *left*).

To investigate the maximum performance, we could expect to achieve with RNA Seq, i.e. if we knew *a priori* which transcripts were most predictive of cell identity, we repeated our analysis on raw and scaled data, selecting only differentially expressed transcripts. We selected two pools of differentially expressed raw transcripts: those with a 2-fold or higher level of expression difference and a q-value < 0.2 or q-value <0.05 (Figure 3.5A; *right*). However, it should be noted that of course differential expression (DE) analysis can only be

carried out by knowing the cell identity beforehand as it requires comparing samples grouped into conditions. Using DESeq2 with a q-value cutoff of 0.2 identified 137 differentially expressed transcripts. Hierarchical clustering on the q<0.2 data set resulted in enhanced clustering, but still failed to faithfully recapitulate cell identity. Hierarchical clustering was greatly improved by using the q<0.05 dataset (Figure 3.5A; *right*) but remains imperfect.

As with the RT-qPCR dataset, we tested a variety of preprocessing and clustering methods to find the best performing clusterings. In parallel to the qRT-PCR data analysis, we used the q<0.05 data (centered and scaled by gene/contig, Ward.D2 and a correlation dissimilarity matrix), and the 22 common cluster determination indices provided by the NbClust package (Figure 3.5B). Of these indices, 14 yielded usable results, with the mode for number of clusters at 3 (5 indices). The highest scoring clusterings were differentially expressed transcripts at q<0.05 using Ward's D with a correlation distance metric. In this instance the raw counts or counts centered and scaled by cell resulted in a Jaccard index of 0.738. The results of cluster estimation differed based on the preprocessing of the datasets. When data were centered and scaled by cell, the mode number of clusters estimated was 2 (4 indices) followed by 10 (3 indices). However, the raw (not scaled) differentially expressed data produced a mode at 4 (6 indices) double the next highest estimate, 3. The KL Index, CH Index, Hartigan Index, Duda Index, Pseudot2 Index, and the Ratkowsky Index agreed with our known number of cell types.

For the supervised machine learning we performed dimensionality reduction, as the full 28,459 contig data set repeatedly failed to yield results in our SVML analyses due to computational limitations. To capture the variation in the full dataset, we transformed the data with PCA and used the first 38 principal components, which accounted for over 99% of the variation. The SVML accuracies on the full RNA seq data set were extremely low, with a maximum accuracy of 67% and most with a median below 50% (Figure 3.5C). SVML accuracies improved by using differentially expressed data. Many of the methods and datasets used producing 100% accuracy for several cross-validation folds (Figures 3.5C; S3.3). It should be noted that no method classified all folds with complete accuracy – most methods ranged between 100% and either 75% or 83%. Curiously we observed that transforming the data drastically worsened the performance of NN (min = 22% max = 66%) and MNN (min = 22% max = 33%) when the data were scaled by cell (Figure S3.4).

Supervised machine learning approaches performed about as well on the q < 0.05 data as on the q < 0.2 data set. However, the median efficacy of some methods did improve over the former. As with the previous data set, performance of NN (min = 22% max = 57%) and MNN (min = 22% max = 33%) decreased when input data were scaled by cell. While these results are encouraging, even under optimal conditions (transcriptomic data, selection of transcripts by

differential expression, ability to use supervised methods) we were unable to consistently classify these neurons with 100% accuracy.

Comparison of qRT-PCR and RNA-seq Results

To ensure that the RNA-seq and qRT-PCR data were producing comparable expression results, we identified 4 different transcripts that were represented both in the DE data set from the RNA-seq and the qRT-PCR data set for the four cell types used in RNA-seq (PD, LP, GM, VD). Overall, there is very strong agreement in expression patterns for all four genes (Figure 3.6A), adding substantial confidence to the quality of both data sets with respect to capturing native expression patterns. However, we then extracted the RNA-seq expression data for all 65 of the transcripts used in the qRT-PCR data set. When we performed hierarchical clustering analysis using these 65 channel and receptor transcripts, the qPCR clusters with nearly 100% success (with the exception of 2 GM neurons) into nodes that contain the 4 known distinct cell types, while the RNA-seq dataset using the same transcripts fails to generate coherent cell type clusters (Figure 3.6B). As we examined this further, we realized that the four transcripts in Figure 3.6A (*ChAT, vAChT, NMDA2B, KCNK1*) represent highly expressed genes that were differentially expressed and showed consistent patterns between qPCR and RNA-seq methods. Other highly expressed transcript types were not differentially expressed (e.g. *NaV, INX1-3*), and therefore do not contribute strongly to distinguishing cell identity. Conversely,

many of the other transcripts in the qRT-PCR data set that were distinct across

cell types had very low levels of detected expression in the RNA-seq data set,

suggesting that the depth of RNA-sequencing could be a major factor in

determining or limiting our ability to use RNA-seq to recapitulate cell identity.

Discussion

Many projects currently attempting to describe neuronal cell types begin

with the acquisition of molecular profiles from populations of unidentified neurons

(Li et al., 2017; Tasic et al., 2016; Usoskin et al., 2015). Our results demonstrate

the limitations of unsupervised methods that rely solely on the molecular profile

to recapitulate neuron identity by showing how poorly predicted clusters

recapitulated known cell identity. However, that is not to say that the molecular

profile is not indicative of identity, as once we employed supervised methods

incorporating known cell identification and expected cluster numbers the results

greatly improved to capturing most of the known identities using only a focal set

of genes. Supervised clustering methods proved superior to unsupervised

algorithms in separating pyramidal neurons from interneurons in the mouse

neocortex based on morphological phenotypes (Guerra et al., 2011). Our results

argue that many studies would benefit from revisiting previous molecular

classifications and incorporating multimodal data alongside the molecular profile

to probe more deeply and separate classes with additional granularity.

LC neurons from the cardiac ganglion of *C. borealis* were arguably the easiest neuron type to be unambiguously separated from the neurons of the STG. The qPCR results molecularly profiled a focal set of genes related to neuron function, consistently separated out neurons that belonged to separate networks (i.e. LCs from STG neurons). However, distinguishing STG neurons from one another was only partially successful, reliably clustering 8 of the 11 cell types, and only after data were scaled by gene. PCA reinforced this with LC neurons having less overlap with STG neurons than STG neurons did with each other. This further supports the notion that broad classes of neurons are easier to separate from one another than neurons of the same class in a single network.

The ease of classification of cell identity is likely to be preparation specific. The retinal ganglia cells of mice show spatial patterning in which cells of the same type are distributed with exclusionary zones around them where no other cells of that type are found, while cells of different types do not exhibit spatial patterning and are more randomly distributed (Sanes & Masland, 2015). A molecular classification neurons in *C. elegans* found that anatomically distinct neurons have correspondingly distinct molecular profiles >90% of the time (Hobert, Glenwinkel, & White, 2016). However, 146 distinct molecular profiles were identified from the 118 anatomically distinct neuron classes, indicating the potential for molecular sub-classification. This classification relied on hierarchical clustering that was carried out solely on identified reporter genes (most

prominently transcription factors and GPCR-type sensory receptors) known to be differentially expressed across the 302 neurons of *C. elegans* from Wormbase.org (Harris et al., 2014) and not whole transcriptome molecular profiles. It is reassuring, but perhaps unsurprising, that the expression of a wide variety of reporter genes known to be differentially expressed across a population of neurons can recapitulate cell identity. But this relies on having a known set-point of neuron cell type to drive the choice of which targets to include in the hierarchical clustering, as differential expression analysis can only be carried out by assigning samples to different populations. The present study supports the results of the *C. elegans* molecular classification in that clustering is most reliable when differentially expressed targets are present, but we have shown that without knowing cell types *a priori by* additional criteria, molecular cell classification can generate unreliable results, particularly with neurons that belong to the same network.

Genes related to synaptic communication were critical for separating GABAergic neurons into different types (Paul et al., 2017). Sets of genes that are regulated together that can be thought of as a "gene battery" have also been shown to be indicative of cell type, such as transcription factors in *C. elegans* neurons (Deneris & Hobert, 2014). The expression of the focal gene set of ion channels, receptors, gap junction innexins, and neurotransmitter-related transcripts examined in this study discriminated neuron types well, using supervised methods taking into account known neuron identity, but this same

gene set did not perform well in the same cell types using RNA-seq , where sensitivity to detect low-abundance transcripts may have occluded cell-type specific expression patterns.

What are the sources of variability that could cloud molecular identification of neuronal cell identity? Most common high-throughput molecular profiling techniques require destructive sampling to acquire mRNA abundances, which generates only a snapshot of the profile at a single point in time. Gene expression has stochastic characteristics (Li & Xie, 2011; Raj & van Oudenaarden, 2008); transcription takes place not continually, but in bursts of expression (Wang, Ni, Wang, & Liu, 2018); and steady-state mRNA abundances are the result of rates of expression, but also degradation and mRNA stability (Perez-Ortin, 2008). Single cell transcriptomes can be altered biologically as a consequence of activity (Benito & Barco, 2015), injury (Llorens-Bobadilla et al., 2015), long-term memory formation (Crocker et al., 2016), differentiation (Olivera-Martinez et al., 2014), and aging (Davie et al., 2018; Moroz & Kohn, 2013), as well as being affected by technical noise (Brennecke et al., 2013).

Beyond cells existing in different types, cells also belong to different transcriptional states under certain conditions, with the major distinction between a cell type and cell state being that state is a reversible condition, where type is more constant and includes neuronal states (Tasic, Levi, & Menon, 2017). Neuron types exist in a continuum, exhibiting variation in expression patterns

within defined cell types, increasing difficulty in discreetly drawing the cutoff of one type from another (Cembrowski & Menon, 2018).

This study is not without its limitations. While we have deeply sampled the transcriptome of individual neurons, we have not measured other gene products not captured in the transcriptome that could drive unique identity, such as non-coding RNA species like miRNA and lncRNA (Sheng, Cao, Niu, Deng, & Zong, 2017). Epigenetic modifications have also been implicated in neuronal cell identity (Mito et al., 2018), which were not considered in this study. Further, we acknowledge that there are numerous other methods and statistical analyses being applied to molecular profiles to distinguish cell type. In this study, we chose to focus on the more commonly employed analyses (PCA, hierarchical clustering, machine learning algorithms) seen in the literature.

While molecular profiling makes a profitable starting point for discovering new neuronal cell types, it is far from the final step in the process. Much like a circuit's connectome alone is insufficient to predict network output and function (Gutierrez, O'Leary, & Marder, 2013), so too the transcriptome alone is insufficient to generate a definitive cell type beyond a superficial level. But also, like a connectome, the transcriptome is a necessary starting point in understanding the functional role of individual neurons and neuron types in a network.
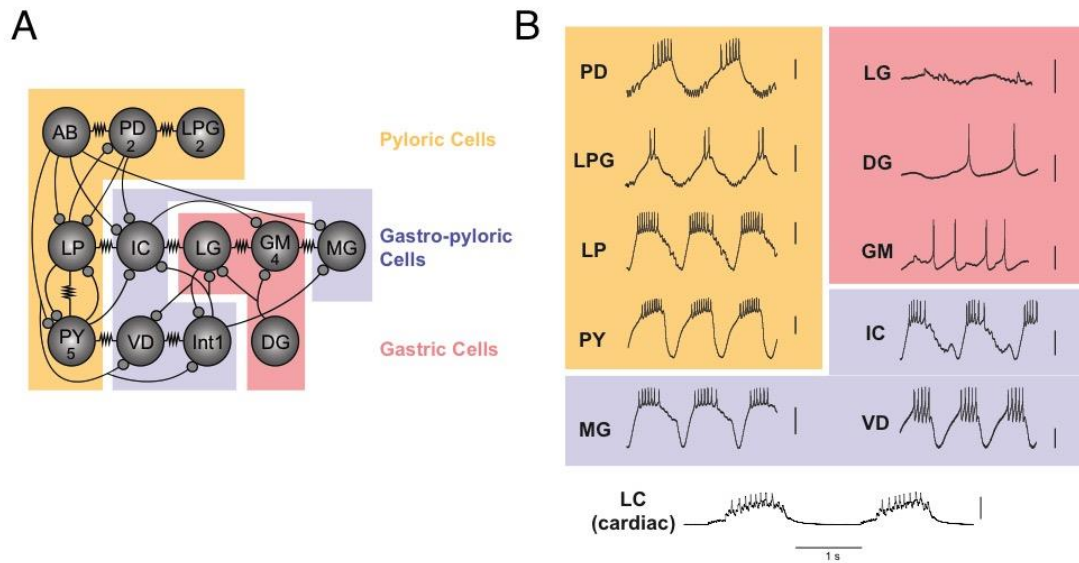
FIGURE 3.1: *The connectivity and intracellular waveforms of the cell types present in the stomatogastric nervous system.* **A)** Circuit map of the stomatogastric ganglion (STG). The STG contains 12 cell types that innervate the pylorus and gastric mill of the crab stomach. These cells are individually identifiable, and their chemical (closed circles) and electrical (resistor symbols) synaptic connections are all known. We used 10 of these 12 cell types (not AB or INT1) for this study, as well as motor neurons of the cardiac ganglion as an outgroup for comparison. **B)** Example traces taken from intracellular recordings of each the 11 identified neuron types used in this study. Neurons are involved in three different networks/circuits in the crab, *Cancer borealis*: the pyloric network (PD, LPG, LP, and PY; orange box), the gastric network (LG, DG, and GM; red box) and the cardiac network (bottom). Some neurons (IC, VD, and MG) participate in both gastric and pyloric network activity, and are noted in the purple

98

box. Large Cell (LC) motor neurons of the cardiac ganglion are used as an

"outgroup" to compare expression patterns of motor neurons from a distinct

ganglion (cardiac ganglion).

FIGURE 3.2: *Heatmap of expression levels for genes of interest measured by multiplex qRT-PCR.* We measured the absolute copy number of 65 genes of interest (rows) from 124 individual STG neurons (columns) of 11 different types (10 STG neuron types: PD, LPG, VD, GM, LP, PY, IC, LG, MG, DG and motor neurons from the cardiac ganglion: LC). Data are normalized as a column z-score for each cell, with highest abundance transcripts in red and lowest in blue.
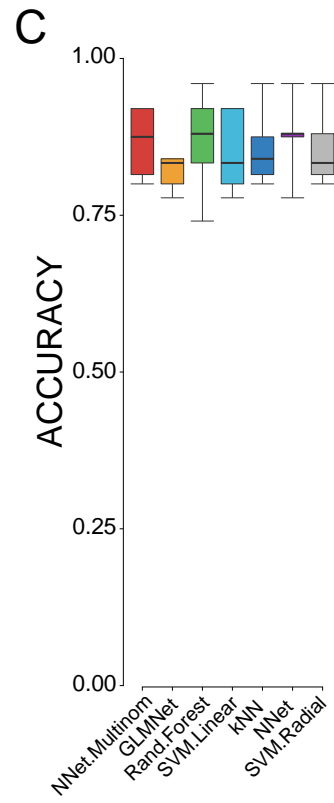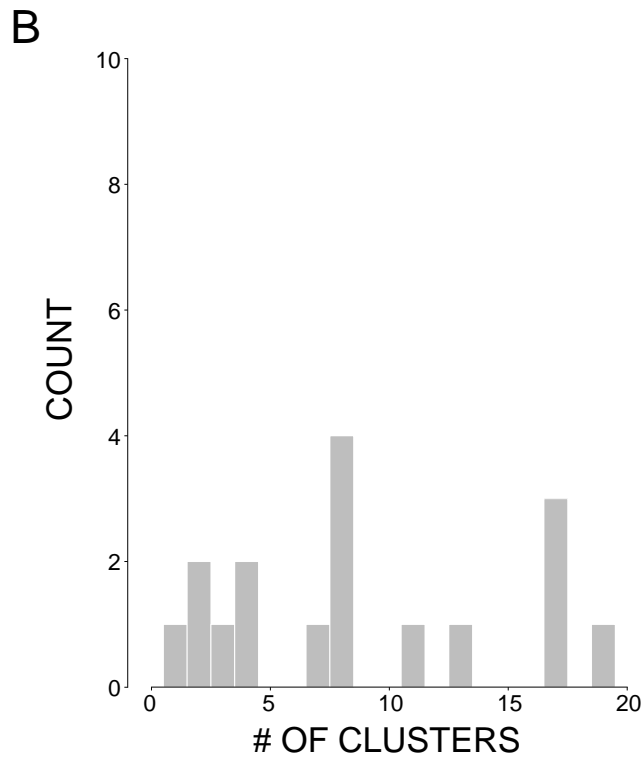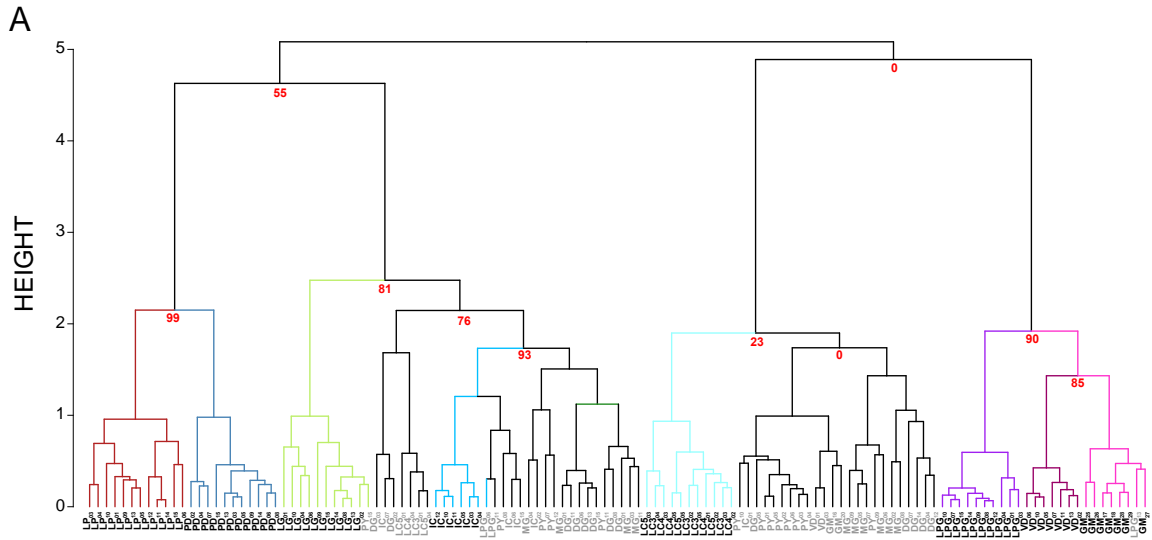
FIGURE 3.3: *Post-hoc recapitulation of cell identity via qRT-PCR expression with hierarchical clustering and supervised machine learning (SVML) algorithms.* **A)** Hierarchical clustering of cell type with correlation as the distance metric, and Ward.D2 as the clustering method for data centered and scaled across genes. Approximately Unbiased (AU) p-values for a given node are noted in red. Each node that has >80% support by AU p-value is color coded, and cell types that form a largely coherent group are noted in bold. Cells that do not appear to cluster by type are noted in gray. **B)** Histogram of predicted number of clusters based on 17 different prediction methods. Only 1 of 17 methods correctly predicted 11 distinct clusters that would represent the 11 different cell types in this assay. **C)** Accuracy of cell type prediction using 7 different methods of SVML for each of the data sets. Box and whisker plots show efficacy of each method across five cross-validation folds.
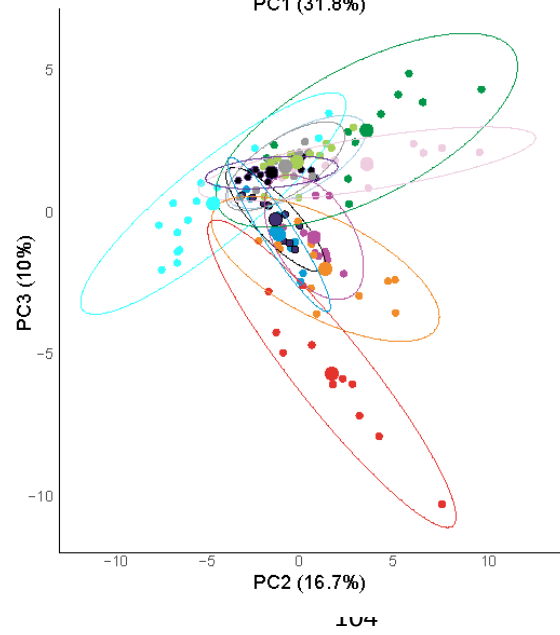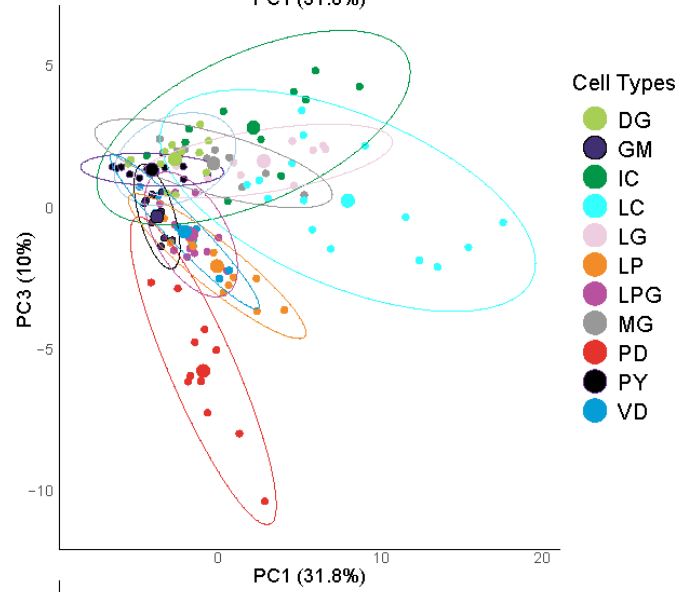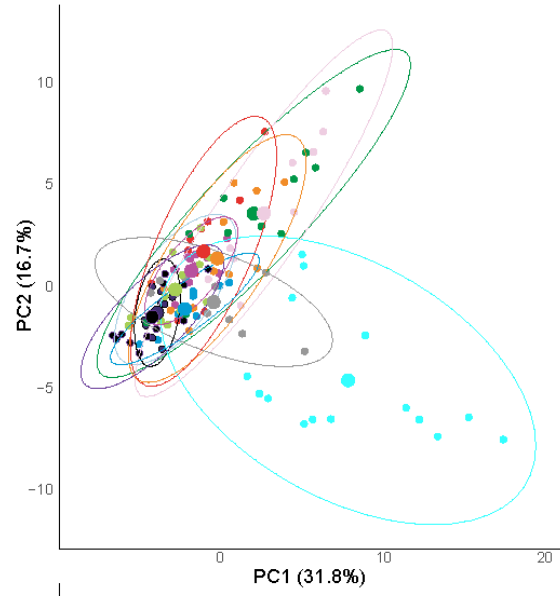
FIGURE 3.4: *Principal Component Analysis (PCA) for qRT-PCR data.* The first two principal components (PC1 and PC2) define the x-and y-axes in the top panel, while PC1 vs PC3 (middle) and PC2 vs PC3 (bottom) are represented in the remaining panels. PC1 accounted for 31.8% of the variance, PC2 accounted for 16.7%, and PC3 accounted for 10% of the total variance across samples.
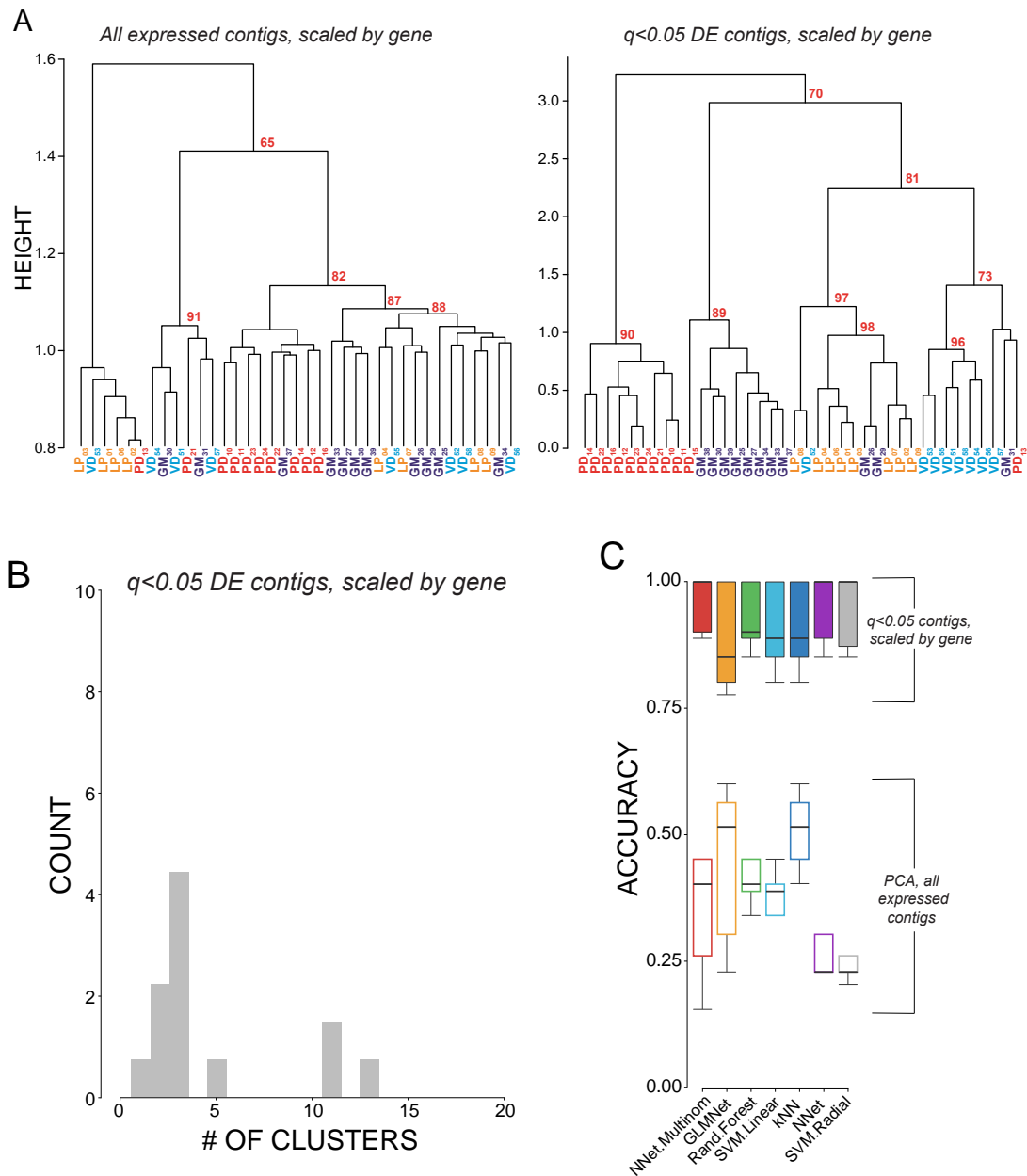
FIGURE 3.5: *Post-hoc recapitulation of cell identity via single cell RNAseq with hierarchical clustering and supervised machine learning (SVML) algorithms.* **A)** Hierarchical clustering of cell type with correlation as the distance metric, Ward.D2 as the clustering method, and data centered and scaled by gene/contig for all expressed contigs (*left*), and differentially expressed (DE) contigs at the q

< 0.05 level (*right*). Each cell type is color coded, and Approximately Unbiased

(AU) p-values are noted for each of the major nodes. **B)** Histogram of predicted

number of clusters based on 14 different prediction methods. None of the 14

methods correctly predicted 4 distinct clusters that would represent the 4 different

cell types in this assay. **C)** Accuracy of cell type prediction using 7 different

methods of SVML for each of the data sets. Box and whisker plots show the

efficacy of these methods to recapitulate cell identity from these two sets of

contigs as estimated by cross validation (5 folds). To assess the efficacy of these

methods without prior knowledge by which to select genes/contigs we used

principal component analysis (PCA) for dimensionality reduction (i.e. >28,000

contigs to 38 PCs) while retaining 99% of the variance.
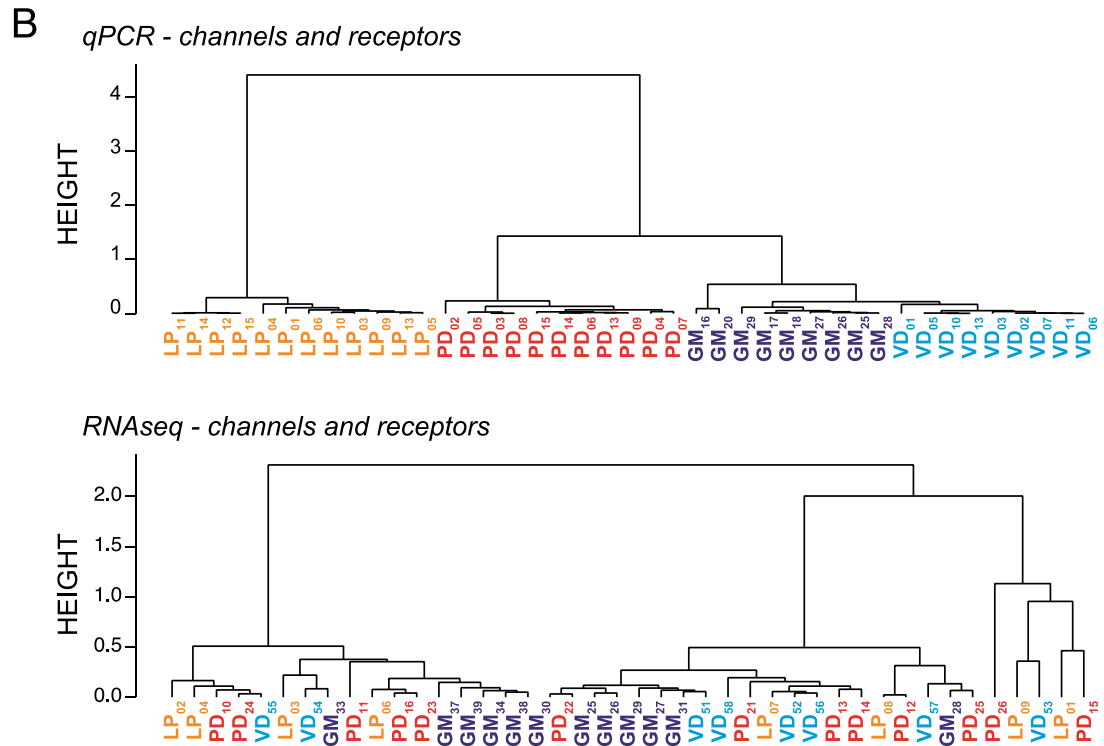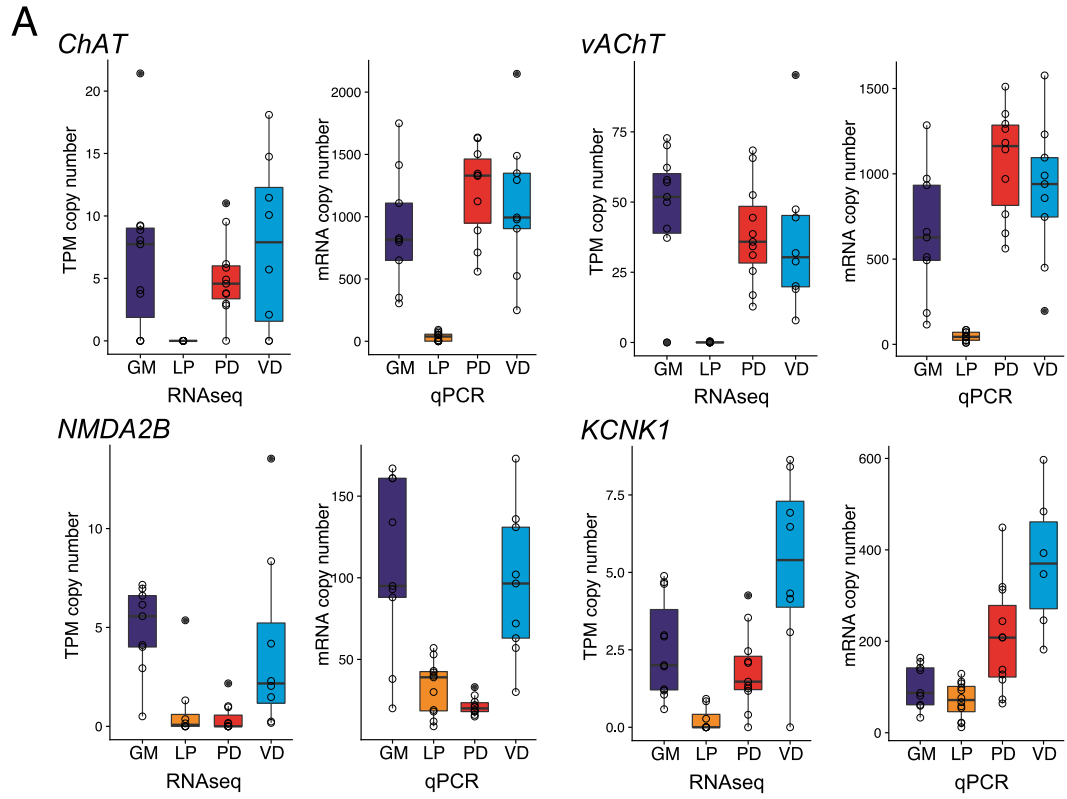
FIGURE 3.6: *Comparison of expression levels and clustering between qRT-PCR and RNAseq data.* **A)** Expression levels of 4 different genes (Choline Acetyltransferase: *ChAT*, Vesicular Acetylcholine Transporter: *vAChT*, NMDA Receptor Subtype 2B: *NMDA2B*, and K^+ Two-Pore-Domain Channel Subfamily K Member 1: *KCNK1*) between the RNAseq and qRT-PCR data sets. Data shown are medians, quartiles and each individual value from a given animal. Each individual data point is also represented as open circles. RNAseq data are presented as Transcripts Per Kilobase Million (TPM) while qRT-PCR data as absolute copy number per cell. **B)** Hierarchical clustering comparison between qRT-PCR (*top*) and RNAseq (*bottom*) for the same 65 genes represented in the genes of interest pool shown in Figure 3.1. Each cell type is color coded, and nodes on trees that contain homogeneous cell types are similarly colored.
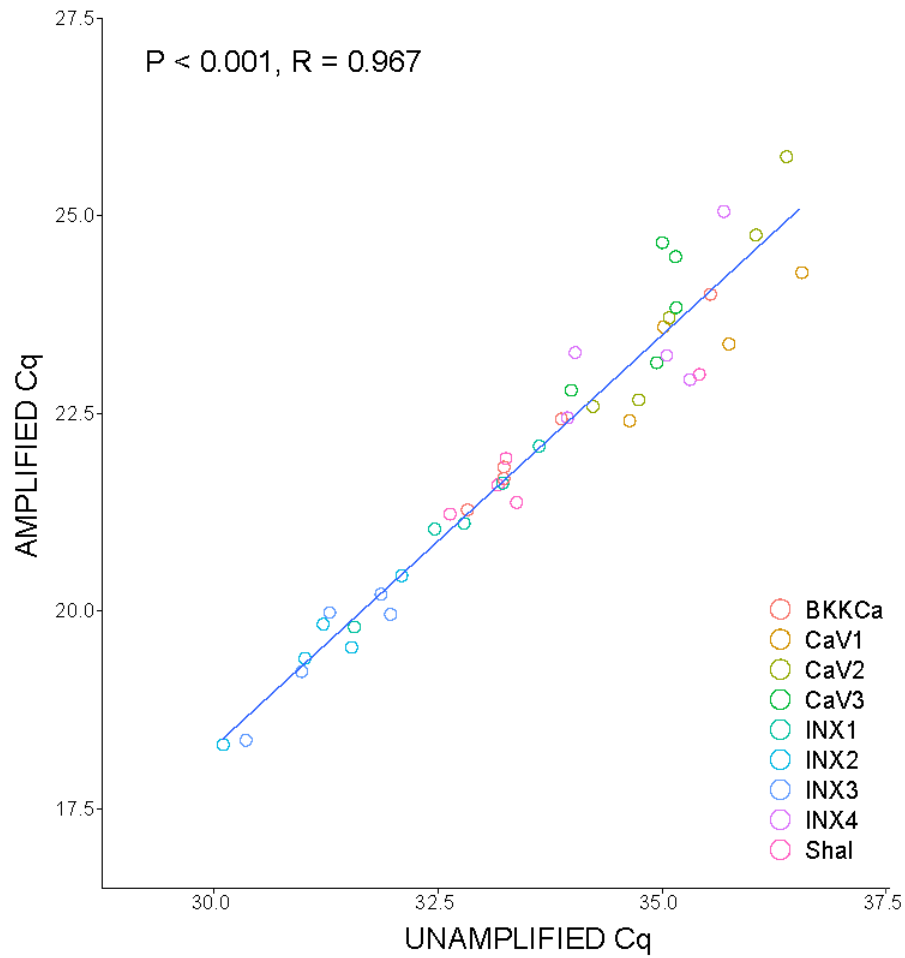
FIGURE S3.1: *Comparison of expression levels of the same single cell samples*

(N = 5) before (Unamplified) and after (Amplified) 14 cycles of preamplification of

cDNA. Data are shown as quantitation cycle (Cq) values. Statistics shown report

values for Pearson's Correlation test.
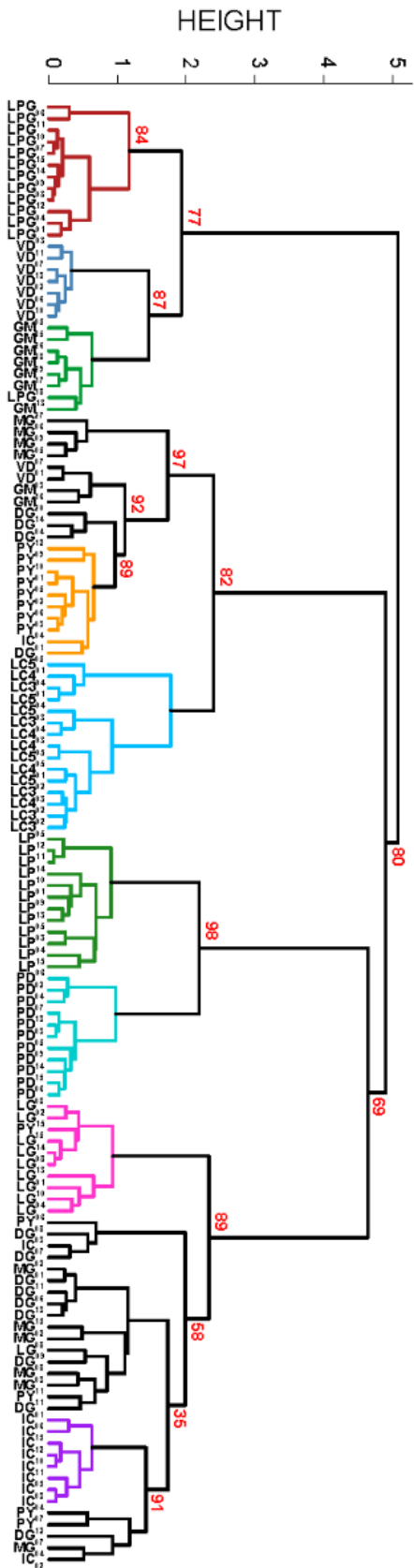
FIGURE S3.2: *"Best" clustering output for qRT-PCR analysis.* Hierarchical clustering of cell type with correlation as the distance metric, and Ward.D2 as the clustering method for data centered and scaled across genes and missing values replaced through median interpolation by identified cell type. Approximately Unbiased (AU) p-values for a given node are noted in red. Each node that has >80% support by AU p-value is color coded.
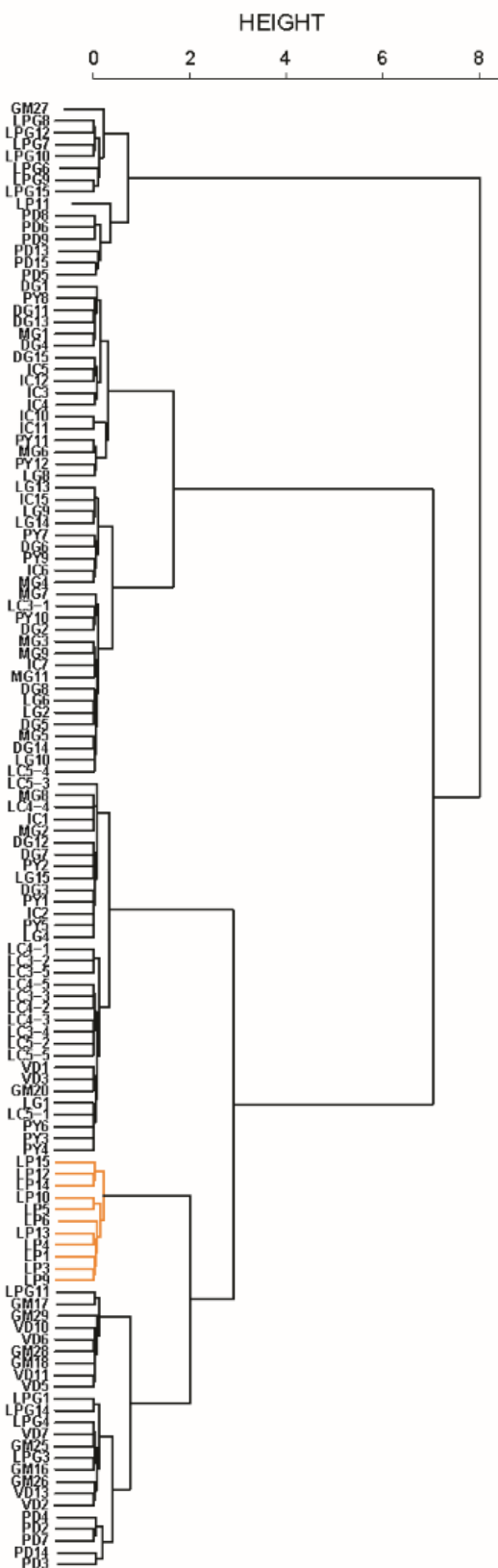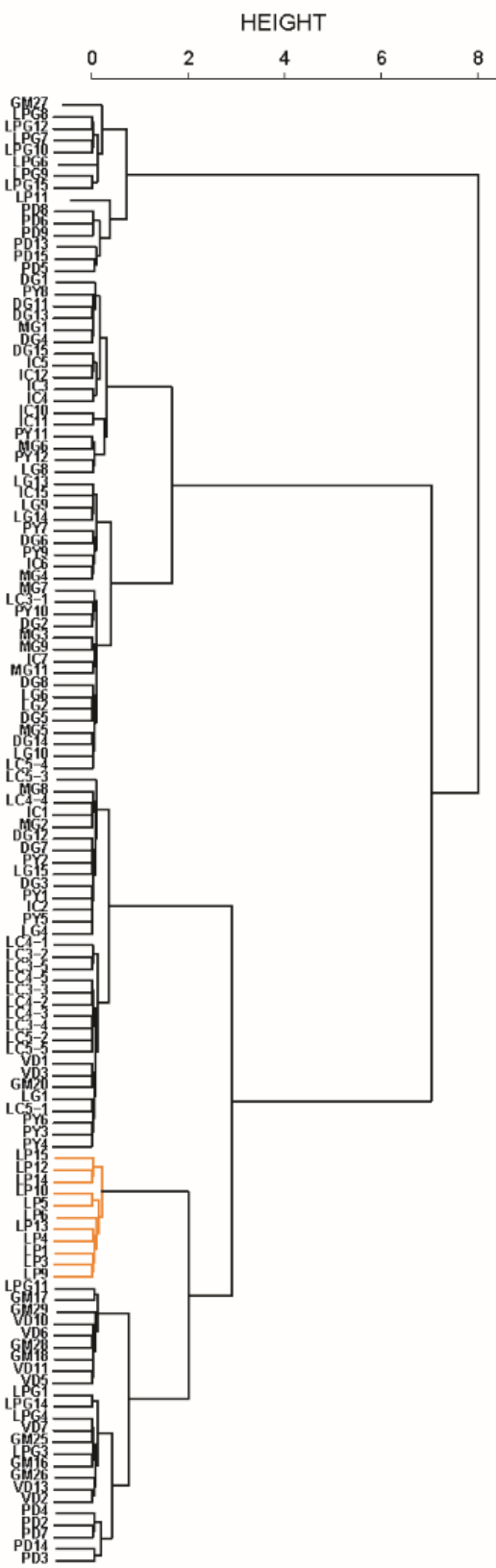
FIGURE S3.3: *Alternative clustering output for qRT-PCR analysis with different data preprocessing.* **A.** Hierarchical clustering of unscaled (raw) data and **B.** Hierarchical clustering of cell type with correlation as the distance metric for data centered and scaled by cell and missing values replaced through median interpolation by gene. Correlation was used as the distance metric, and Ward.D2 as the clustering method for both analyses. Cells are color coded by their *a priori* known cell identity. Only one strongly coherent cluster for LP (orange) was noted in these analyses.
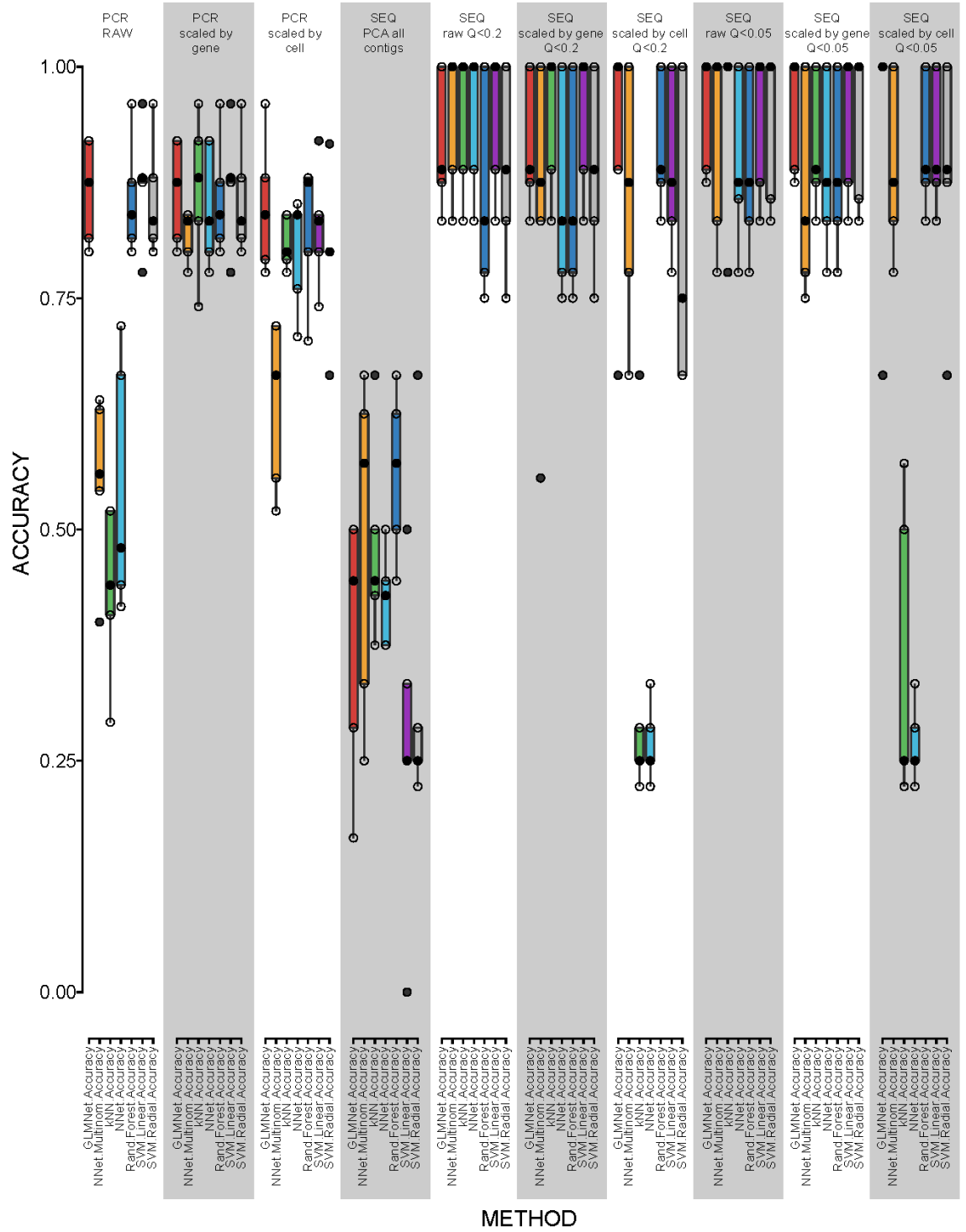
FIGURE S3.4: *Accuracy of cell type prediction using 7 different methods of*

SVML for all of the data sets. Preprocessing of data included no scaling ("raw"),

centering and scaling across cells ("scaled by cell") and centering and scaling

across target contigs or genes ("scaled by gene"). RNAseq data included all the expressed contigs with principle component analysis (PCA) for dimensionality reduction (i.e. >28,000 contigs to 38 PCs), as well as differentially expressed contigs at the q<0.2 and q<0.05 levels (2-fold difference). Box and whisker plots show the efficacy of these methods to recapitulate cell identity from these two sets of contigs as estimated by cross validation (5 folds).

CHAPTER 4

Corresponding molecular and electrophysiological properties of neurons from *Cancer borealis*

Introduction

Traditional electrophysiology preparations are witnessing a coming-of-age with expanded molecular toolkits to combine their experimental accessibility with broad high-throughput molecular techniques (Northcutt, Fischer, Puhl, Mesce, & Schulz, 2018; Northcutt et al., 2016; Smith et al., 2018; Tamvacakis, Senatore, & Katz, 2018). The crustacean stomatogastric ganglion (STG) has a long history of study, revealing general principles in circuit dynamics, neurotransmission, neuromodulation, rhythmic motor pattern generation, connectomics, and variability (Daur, Nadim, & Bucher, 2016; Marder & Bucher, 2007; Otopalik, Sutton, Banghart, & Marder, 2017). However, early molecular approaches in this system were limited to examining the expression of only a few genes across a few cell types (Schulz et al., 2006, 2007; Temporal et al., 2014a). Since the resulting activity of a neuron is the product of the expression of ion channel

genes, we sought to further probe the expression profile of neurons across a range of cell types to understand how these patterns of mRNA abundance relate to the properties of individual cell types. Distinct activity patterns have been shown to drive the expression of different ensembles of genes (Tyssowski et al., 2018). To drive a consistent activity pattern over the lifetime of a neuron requires homeostatic feedback to maintain excitability and synaptic strength (G. G. Turrigiano & Nelson, 2004). Therefore, one possibility of the mechanism by which neurons achieve maintenance of activity is though the differential regulation and expression of genes most directly related to neuronal function.

The STG of *Cancer borealis* is comprised of 25-26 neurons (Kilman & Marder, 1996) that control the musculature of the foregut. These neurons are unambiguously identifiable from animal to animal through simultaneous intra- and extracellular recordings, corresponding neural activity with anatomical projection (Maynard & Dando, 1974). The STG neurons are connected into two intertwined networks that drive the filtering (pyloric) and breakdown (gastric mill) of food through rhythmic generation of motor patterns. The pyloric rhythm is a fast triphastic motor pattern that relies on the intrinsically oscillating Anterior Burster (AB) interneuron electrically coupled with the two Pyloric Dilator (PD) neurons, all three together comprising the pacemaking kernel (Marder & Bucher, 2007). This kernel then relies on reciprocal inhibition through cholinergic and glutamatergic synapses with the lateral pyloric (LP) and pyloric (PY) neurons to generate the three phases of muscle contraction elicited from the rhythm. Also contributing to

the pyloric rhythm are the ventricular dilator (VD) and inferior cardiac (IC)

neurons that fire in-time with PY and LP, respectively. The gastric mill rhythm is a

slower rhythm that controls the actions of the teeth in the stomach and is

comprised of the lateral posterior gastric (LPG), lateral gastric (LG), medial

gastric (MG, dorsal gastric (DG) and gastric mill (GM) neurons. Interestingly,

many STG neurons can participate in both rhythms under various

neuromodulator states to generate distinct behaviors (Weimann et al., 1991).

With identifiable neurons comprising multiple networks with unique

electrophysiological properties, we sought to examine a broad range of the cell-

type specific steady-state mRNA expression profile for 11 cell types in the STG

across 57 genes related to neural function, including ion channels, receptors, gap

junction, and neurotransmitter-related sequences with absolute quantitation using

qRT-PCR. We further examined correlation networks of all 57 mRNAs to glean

information about potential co-regulation of expression patterns. Our results

demonstrate cell-type specific patterns of mRNA abundances that relate to

known the electrophysiological and network properties of STG neurons.

Materials and Methods

Tissue collection and RNA preparation

Adult Jonah Crabs were bought from the Fresh Lobster Company (Gloucester, Massachusetts, USA) and kept in chilled (10–13°C) and filtered artificial seawater tanks. Experiments were carried out in the spring and summer of 2017. Crabs were anesthetized on ice for 30 minutes prior to dissection. During experimentation, the stomatogastric nervous system (STNS) was dissected from the animal and secured on a Sylgard-coated dish (Dow Corning) with chilled (12°C) physiological saline (composition in mM/l: 440.0 NaCl, 20.0 MgCl2, 13.0 CaCl$_2$, 11.0 KCl, 11.2 Trizma base, and 5.1 maleic acid pH = 7.4 at 23 °C. in RNase-free water). Once the stomatogastric ganglion (STG) was desheathed, individual neurons were identified through matching the timing of activity during intra- and extracellular recordings using standard practices for *C. borealis* (Hooper et al., 1986; Weimann et al., 1991). The *C. borealis* neurons in targeted for this study include the STG neurons LP (lateral pyloric), PY (pyloric), PD (pyloric dilator), LPG (lateral posterior gastric), GM, (gastric mill), IC (inferior cardiac), LG (lateral gastric), VD (ventricular dilator), MG (medial gastric), DG (dorsal gastric), and AB (anterior burster), as well as the sensory neuron AGR (anterior gastric receptor) and the LCs (large cells) of the cardiac ganglion. Identified neurons were obtained in the same manner as previously described (Schulz et al., 2006). In brief, isolate a Vaseline well surrounding the ganglion was incubated for 10 minutes in ~1mg/mL protease (Sigma, St. Louis, MO), the well was washed with fresh physiological saline, and a 70% solution of chilled ethylene glycol in saline was applied. The dish was subsequently frozen for 30 min, and each identified neuron was manually placed in 400 μL lysis buffer

(Zymo Research) in a cryogenic microcentrifuge tube and stored at -80°C until RNA extraction. The Quick-RNA MicroPrep kit (Zymo Research) was used to extract total RNA per the manufacturer's protocol.

Multiplex Primer and Probe Design

Multiplex primers and probes used in this study were previously validated (Figure S4.1). In brief, the RealTimeDesign™ qPCR Assay Design Software from LGC Biosearch Technologies (Teddington, UK) was used to generate custom for custom assays targeting *C. borealis* genes of interest. Multiplex cassettes were designed to ensure minimal interference and off-target interactions in simultaneous qPCR reactions.

cDNA synthesis and pre-amplification

Following RNA extraction, total RNA from individual neuron samples was reverse transcribed into cDNA with qScript cDNA SuperMix (QuantaBio, Beverly, MA, USA) in 20 μL reactions using random hexamer and oligo-dT primers per the manufacturer's protocol. PerfeCTa PreAmp Supermix (QuantaBio) was used to pre-amplify half of each cDNA pool (10 μL) in a 20 μL reaction per the manufacturer's protocol using a 14-cycle RT-PCR reaction primed with a pool of target-specific primers. This preamplificaiton was necessary to generate a cDNA

pool large enough for carrying out 15 multiplex qPCR reactions per individual

neuron sample. Preamplificaiton bias was minimal, as determined by comparing

the preamplified and unamplified target abundances.

Quantitative single-cell RT-PCR

To quantify 65 unique gene products across 15 multiplex assays,

preamplified cDNA samples were diluted to 150 µL final volume in nuclease-free

water. Reactions took place in 10 uL triplicate reactions on 96-well plates using a

CFX96 Touch™ Real-Time PCR Detection System from Bio-Rad (Hercules, CA,

USA). The qPCR reactions were carried out with a hot start of 95°C for 3 min

followed by 40 cycles of two temperature steps: 95°C for 15 sec and 58°C for 1

min. Measures of fluorescence intensity were acquired at the end of each cycle.

The concentration of each primer in the multiplex reaction was 2.5 µM, and probe

concentrations were 0.3125 µM.

Standard curves were generated for each RT-qPCR multiplex assay to

quantify absolute mRNA abundances. Custom gBlock gene fragments

(Integrated DNA Technologies, Skokie, IL, USA) were used in serial dilutions

from 1×10^6 to 1×10^1 copies for each standard curves. Curves were verified to

be linear and reproducible. The efficiency and slope obtained from the standard

curves was used to quantify absolute mRNA copy number, with corrections for cDNA dilutions and the 14-cycle preamplification.

Correlation Network Analysis

Correlation analyses were generated using pairwise Pearson correlations to create correlation matrices using R (v3.5.2) in RStudio (v1.0.136). The R software package qgraph (v1.5) with the psych package (v1.8.12) was used to generate correlation network graphs from the pairwise correlation matrices with an R-value cutoff of 0.75. This cutoff was determined through examining thresholds of various stringencies to find which produced the most variability across the cell types examined in this study (see Figure 4.6). Network analyses were chosen for visualization over other methods due to the large number of pairwise correlations between 57 genes (1596 total pairwise correlations).

Results

The neurons of the STG of *C. borealis* drive the rhythmic contractions of the foregut, have unique electrophysiological properties, and are unambiguously identifiable based on simultaneous intracellular and extracellular recordings (Figure 4.1). The timing of intracellular neural activity is matched to recordings from nerves that project to muscles of the foregut, which allowed for the

identification of 11 types of STG neurons (anterior burster, AB; pyloric dilator, PD; lateral pyloric, LP; pyloric, PY; inferior cardiac, IC; ventricular dilator, VD; lateral posterior gastric, LPG; lateral gastric, LG; medial gastric, MG; dorsal gastric, DG; and gastric mill, GM) that were examined in this study. The intracellular activity of four of these neurons (PD, LP, GM, and VD) is shown in Figure 4.1 with corresponding extracellular activity on the lvn, dgn, and mvn nerves, respectively. By examining the molecular profile of these neurons, we determined cell-type specific patterns of mRNA abundance that supports known electrophysiological properties and neurotransmitter phenotypes, as well as overall mRNA correlation patterns.

Network Level Trends in mRNA Abundances

By broadly probing the expression of gene families across the entire array of single neurons in the STG, we reveal trends in mRNA abundance patterns that are relevant to the performance of the network as a whole. That is, we can examine members of each gene family related to neural activity to determine which are most widely used across STG neurons and which members are hardly present. Reasoning behind this examination is that different member subtypes examined in this paper could be present in higher abundances in other parts of the crustacean nervous system and not as relevant to the function of the STG.

Figure 4.2 shows the absolute mRNA abundances of STG neurons (N = 119) for all ion channel, receptor, gap junction, and transmitter-related genes examined in this study. The innexin gene family comprises the gap junctions in invertebrates (Phelan, 2005) and makes up the electrical synapses among STG neurons. The innexin family has some of the highest level of mRNA abundances out of all the genes examined here at nearly an order of magnitude greater than many of the channels and receptors. From the four innexin subtypes, *INX4* clearly had the lowest abundance (median = 37) across STG neurons with *INX1, INX2,* and *INX3* having median counts in the thousands. Of the ion channels, the sodium channel *NaV* and the hyperpolarization-activated *HCN* (I_H) channel were the highest expressed. However, *NaV* was highly expressed across all STG neurons (median = 1399) while *IH* was driven by high outliers from expression in AB neurons. Other overall trends in ion channel mRNA abundances included higher *BKKCa* expression than most other ion channels. The remaining ion channels fell into a range of expression in the hundreds of counts per cell. Among the *KCNH* family of EAG-like voltage-gated potassium channels, *KCNH1* was highest and *KCNH3* was lowest among STG neurons. The two-pore domain *KCNK* family only contains two members in *C. borealis*, of which the *KCNK1* member has higher abundance in STG neurons. The *Shaw-like* voltage-gated potassium channels showed higher expression of *Shaw1* than *Shaw2.* Overall expression of the transient-receptor potential (TRP) channels was quite low (all medians less than 100) compared to other channel types.

Receptor mRNA abundances in STG neurons were at similar levels to that
of channels with median expression values typically in the hundreds of counts
rather than thousands. Among the biogenic amine receptors, *Dopa-1Br* is the
most abundant dopamine receptor, while *HTR7* is the serotonin receptor most
often used. For GABA receptors, the metabotropic *GABAB-R1* has the greatest
abundances compared to the other four identified GABA receptors in *C. borealis*.
Only muscarinic acetylcholine receptors were examined in this study with
mACHrB generally being more highly abundant in STG neurons than *mACHrA*.
Among the glutamate receptor subtypes, *NMDA-2A* stands out as the most
abundant across STG neurons. Markers for neurotransmitter phenotypes were
highly expressed across STG neurons making identification of cholinergic and
glutamatergic neurons readily available.

Cell-type specific mRNA Abundances

With unambiguously identifiable neurons, we were able to capture the cell-
specific expression profile of 57 mRNAs related to neural activity and circuit
function. For innexin expression across STG cell types, the trends seen at the
combined network mRNA were maintained at the single cell level (Figure 4.3).
That is, *INX2* had the highest abundances across all the cell types examined,
and *INX4* was much more lowly expressed. However, with the resolution of cell
type, it can be seen than *INX4* is more highly abundant in PD neurons than other
neuron types. All STG neurons utilize electrical synapses to couple the firing

patterns of various neurons for different patterns of network output (Otopalik,

Lane, Schulz, & Marder, 2017). Our data suggest that these neurons do not use

differential innexin mRNA ratios to achieve coupling between only specific

neuron types.

For TRP channels, notable differences across cell types include high *TRP-*

A1 in MG neurons compared to all other cell types. VD neurons had higher *TRP-*

M3 and *TRP-M-like* abundances than other cell types. Amine receptor subtypes

showed *DAR1A* at low abundances across the gambit of cell types, but *DAR2*

had particularly high abundances in LG neurons. *Dopa-1Br* was also particularly

high in LG, IC, and LPG. Histamine receptors showed generally lower

abundances across cell types with notably low expression in LG neurons. For

serotonin, *HTR1A* was highest in IC neurons while *HTR7* was highest in PD, LP,

and LG neurons. Among the GABA receptor family, *GABAB-R1* was most

abundant, particularly in PD and LP neurons. The transmitter receptors for

acetylcholine examined in this study focused on muscarinic type acetylcholine

receptors (Figure 4.4). Most neurons had *mACHrB* in greater abundance than

mACHrA, with exceptions being AB and LPG, although overall *mACHr*

abundance is quite low in both of those neuron types. Similarly for glutamate

receptors, nearly every neuron had *mGluR5* in greater abundance than *mGluR1*

except for AB and LP. The only neuropeptide receptor examined here was the

crustacean cardioactive peptide (CCAP) receptor. Interestingly, *CCAPr* receptor

expression entirely overlaps with glutamatergic neurons and is essentially absent

in cholinergic neurons. Further, the patterns of expression are highly correlated across all glutamatergic neurons. The glutamate-gated chloride channel *GluCl* was much more abundant in PD neurons than other cell types, followed by LP and LG. Kainate-like glutamate receptors were overall more lowly expressed in all neuron types compared to NMDA-like receptors. *NMDA-2A* is the most abundant across every neuron type examined and is particularly rich in IC and LG neurons. For genes related to the synthesis, breakdown, or transport of neurotransmitters, acetylcholinesterase (*ACHE*) was present across all cell types whether the neuron was cholinergic or glutamatergic. However, only vesicular acetylcholine transporter (*vAChT*) and choline acetyltransferase (*ChAT*) mRNA was found to be in high abundance in the cholinergic neurons GM, LPG, PD, and VD. Conversely, vesicular glutamate transporter (*vGluT*) was present in high abundance in glutamatergic neurons (AB, DG, IC, LG, LP, MG, and PY).

The KCN family of potassium channels in STG neurons showed differential expression across cell types. *KCNH1* was more abundant than *KCNH2* and *KCNH3* in all cell types, and IC and LG neurons had the most *KCNH1.* For the *KCNK* subfamily, a similar pattern emerged with *KCNK1* as the higher abundance compared to *KCNK2* across cell types with IC and LG having greatest abundance. *KCNQ2* was highest in LPG while *KCNT1* was highest in LG. Sodium channel (*NaV*) abundance was relatively high compared to other channels and consistent in abundance across cell types. The hyperpolarization-activated inward current (*IH*) was remarkably high in the AB neuron type

compared to all other types. This is likely the result of the AB neuron being an intrinsically oscillating pacemaking interneuron in the STG that drives the pyloric network, and the I_H produced in AB helps to drive these intrinsic oscillations (Jorge Golowasch et al., 2017). The large conductance calcium-activated potassium channel (*BKKCa*) was highest in LG, IC, and LPG. For the voltage-dependent calcium channels *CaV1* and *CaV2*, the expression was consistent across cell types and little difference was seen between the two transcript types. The inward rectifying potassium channel (*IRK*) was lowly expressed across all STG neurons with LP showing the highest amount if *IRK* mRNA. The sodium leak channel *NALCN* was highest in LG, LP, and IC. Among the subfamilies of delayed rectifying potassium channels, *Shal, Shab,* and *Shaker* were present in all neuron types. VD and LG displayed very little *Shaw* mRNA. The only neuron that expressed *Shaw2* at an appreciable level is LPG, which can nearly be presented as a marker gene for LPG in the STG. Compared to *BKKCa*, expression of *SKKCa* was relatively small across STG neurons. However, VD, IB, and AB still show respectable abundances of *SKKCa* mRNA amounts, especially in AB where the amount of *SKKCa* is greater than *BKKCa*.

Cell-type specific mRNA Correlations

Beyond the absolute abundances of mRNAs, correlated patterns of mRNA expression have been shown to be cell-type specific (Schulz et al., 2007). Here we determined the pairwise correlations of all 57 mRNAs quantified across all 11

cell types examined in this study (Figure 4.6). These network plots of pairwise correlations display cell-type specific patterns of mRNA abundances that add an additional layer to the uniqueness of identified STG neurons. The total number of pairwise correlations varied dramatically across cell types, with only 40 pairwise correlations above threshold in LPG and 380 correlations for VD (out of a total possible 1596 correlations). Regardless of cell type, strong correlations were present across all pairwise comparisons of *INX1*, *INX2*, and *INX3*, suggesting a co-regulation of abundances for these gap junction mRNAs. The correlation between *INX1* and *INX2* was particularly strong ($R^2 = 0.948$). These three innexin subtypes are also the most prevalent in STG neurons out of the six subtypes identified in *C. borealis* (Shruti et al., 2014). In general, ion channel transcripts showed more correlations with other channels than with other transcripts, as particularly exemplified in DG, LG, LP, PD, PY, and IC. High levels of ion channel correlations has also been previously observed in the large cell (LC) neurons of the cardiac ganglion in *C. borealis* (A.-E. Tobin, Cruz-Bermúdez, Marder, & Schulz, 2009). *CCAPr* was highly correlated with *vGluT* in glutamatergic neurons ($R^2 = 0.911$), but not in cholinergic neurons ($R^2 = 0.281$). Our data recapitulate what is previously known correlation patterns seen among the *shaker* family of transcripts with *NaV* and *BKKCa* (Schulz et al., 2007). Specifically, there is a 91.4% concurrence with previously identified correlations in GM, IC, LG, LP, and PD with the correlations determined in the present analysis. To better visualize the overall level of correlation present across cell types, we have represented the total pairwise correlations as a cumulative frequency of R values (Figure 4.7).

Similar trends can be seen where LPG has a greater abundance of weak mRNA correlations while VD neurons display higher amounts of strongly correlated abundances.

Discussion

To what degree does the molecular profile of a neuron impact the unique features of its activity? The strength of the relationship between mRNA abundance and protein abundance has been determined to be quite variable depending on the system and methodologies used to examine it. It has been argued that only 40% of the variance in protein concentrations can be explained by the mRNA abundances (Ghazalpour et al., 2011; Kristensen, Gsponer, & Foster, 2014; Schwanhäusser et al., 2011; B. Zhang et al., 2014). However, more recent technologies and improved mathematical modeling suggests that in considering steady-state differences between protein levels, mRNA abundances largely explain the variation in protein concentration (Li & Biggin, 2015; Vogel & Marcotte, 2012) reviewed in (Liu, Beyer, & Aebersold, 2016). Further, when considering ion channel proteins, the conductance carried by the ion channel has been shown to have strong correlations with the mRNA that codes for that channel (Schulz et al., 2006).

When comparing the electrophysiological and transcriptional profiles of STG neurons, it is important to consider whether the neuronal electrical activity is the result of intrinsic properties or synaptic connectivity. Isolated cultured neurons are capable of transitioning from tonic firing to bursting over 3-4 days (G Turrigiano et al., 1994), and the transition to endogenous bursting in decentralized neurons is reliant upon *de novo* transcription (Muriel Thoby-Brisson & Simmers, 2000). The mRNA expression profile can also be useful in examining both levels of neuronal properties in that ion channel mRNA composition can be most relevant to intrinsic properties (such as high *IH* in AB neurons), while neurotransmitter phenotype and receptor mRNA expression are strongly indicative of synaptic activity, both for producing and receiving signals. For instance, having this molecular profile across STG neuron types revealed how *CCAPr* is essentially only expressed in neurons that also are glutamatergic, which has yet been unexplored. Further, the mRNA expression gives credence to exploring the conductances related to *SKKCa* in greater detail, as the mRNA expression is significant, while the electrophysiological aspects have been understudied.

What is the fewest unique mRNAs that must be measured to tell STG neurons apart? It is often the exception that certain genes are expressed in a binary "on/off" manner across cell types, but rather gene expression typically results in a continuum of mRNA abundances (Hobert et al., 2010). However, combining the expression of genes such as those that define neurotransmitter

phenotype with those that, while not expressed in a binary manner, do exhibit strong differential expression, we can generate a panel of qRT-PCR assays that correctly discriminate neuron types. First broadly separating glutamatergic from cholinergic neuron types can be done by measuring *vGluT* abundance alone should be sufficient with high abundance in glutamatergic neurons and nearly no presence in cholinergic neurons. Measuring *GABAB-R1* expression could further isolate PD from other cholinergic neurons, as well as separate the glutamatergic into high (LG, LP, and IC) and low (PY, DG, AB, and MG) groups. Further subdivisions could take place by needing to measure only *BKKCa, mACHrB, IH, TRP-A1, DAR2,* and *Shaw2* to isolate the rest of the STG neurons into unique bins. This means that the expression of only 8 mRNA species would be necessary to separate out 10 STG cell types, which supports the notion that the expression of clusters of genes are most important for these STG neurons.

This study is not without its limitations. While neurons are acutely extracted from the network and assumed to be in a steady-state of mRNA expression, this is an assumption. It is entirely possible that different neurons across preparations could be undergoing transitions from one state of expression to another, producing different patterns of mRNA expression. This is one potential explanation to the amount of variability in mRNA expression seen within the same STG neuron types (Schulz et al., 2006). However, compared to extraction methods, such as fluorescence activated cell sorting (FACS), our dissection procedure gives much less time for changes in the transcriptional

landscape of these neurons to take place (Richardson, Lannigan, & Macara, 2015; van den Brink et al., 2017).
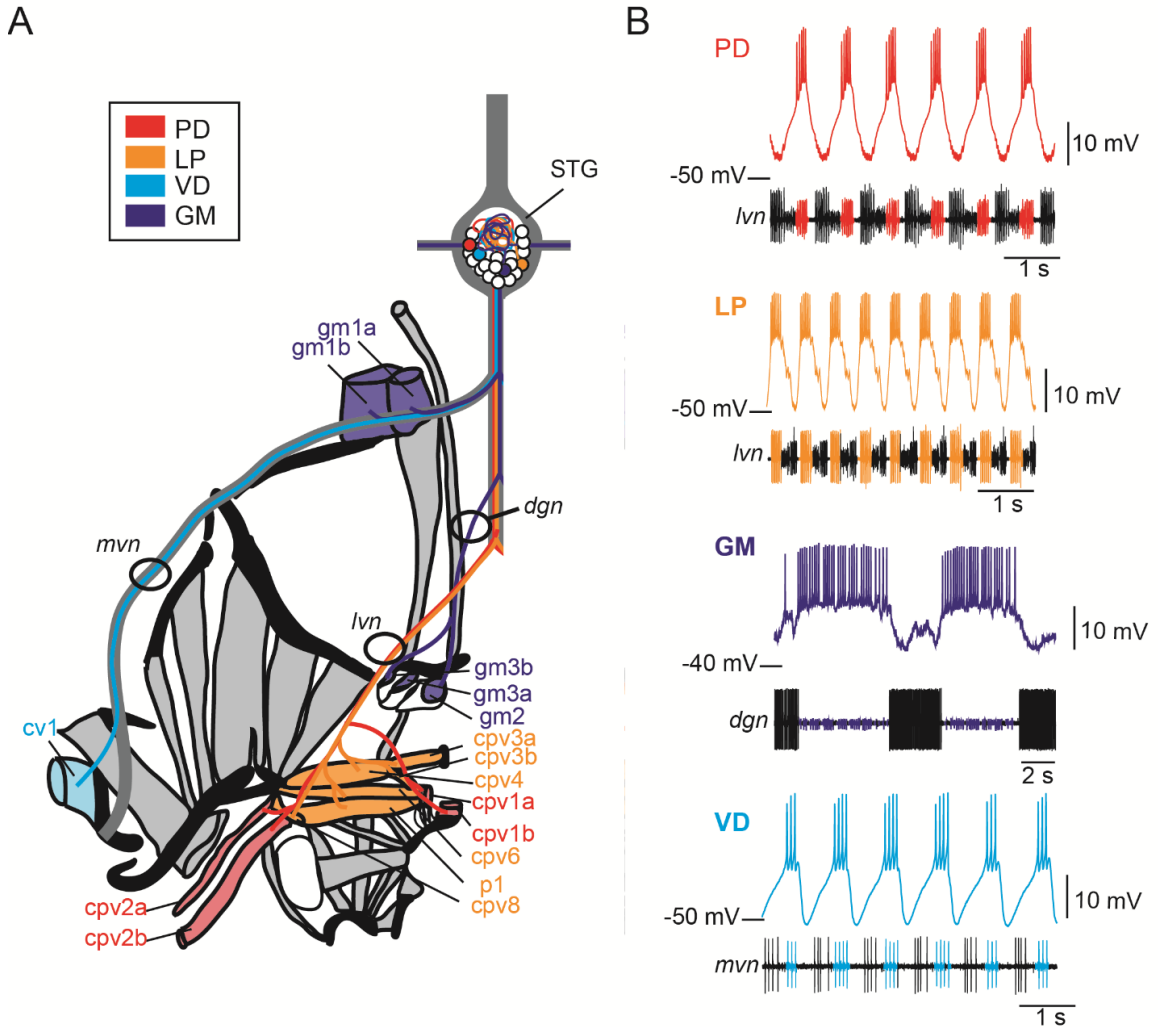
FIGURE 4.1: *Muscle innervation and firing patterns of four STG neuron types.* **A.** Anatomical schematic of one side of the foregut of *Cancer borealis* showing the axonal projections of PD, LP, VD, and GM neurons to specific muscle groups (in the animal, these are symmetrical bilateral innervations patterns). Relevant axons and muscles are color-coded by neuron type as shown in the key. Muscles that are not innervated by the neurons of interest are indicated in gray. Nerves

containing axons from these four neurons are circled and denoted with black text.

B. Stereotyped intracellular voltage waveforms matched with concurrent extracellular spiking patterns in nerves known to contain the respective neuron's axon. Color-coded as in A. *mvn* = medial ventricular nerve; *dgn* = dorsal gastric nerve; *lvn* = lateral ventricular nerve.
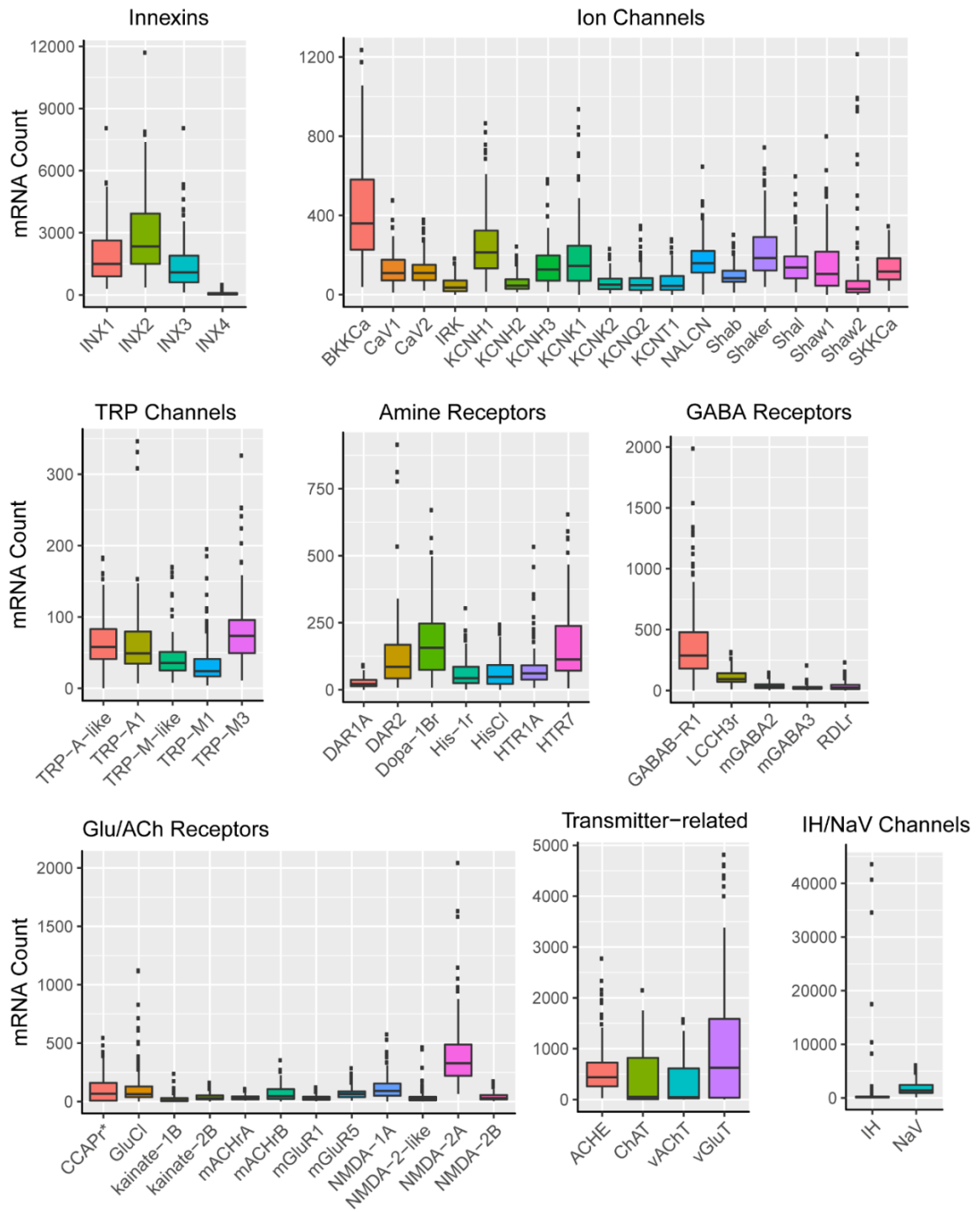
FIGURE 4.2: *Combined STG neurons mRNA abundances across genes related to neuronal function.* The absolute mRNA abundances were taken from STG

neurons (N = 119) across all STG cell types and analyzed for trends across

genes. Box plots are displayed with data points representing outliers beyond the

outer quartiles. The y-axis for each plot is scaled according abundance

differences across gene groups.
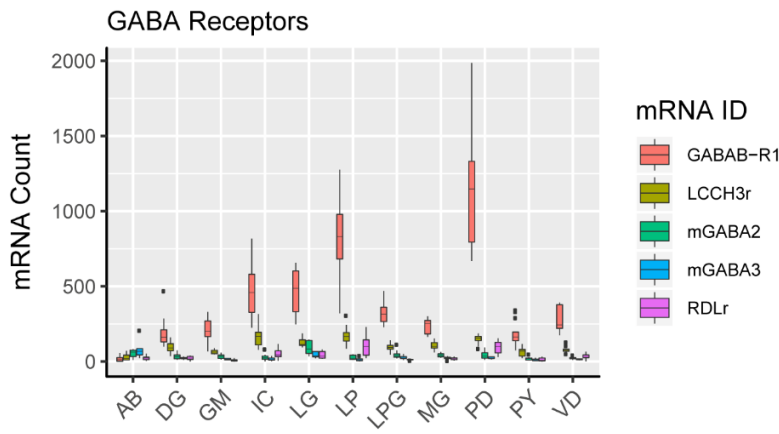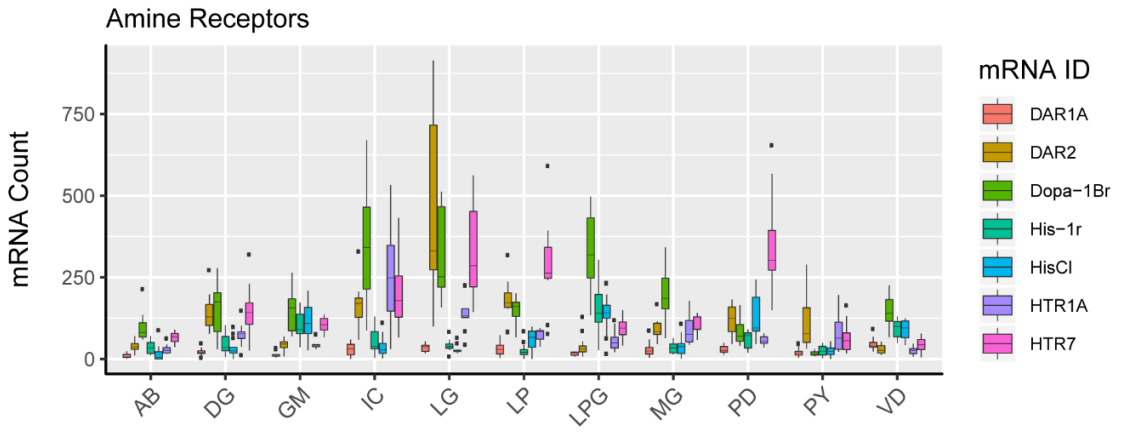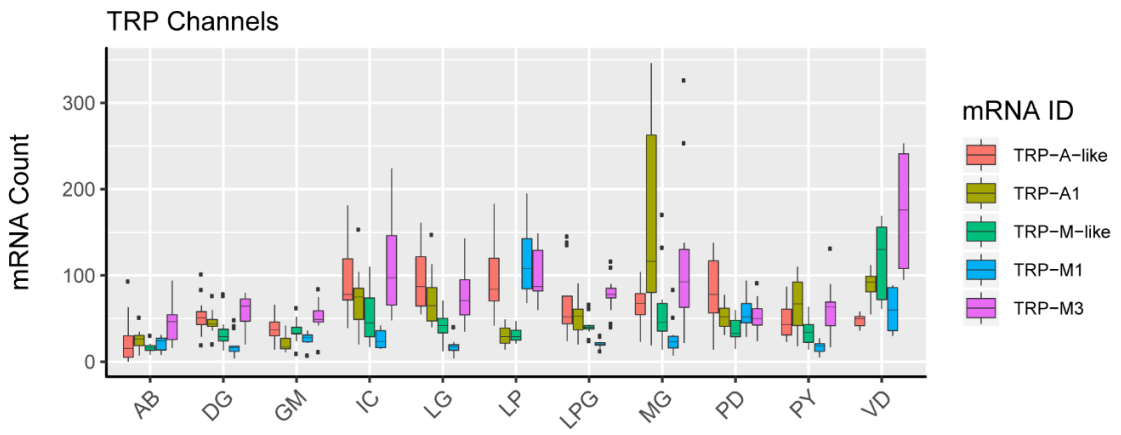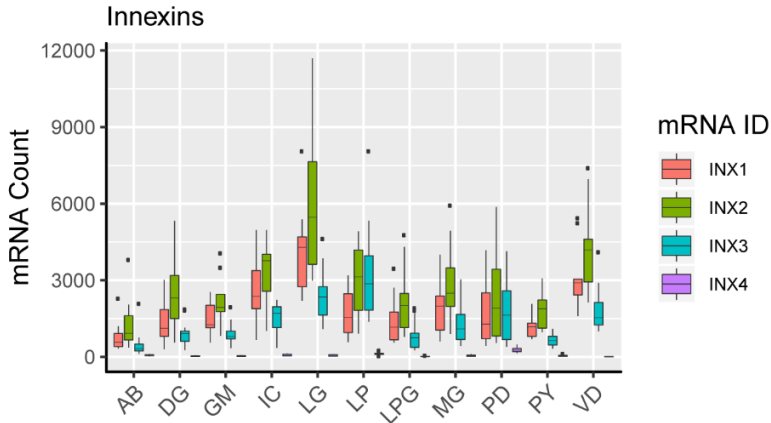
FIGURE 4.3: *Cell type specific mRNA abundances across innexins, TRP channels, GABA receptors, and amine receptors.* STG neurons are groups based on their identified cell type along the x-axis. Sample size for each group is represented as follows: AB = 8, DG = 13, GM = 9, IC = 11, LG = 10, LP = 12, LPG = 13, MG = 10, PD = 11, PY = 13, VD = 9). The y-axis for each plot is scaled according abundance differences across gene groups.
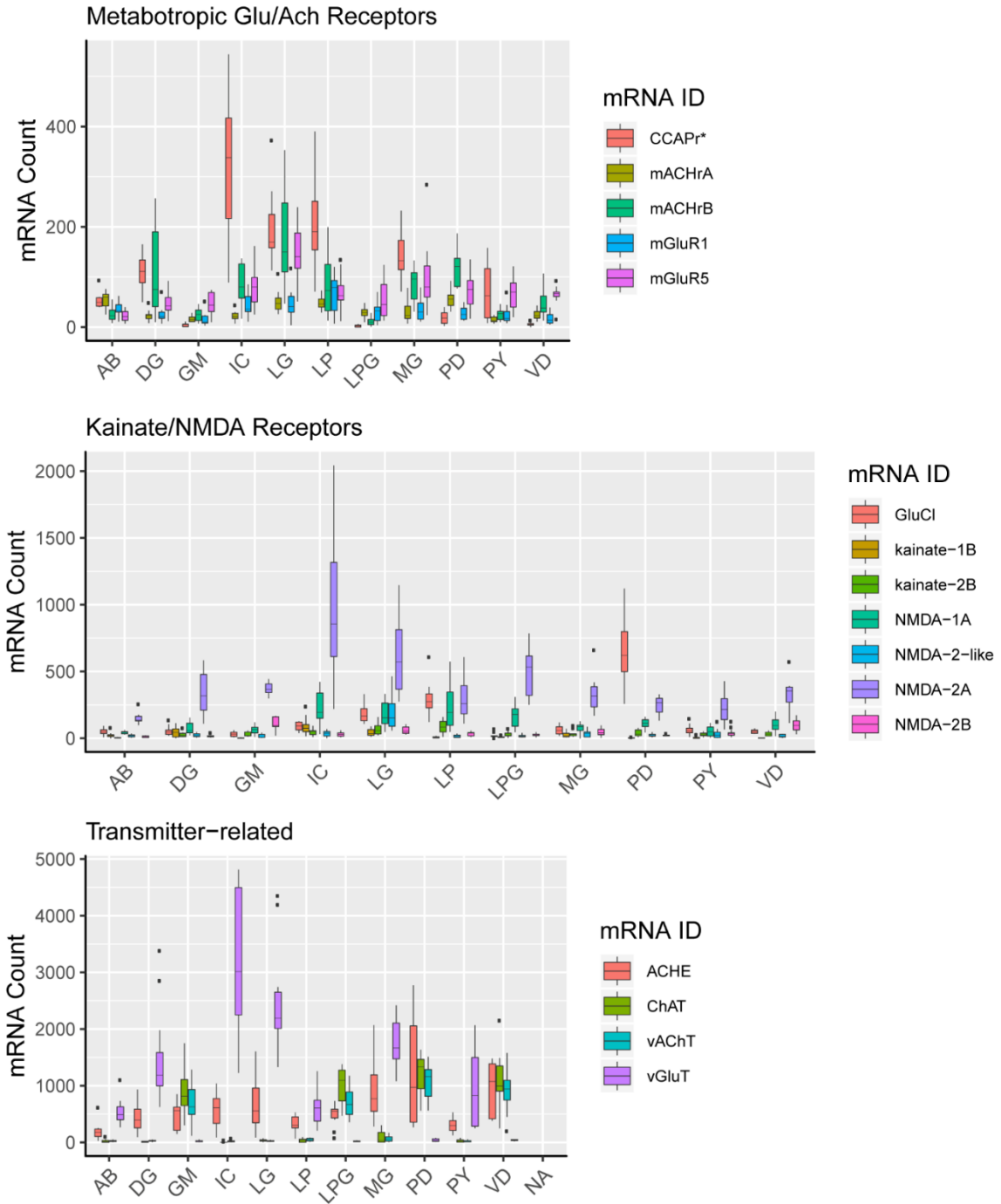
FIGURE 4.4: *Cell type specific mRNA abundances across metabotropic glutamate and acetylcholine receptors, kainate- and NMDA- like glutamate receptors, and transmitter-related genes*. Plots are generated in the same manner as previous figures.
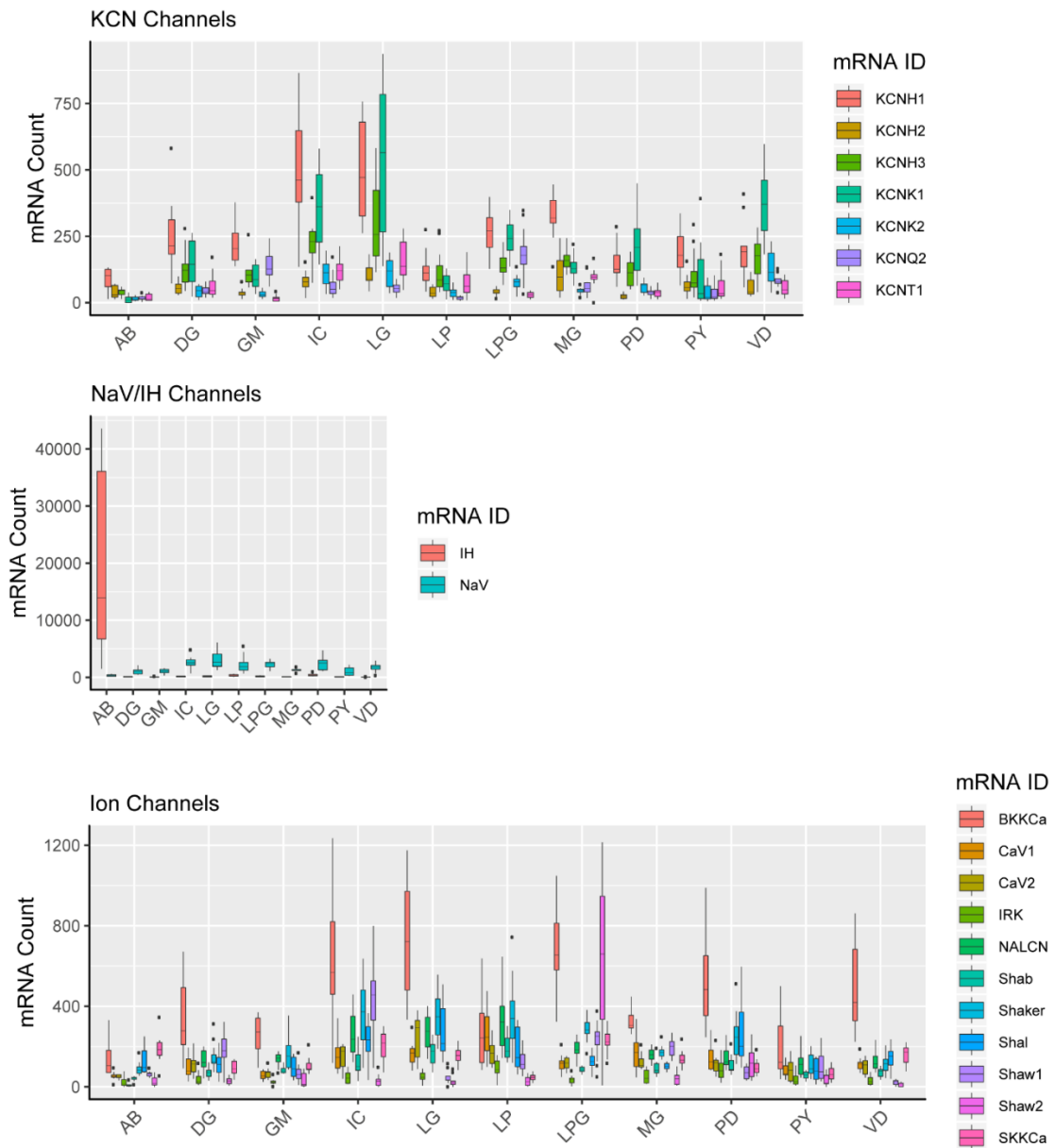
FIGURE 4.5: *Cell type specific mRNA abundances across ion channels.* Plots are generated in the same manner as previous figures. *IH* and *NaV* were separated out from other channels due to their much greater mRNA abundance across all cell types (as is the case with *NaV*), or in AB neurons (as is the case with *IH*).

AB

Ion Channels
TRP Channels
Amine Receptors
GABA Receptors
Glu/ACh Receptors
Transmitter-related
Innexins

NaV
NALCN
CaV1
CaV2
Shaker
Shal
Shab
Shaw1
Shaw2
BKKCa
SKKCa
IRK
IH
KCNH1
KCNH2
KCNH3
KCNK1
KCNK2
KCNQ2
KCNT1
TRP-A1
TRP-A-like
TRP-M1
TRP-M3
DAR
DAR1A
Dopa-1Br
His-1r
HisCl
HTR1A
HTR7
GABAB-R1
LCCH3r
RDLr
mGABA2
mGABA3
mAChrA
mAChrB
mGluR1
mGluR5
kainate-1B
kainate-2B
GluCl
NMDA-1A
NMDA-2A
NMDA-2B
NMDA-2-like
ChAT
vAChT
vGluT
CCAP-r
INX1
INX2
INX3
INX4

MG

DG

LG

PD

GM

LP

PY

IC

LPG

VD

142

FIGURE 4.6. *Cell type specific network correlation analysis*. The Pearson

correlations of mRNA abundance are displayed as a network among gene

groups. Lines designate a correlation present between two mRNA species with

an R-value threshold of 0.75. Line thickness increases proportionally to the

strength of the correlation.
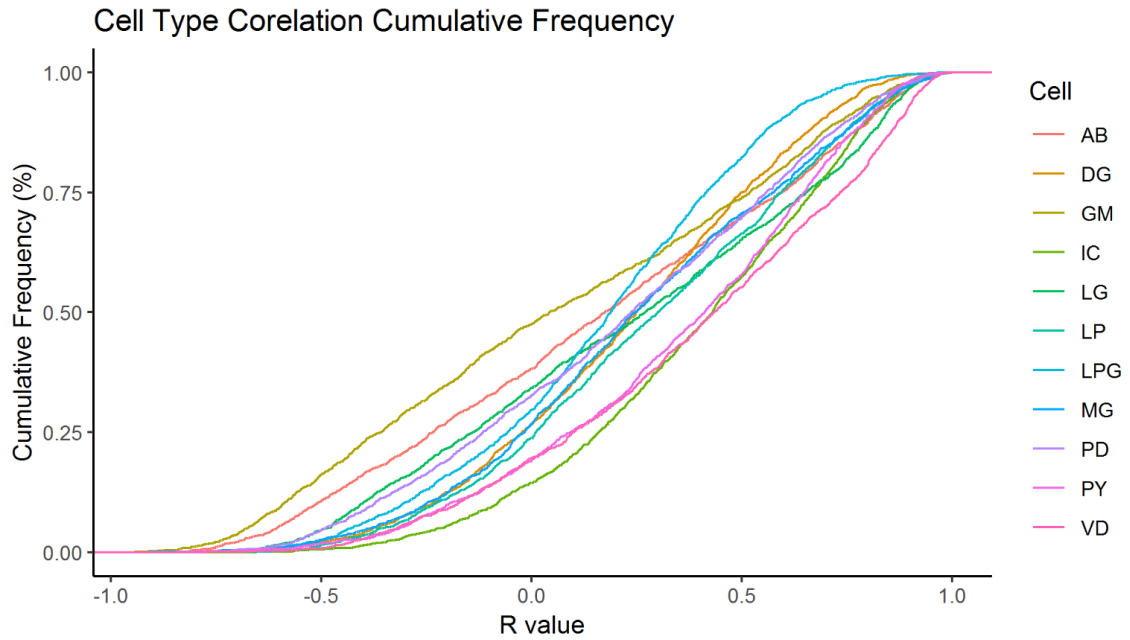
FIGURE 4.7: *Cumulative frequency of cell-type specific correlations.* The

cumulative frequency of R values for all possible pairwise correlations (1596

correlations total) are given across the cell types of the STG neurons. The

cumulative frequency is given as an additive percentage of correlations present

at each given R value.

TABLE S4.1: *Target primer and probe sequences for qPCR Multiplex assays.*

| Gene	Forward **/** Reverse Primer 5'-3' sequence	Probe 5'-3' sequence
HTR1A	AACCGCTGTGGTAGTTTCCA / TGCTCGTTAACCCGGACTAAG	AGCGCCTTTATTTGGCTGGAAGGA
HTR2	TCCGCCTCCATCAAGTTTGT / GCACGTTGGCGGATGAAGAAC	TCATCGAAGAGACACGGGAGGACC
HTR7	ACGGCGATGGCTCCATCTG / CGGTGAGCGGGATGGTAGAAG	TGAGGTGTGCAACAACTTCTGGTACC
DAR2	GAAGCCGAAAGTGAGTGAGATCA / TCCGAACTAAGCGCTGTTTC	TGTGATCGAGAATGTGACACAGACGA
DAR1A	GGCGCCTGTCCATTCACT / CGCGGTAGATGCGGAAGTAAG	ACCTTGTGTTCTCCTCTACTATCAGCTTCT
HisCl	TGCCTGCCAGAGTAACCTTAG / GACAGGTGGGAGGGATTTCTG	TTCTCACTCTGTCAACCCAACACGC
GABAB-R1	TCGTCTCGTTTGCCATCATC / GGTGCCGAACCTCAATGATC	TCTGCTGCTTCCTCTCCATGGCT
LCCH3r	TGACGGCTCCATCACCTATGG / TTGGGTGTCGAGTGGATAGTAG	TTCACCACTACGTTGGCCTGCAT
RDLr	TGGTGTTTGCCTCGCTTCTAG / TCCGCTGTTCTGCTAACTTC	AATACGCCGCGGTGGGCTACAT
GluCl	ACGGAGGATCTGGTGTTTCTG / ACCCGTGTTGGTCTTGCTGTT	TACAGGTGACCAAGAACCTTCACC
NALCN	TCGCTTCCACGGTGTACATTC / GCGGTGCCTTTGTTCTCAG	TCTTCGTCTTCCTTGGCTGCATGA
CACNAB	GCAGCTGGCCAAGACTTCTTT / AGACGCTGCAATACCTTAGGA	AGCGCCCATCCTCGTGTACCTTAAG
IRK	TACAGTGGCGTTGGACTCTAC / TCCACCACACCAAGGCAAATAG	TCGTGTTCGCTATGTCATTCATCAGC
IH	TCGGTGCCACTAGACTACATC / GACCCGCGTGGAGAATCTG	TCCTCATCTTCAACCAGGACTTCAGC
SKKCa	GCATCGGAGCATTGAACAGAA / GCCCGGACAGATAGTCATCAG	CAACTTCAACACTCGGTTTGTCCTCAA
INX1	TGGAGCGTCATGATGCATTC / GAGCAGGATGGCAAGGATCAC	TGCTGCCTCTCAACATTCTTAACGAA
INX2	GGCTGTGGTGTCTGGTGTAG / GCGAGAGCGTGTCCTTAACAG	CTGCTGTACCGCCTCGCCACTTT
INX3	TGTCGGCCCTAGTGAAAGAG / GGTACCGTGGGATGTAGAACA	TGACGAGATTGTGTACCACGCTTAC
INX4	CTGGCGTTCAGCCTCATTGTC / CACGTCCTCTGGGATCTCCTTAG	CACGCGTCAGTATGTCGGGAACC
INX5	TGCCTTCCCTGCTGGATAA / GCGTCACCCATTGGTAGTAAC	AGGTGGCTCATCCAGGCATCGGT
Shaw1	CGCGTCACTCCTCAGGACTT / CCCAGCACCAGGAAGAACAC	TGATACAGACTTTCCGTGCATCCGC
Shaw2	GAACGCCATCAAGCACTATCATC / ATGGCGCCCGACAGCTTAG	TGGCTTGAAGGACGGTCTCACA
NaV	TCAACGGGAGGTACCATAAGTG / TCGCTGTTCACCCAAGAGTAG	CGGAGGGATTGAAGCTCAACGCA
Shaker	GAGGCTCAGAAGACCAGTCAAC / TGCGATATCACCGAGCTCAT	CACTCGATGTCTTCGCGGAGGAGAT
Shab	GAGCCGGACAGACAGGAAC / TGCGCCTCCTTCTGTAGTC	AAGAACCACGAACACCACATGGGTC
CaV1	CCAGGCCTTCTACTGGCTCATT / GCTGGCGATAGTGCTCACTG	TGTGCTCGTCTTCCTCAACACGG
CaV2	ATCCGGCGGACAGTAAAGC / GTTCGGCAGCAACACAAAC	TGGTTCTACTGGTTCGTCATCATACTTGT
CaV3	TGGCTGCCACCGATACTTC / CAGCACAATGCCCACAACTG	CAGGACAGAGATGGAACCAGTTGGA
Shal	GACACCACCTTCACCTCCATTC / GAACCATGTCGCCGTATCCTA	CGGCGTTTTGGTACACCATTGTCAC
BKKCa	GCTCAAACTCGGCTTCATTG / CTGCGTGTCTGGAGAAGTTT	AGAATCCCGGCGCTAAACATGACT
mGluR1	GCATCGTGGCTCATCTTTG / GGCCAACGTGACCACTCTAAT	ACCTGTCAGCGGGAGTCACTGATG
mGluR2	TCCGCAGGTGAGTTTCTTC / CCATGGCTTTCACTTGGTAATGG	CCACGTCGCCTGAACTGAGTAACA
mGluR4	GCGCGTTGATTCCGGTACT / CCACTCATCGTCCTCAACTTC	AAGTCTTCCCGCTGGACTACGAAC
mGluR5	GCCTGTCCTTTGCCATGATC / TGCGCATCGTGATCTTCTTG	CGCTTGTCACCAAGACCAACCGC
mGluR7	ACCGCGCTCGCAGATTGTC / TGGCTGGCGTTTCCACTATC	TTCTGGACTGGTGAGTGTTCAGCT
mACHrA	GGTGTCGATGCCTTTGTTCAC / GCCAGCCAGGTGTCACATAT	TGTACACGCTGATGGGATACTGGC
mACHrB	GCGGACCATCTCCATCATTC / TGGTGCACTGACAGAAGCT	CCTTTGTGGCCTGTTGGACGCC
mGABA2	TGCCTACAAAGGTCTGCTAATGG / TCGTTCAAGGCGGGTATGTT	TGGTTGTTTCCTGGCCTGGGAAA
mGABA3	CGATCCCATGACGAGACACAT / TCCACCTGTGGCTGATAGAC	CCTCACCTTAGAAGTGTCAGCCGA
mGluR3	CACCGTGTATGCCGTCAAGAC / CCCGGTGCCGAAGTAGATG	CCCCGAGAATTTCAACGAGAGCAAGT
TRP-M3	CCGCACCATCTACGAGAAC / TGCTGGCCTGGAAGATGT	TGCTCAAGTCTCCTCGTCTTCACC
TRP-A-like	TCGCGACCTTCCTCAAATTC / CGGTACCTGAGTCCTCAACAC	CACGGTCTTCTTACTCTTCCTCATCGC
TRP-A1	CTGCCAAGTACGGTCGTTACAC / CCCTCGTCATTGCACTCGTTA	ACGTCAGCTTGTGGAGTCTCTGAA
TRP-M1	GAGGGCGGACCTCAAACTATC / TGTCGGCTGCTCTTCCTGTT	CGTCAGGTGCTGGAGTATGTCACTG
TRP-M-like	GACGGGACGCAGATCCTCTT / GAGTGCTTGGGCTGTTAGGT	ACGGTATACGGTTGGCTATTTCCCA
Dopa-1Br	CGCAAGATTGGCAACCTCTTC / ACGAGGGCAAAGCTCATCAC	TGTCCTTGGCTATCGCTGACCTCTT
5HTR-1Br	TGACGCAGGTGGACTACATTC / GGAACGACACCACCCAGATC	ACCGATCGCCGCGTCGCATC
His-1r	TGCCTGCCAGAGTAACCTTAG / GACAGGTGGGAGGGATTTCTG	CATCACTTCTCACTCTGTCAACCCAACA
His-2r	CCGCCACAGTCTCAAGGTAATC / GCGTAGGTCATGGAACTCTCATC	ACGGTAGTCTACTTCCACGTCACA
His-3r	ATCCGCCGCAACAAAGCAT / GAGAGCGAAGGAGGTTGGAA	TGATGGTGGATCGAGTCTCAAGATATGTAT
kainate-1A	CAGGTCGGAGTGCAGTAAAGAC / GCCACCAGTCAGGATGTAGAAG	CGATGACCACCCAGACGAGTGC
kainate-1B	TGAGCAGAACGAGATCGAGTATG / CGCCACATGTTCTGATACGTC	AGGGCGGGTCTACCATGGCCTT
kainate-2A	CGCATGGAGTCACCTATTGAGA / AGGCGAAAGTGGTGCCAGTTG	TGAGGACCTTGCTAACCAGGACAA
kainate-2B	GCACGGCAAGTTTGACAAGAAG / TGCTCCCTCTCGTAAGTGATG	AACGGCATGATTGGGCAGCTGTT
kainate-2C	GGCTTGGTCAGGGAACTCAAG / GCTCTCCCTCGCGTAGTTG	TGATCTAGCGGTGGGTTCTATGACTA
NMDA-1A	GCCGTCAAATCAGGGAGGTT / ACCGGCGGTTACCAGTTCAC	AGGCGTTCATCTGGGACAGTTCACGT
NMDA-1B	ACAGCCAAGACGAAGAAGAC / CCGCTGTTCAGGATGACAGA	TGAGTTCATGGCCATCTCGGAGTC
NMDA-2A	TCGGGTTCGTTCCCTTCAC / TGATGCCGTCCGTGATAGAAG	TGAGACCATCCTTGCCAAGCACC
NMDA-2B	GCAAGGGTCACCATCAGACA / CGCTGTGAGCCATGATGTAGGTA	TGGAGAAACAACTTGAGGCCAATGGA
NMDA-2-like	GCGTTGGAGCAGTTCATGTC / GCCACATACTGACGGGAAGTAC	CACGTTCCTGCTGTTGGGTTGTG
KCNK2	GACGCCTTCTACTACTGCTTCATC / GAGGGCGTTCTCCTTCTGTAG	CCCTCACTACCATTGGCTTCGGG
KCNK1	TGGCGAACGACTCAACAAAG / GTCCTGGCACTTAAGGATCTTC	CTCCTCCATCGGCATCAGGCAA
KCNQ1	GAGCCTCCTTGGGAAACCTATC / CCGCTCCAGGAAGTTGTAGAC	CTCTCGCAGGGACGTCCGCTAC
KCNQ2	GCTGCCATGTTGATCCAGTG / CCACGTTGCTGTAGAGTTGAAG	TGTGGCGTTGTTATGCTGCAGATAA
KCNH2	CACCGCGAGATCCTTTCAC / CCTGATGTGGAGGCTGAGTAG	CATCTTCGAGACAGCGTCGCAGG
KCNH3	GAGGCGACGTACTTACCTCTATG / AGTGGCGTACATGCAAGGATTC	ACTTCATCTCAAGAGGCTCGCTAGA
KCNH1	GGTCACGTCACCACCATCATC / CCGCACGTTGTTGAGCATTTC	ATGACCTCCGCCACCGCCAAGT
KCNT1	CGTCCAGACCATGTTCAAGTTC / AAGCGCATGTTGGACGACTG	TCCCCAACATCAAAATCATCACGGA
CCAPr	GCCCTTCTCCTCTCCAAATCAC / GTCGGTGAGAACGCTGATGAG	CCAGGACCAACTTCTTCATCATGCATCT
vGluT	GCGTTCGTGGACCTTCTAC / TCAGCCACCCTGTAATGGAA	ATCACAGCCAACCTACTTCAGCGAG
ChAT	GGACCGCCTGGCTAAGTAC / TCGCGGAGTCCCATAAGG	AGGCGGCGCTCAAGCTTCAGAC
vAChT	GCGTCAGCTGCTTCTTCCT / CAGCAGTGCCGTGTCTATGAG	TTCGCCAGCAACTACTGGGTGTT
ACHE	GGGCAACATGGGCATGTAC / GGTCACCACCGAAGAATTCAATG	AGGCGCTGGCCATCAAGTGGATAC

CHAPTER 5

Explorations in molecular mechanisms underlying mRNA correlations of the stomatogastric neurons system of *Cancer borealis*

Introduction

The regulation of eukaryotic mRNA abundances is a process that involves diverse molecular mechanisms to manipulate the expression and stability of mRNA species, and it is the rate of this synthesis and turnover that drives mRNA expression patterns (Perez-Ortin, 2008). The rates at which particular mRNA transcripts are degraded has been shown to be variable and transcript-specific (Herrick, Parker, & Jacobson, 1990; Sachs, 1993). The regulation of mRNA in neurons has an additional layer of complexity added where activity, in the form of calcium signaling, drives changes in mRNA expression (Greenberg, Ziff, & Greene, 1986b; Greer & Greenberg, 2008; West et al., 2001). In order for a neuron to generate consistent patterns of activity, the conductances of inward and outward currents need to be balanced, but there is much variability in the absolute values of conductances required to generate a particular activity pattern (Golowasch, Goldman, Abbott, & Marder, 2002). Computational modeling has

146

suggested that simple tuning rules in neurons could regulate the rates in which mRNAs are synthesized and degraded to arrive at the appropriate solution for generating activity (O'Leary et al., 2013). This regulation results in mRNA expression patterns that are correlated with the conductances suspected to be carried by the ion channel protein produced from that mRNA (Schulz et al., 2006). Further, correlations between mRNAs have been shown to be cell-type specific (Schulz et al., 2007), present across multiple systems (Tobin et al., 2009), and often dependent on neuronal activity (Temporal et al., 2014a).

What is the molecular mechanism by which correlated levels of mRNA abundances arise in neurons? Certainly a simple explanation for this phenomenon is transcription factors that concurrently drive the expression of multiple genes, resulting in correlated mRNA abundances from similar levels of transcription. Conversely, transcription could be loosely regulated with stricter regulation occurring at the post-transcriptional levels. The rates at which mRNAs are signaled for degradation could lead to correlated mRNA abundances through parallel regulatory sites across mRNAs. Targeted mRNAs having similar regulatory sites to which RNA-binding proteins (RBPs) or microRNAs could bind is a potential mechanism for controlling levels of groups of mRNAs. RBPs can influence mRNAs through signaling to the core degradation machinery (Pérez-Ortín, Alepuz, Chávez, & Choder, 2013), while miRNAs utilize the RNA-induced silencing complex to increase degradation rates (Aalto & Pasquinelli, 2012). Indeed, the binding of RBPs to mRNAs leads to changes in stability that can

either increase or decrease the half-life of groups of mRNAs (Hasan, Cotobal, Duncan, & Mata, 2014). These regulatory binding sites are often found at the 3' untranslated region (3'-UTR) of the mRNA (Heaton et al., 1992).

Here we sought to better understand the molecular underpinnings of how these correlated patterns of mRNA expression are generated and maintained using the stomatogastric ganglion (STG) from *Cancer borealis* to manipulate the molecular profile of identified neurons. We used a multifaceted approach to tackle this question by perturbing STG neurons through attempting to disrupt preexisting correlations or forming new mRNA correlation patterns not previously found in the endogenous neuron.

Materials and Methods

Tissue dissection and RNA extraction

Adult male Jonah Crabs were purchased from the Fresh Lobster Company (Gloucester, Massachusetts, USA) and housed in filtered and cooled (10–12°C) artificial seawater concrete aquaria. Experiments were carried out throughout 2018 and in the spring of 2019. Anesthesia was induced in animals by submerging in ice for 30 minutes before dissection. Following anesthesia, the stomatogastric nervous system (STNS) was removed from the crab and pinned

out on a Sylgard-coated dish (Dow Corning) with cooled (10-12°C) physiological saline (composition in mM/l: 440.0 NaCl, 20.0 MgCl2, 13.0 CaCl$_2$, 11.0 KCl, 11.2 Trizma base, and 5.1 maleic acid pH = 7.4 at 23 °C. in RNase-free water). Upon desheathing the stomatogastric ganglion (STG), pyloric dilator (PD) neurons were identified by aligning the timing of intra- and extracellular activity with associated recordings, as is standard practices for *C. borealis* (Hooper et al., 1986; Weimann et al., 1991). PD neurons were acquired using techniques previously described (Schulz et al., 2006). In short, a Vaseline well was erected around the ganglion for incubation in ~1mg/mL protease (Sigma, St. Louis, MO) for 10 minutes, fresh physiological saline was perfused to wash the well, and for stability in pulling neurons a solution of cold ethylene glycol in saline (7:3 ratio) was applied. Following this, the preparation was placed in a -20 °C freezer for 30 min, and each neuron was manually extracted and added to a cryogenic microcentrifuge tube that contained 400 µL lysis buffer (Zymo Research) and stored in a -80°C freezer until extraction of RNA. Per the manufacturer's protocol, total RNA was extracted with the Quick-RNA MicroPrep kit (Zymo Research).

cDNA synthesis

Following extraction of total RNA, samples from individual neurons were reverse transcribed into cDNA using qScript cDNA SuperMix (QuantaBio, Beverly, MA, USA) with oligo-dT and random hexamer primers in 20 µL reactions per the manufacturer's protocol. A preamplification step was used to generate

cDNA for several qPCR reactions using PerfeCTa PreAmp Supermix (QuantaBio) in a 20 µL reaction per the manufacturer's protocol. Reactions used a 14-cycle PCR reaction with primers targeting specific transcripts to be subsequently quantified. Preamplificaiton bias on these targets has been previously determined to be minimal.

Quantitative single-cell RT-PCR

Multiplex assays targeting *C. borealis* genes of interest used in this study were validated previously. In short, custom multiplex assays were generated using the RealTimeDesign™ qPCR Assay Design Software from LGC Biosearch Technologies (Teddington, UK). Targets were assembled into cassettes with similar sequences reduce interference and off-target interactions in simultaneous qPCR reactions. Preamplified cDNA samples were diluted in nuclease-free water to 150 µL, as is standard practice in these multiplex assays. Triplicate reactions of 10 uL were generated using 96-well plates on a CFX96 Touch™ Real-Time PCR Detection System from Bio-Rad (Hercules, CA, USA). The cycling protocols for the qPCR were as follows: a hot start of 95°C for 3 min proceeded by forty cycles of two alternating temperature steps: 95°C for 15 sec and 58°C for 1 min. Fluorescence measurements were taken following the elongation steps of each cycle. The primer and probe concentrations for reactions were 2.5 µM and 0.3125 µM, respectively.

Absolute mRNA abundances were determined using standard curves previously generated for each RT-qPCR multiplex assay. Serial dilutions of custom gBlock gene fragments (Integrated DNA Technologies, Skokie, IL, USA) containing amplicon targets were diluted from 1 X 10^6 to 1 X 10^1 copies in for qPCR reactions, and standard curves were further established to be linear and reproducible. Absolute mRNA abundances were calculated with the efficiency and slope produced from the standard curves with corrections for the 14-cycle preamplification and cDNA dilution. Correlations in mRNA abundances were analyzed using Pearson correlation statistics.

eGFP-Shal 3'-UTR fusion mRNA construction

Identified endogenously expressed *Shal* 3'-UTR mRNA sequence from the *C. borealis* transcriptome was confirmed by RT-PCR on STG neurons and subsequently Sanger sequenced to confirm nucleotide sequence. This sequence was used to generate a fusion between the coding sequence of eGFP (sequence from Addgene public plasmid repository) and the 3'-UTR of *Shal.* This sequence was generated by creating a custom DNA gBlock sequence from Integrated DNA Technologies (San Jose, CA, USA) that was driven by a T7 promoter (5′-TAATACGACTCACTATAG-3′) and contained 5'UTR sequence from 60 random nucleotides for spacing. *In vitro* transcription of sense mRNA was generated off

this gBlock construct by using the HiScribe T7 ARCA mRNA Kit with tailing (New England Biolabs Inc., Ipswich, MA, USA). Per the manufacturer's protocol, DNase I treatment was used to degrade gBlock remaining in the reaction after *in vitro* transcription. Synthesized mRNA was purified by the Quick-RNA Microprep Kit (Zymo Research, Irvine, CA, USA), size verified by gel electrophoresis with a reducing RNA dye, and concentration was measured by a Nanodrop-1000 (Thermo Fischer Scientific, Waltham, MA, USA).

Actinomycin-D Treatment

The transcription inhibitor actinomycin-D (dactinomycin) 93% was acquired from Thermo Fischer Scientific (Waltham, MA, USA). Treatment of STG neurons with actinomycin-D was carried out by using a Vaseline well surrounding the STG to isolate treatment and avoid blocking transcription of the commissural and oesophageal ganglia. Actinomycin-D was dissolved in saline at 5 µg/mL and applied to the STG well for either 24 or 72 hour intervals before neurons were extracted for RT-qPCR. In separate experiments, actinomycin-D was administered to STG neurons directly through pressure injection through a microelectrode using a Picospritzer II from General Valve Corporation (NJ, USA) at a concentration of 20 µg/mL in 200µM KCl and 0.5% phenol red (Sigma, St. Louis, MO). Efficacy of actinomycin-D disruption of transcription was measured as relative changes in 18S rRNA abundances across time controls, as well as relative changes in targeted mRNA.

INX1 and INX2 mRNA synthesis

Innexin 1 (*INX1*) mRNA was generated from a pGem-T plasmid vector system (Promega, Madison, WI, USA) that contained a full length insert for *C. borealis INX1* sequence. *In vitro* transcription was carried out as described above using a T7 RNA polymerase to generate sense mRNA from the DNA vector. *INX2* mRNA was synthesized using primers with attached barcodes and promoters that amplify full length *INX2* sequence. Barcodes allowed for downstream qRT-PCR to differentiate injected mRNA from endogenous mRNA of the same type. An endpoint RT-PCR reaction was carried out with the barcoded primers on *C. borealis* cDNA generated from whole brain RNA. T7 RNA polymerase was also used as the primer contained an upstream T7 promoter site for generating sense mRNA. Synthesized mRNA was purified by the Quick-RNA Microprep Kit (Zymo Research, Irvine, CA, USA), size verified by gel electrophoresis with a reducing RNA dye, and concentration was measured by a Nanodrop-1000 (Thermo Fischer Scientific, Waltham, MA, USA).

INX1 and INX2 mRNA injection

Following synthesis, mRNA was diluted 100-fold to 7ng/µL and mixed in 1 µL of 200 mM KCl in 0.5% phenol red for injection. This injection solution was

pipetted into a glass filamented, thin-walled microelectrode (1.0mm x 0.75mm; 4''

long) A-M systems (Sequim, WA, USA) pulled on a P-97 Micropipette Puller from

Sutter Instruments (Novato, CA, USA). A pulling protocol that generated long

electrode tips was used to allow for microelectrodes that maintained a high

resistance even after the tip was sheared to allow for mRNA pressure injection.

Electrode tips were sheared with fine wire pressed to the tip, and tip diameter

was subsequently measured as a function of resistance. Electrodes were used

for mRNA injection only if the electrode resistance was higher than 20 MΩ (to

reduce cell damage from impalement) and phenol red solution could be

visualized exciting the electrode following pressure pulses from the Picospritzer.

Following impalement of neurons, a 20 PSI pressure pulse was administered

with visualization of the phenol red mRNA solution entering the cell.

Results

Common 3' Untranslated Regions alone are insufficient to generate new mRNA correlations or disrupt existing mRNA correlations in PD neurons

In order to determine if the correlation patterns seen in PD neurons in the

STG of *C. borealis* are the result of common regulatory sites in the 3'-UTR in co-

regulated mRNAs, we created hybrid *eGFP* mRNA transcripts that contained the

3'-UTR of the A-type potassium channel *Shal* (Figure 5.1A), which is known to

form multiple correlations with other mRNAs in PD neurons, particularly with *IH* and *BKKCa* transcripts (Schulz et al., 2007). Pressure injection of hybrid *eGFP-Shal* 3'-UTR mRNA into PD neurons produced visible green fluorescence after 24 hours, indicating successful injection and translation of the hybrid mRNA (Figure 5.1B). However, correlations with the hybrid mRNA and endogenous *Shal* mRNA were not produced (R = -0.097) after 24 hours following qRT-PCR measurements of single neuron mRNA abundances (Figure 5.2). Further, the endogenous correlations seen in PD neurons between *Shal* and *IH* mRNA abundances were not disrupted (R = 0.97) from hybrid mRNA injection.

Actinomycin D disrupts only certain correlations but does not reduce mRNA or rRNA abundances in STG neurons

To investigate if the mRNA correlations seen in STG neurons are the result of regulated transcription that gives rise to specific mRNA ratios, we sought to block transcription and measure the impact on correlated mRNAs. Actinomycin D (ActD) was chosen as a transcription blocker because it is known to disrupt transcription in crustaceans (Eugenia Pedreira, Dimant, & Maldonado, 1996; W.-D. Krenz et al., 2015; Muriel Thoby-Brisson & Simmers, 2000; Zmora, Trant, Zohar, & Chung, 2009). Bath application of STG neurons with 5μg/mL ActD added to the saline in a Vaseline well surrounding the ganglion resulted in a loss of certain correlations in PD neurons, while leaving others intact (Figure 5.3). The *IH-Shal* correlation is robust in PD neurons, as shown in Figure 5.2 (R = 0.97),

but was disrupted following bath application of ActD (R = 0.41). However, other mRNA correlations known to be present in PD neurons persisted, such as the *Shab-Shaker* (R = 0.93) and *INX1-INX2* (R = 0.99) correlations. Most surprisingly, there was no reduction in the amount of rRNA present in the whole STG after treatment.

Injection of innexin mRNA into STG neurons disrupts endogenous innexin correlations

In the neurons of the STG, the innexin genes are responsible for generating the gap junction proteins required for the electrical synapses between different cell types. Regardless of cell type, there are strong correlations in the mRNA abundances of *INX1, INX2,* and *INX3* (See chapter 4). In order to probe the regulation of these innexin mRNA correlations, we disrupted the amount of one innexin mRNA to see the impact on the correlations and overall abundances of the other two innexins. RT-qPCR was used to measure barcoded *INX1* that was injected separately from endogenous *INX1* through subtracting total measured *INX1* from barcoded *INX1*. Through pressure injection of barcoded *in vitro* synthesized *C. borealis INX1* mRNA, a disruption of mRNA correlations in STG neurons was seen at 24 hours that persisted at 72 hours as well (Figure 5.4). For the *INX1-INX2* correlation, the strength of the correlation was greatly reduced following *INX1* mRNA injection at the 24hr (R = 0.55) and 72hr (R = 0.46) relative to the random mRNA injected controls (R = 0.92 and R = 0.86,

respectively). To determine indirect effects of *INX1* injection, we examined the *INX2-INX3* relationship. Unlike the *INX1-INX2* relationship, the *INX2-INX3* relationship was not disrupted at 24hrs (R = 0.95), but degraded by the 72hr time point (R = 0.64). However, this could be due to the reduction in absolute mRNA abundance seen for *INX2* and *INX3* at 72hrs post *INX1* injection (Figure 5.4C).

To determine if the effects seen by injection of *INX1* mRNA are bidirectional, we synthesized and injected barcoded *INX2* mRNA into STG neurons to examine innexin mRNA correlations and abundances (Figure 5.5). Similarly, *INX2* injections disrupted mRNA correlations of endogenous *INX2-INX1* at 24hr (R = 0.17) and 72hr (R = -0.039) time points compared to vehicle injections (R = 0.92 and 0.86, respectively). Further coinciding with the *INX1* injection results, the *INX1-INX3* relationship was weakened following *INX2* injection at the 24hr time point (R = 0.76) and disrupted by 72hr (R = 0.37). Finally, a reduction in *INX1* mRNA abundance was seen at 72hr following *INX2* injection, similar to how *INX2* was reduced at 72hr following *INX1* mRNA injection.

Discussion

The regulation of coordinated of gene expression is an integral part of cellular function required to give cells their unique properties. While coordination

takes place at all levels of gene expression, we have focused here on the transcriptional and post-transcriptional regulation of mRNA abundances that give rise to correlation patterns between transcripts. The 3'-UTRs of mRNAs often contain multiple regulatory sites for miRNAs or RBPs to bind and alter the rate of translation and mRNA stability (Barrett, Fletcher, & Wilton, 2012; Pichon et al., 2012). It has been proposed that multiple mRNAs containing common elements in their 3'-UTRs could act in a post-transcriptional operon, particularly where the mRNAs encode functionally related proteins, through co-regulation of miRNA or RBPs that could give rise to correlated levels of mRNA abundance (Tenenbaum, Christiansen, & Nielsen, 2011). It is the proteins that an mRNA interacts with that most directly determines what will become of it, whether that be translation, degradation, transport, or other fates (Moore, 2005). Indeed, both RBPs and miRNAs could be simultaneously involved in the regulatory processes, as it has been shown that miRNA binding mRNA can alter its affinity for RBPs to bind and regulate the mRNA (Doyle & Tenenbaum, 2014).

Here we examined the role of the 3'-UTR in the formation of correlated levels of mRNA abundances in STG neurons of *C. borealis* by injection of hybrid mRNAs that contained identical 3'-UTR sequence to endogenous transcripts. Since injection of *eGFP-Shal 3'-UTR* mRNAs into PD neurons did not form new correlations, nor did it disrupt existing correlations with *Shal* mRNA, this data suggests that common 3'-UTR sequence alone is insufficient to drive correlated patterns of mRNA. While the 3'-UTR is a prime target for regulation, there are

also regulatory sites in the coding region and 5'UTR that can influence mRNA abundances (E. K. Lee & Gorospe, 2011; Leppek, Das, & Barna, 2017).

We also examined the role of transcription in driving correlated mRNAs in PD neurons. Gene expression under the control of commonly shared promoters, or unique promoters driven by the same transcription factors, could give rise to mRNA correlations (Allocco, Kohane, & Butte, 2004). ActD, an antibiotic compound isolated from *Streptomyces parvullus*, has a long history as a transcription blocker (Perry & Kelley, 1970; Sobell, 1985) that works through DNA intercalation and is often used to study mRNA stability and turnover (Lai, Arvola, & Blackshear, 2019). In crustaceans, ActD has been used to study apoptosis (Leu, Kuo, Kou, & Lo, 2008), regulation of transcription in mitochondria (Eads & Hand, 1999), and changes in intrinsic neuronal excitability (Thoby-Brisson & Simmers, 2000). Here, administration of ActD disrupted mRNA correlations between *Shal* and *IH* while leaving others intact, suggesting different mechanisms by which correlated mRNA abundances arrive in PD neurons. However, the efficacy by which ActD blocked transcription is questionable as there was no significant reduction in the abundance of 18s rRNA following treatment. Class I transcription of rRNA by RNA polymerase I is most sensitive to Actinomycin D treatment in the inhibition of eukaryotic transcription, its mechanism of action is fast, and it can readily pass membranes (Bensaude, 2011). However, the stability of mRNA and rRNA in *C. borealis* is unstudied, and

RNA half-life could be sufficient to not see a significant reduction in RNA abundance following blockade of transcription for highly expressed RNAs.

If there is tight regulation in the abundances of correlated mRNAs, how does a rapid change in the abundance of one mRNA change correlation patterns? In the STG, the innexin gap junction proteins are highly correlated with each other in a non-cell type dependent manner (See Chapter 4). In *Drosophila, Dm-INX2* and *Dm-INX3* are co-expressed and together likely form heteromeric gap junctions with more than one gap junction protein in each hemi-channel (Stebbings, Todman, Phelan, Bacon, & Davies, 2000). Separate injection of *INX1* and *INX2* mRNA both produced similar results where the endogenous correlations were disrupted between other innexin types, as well as a reduction of the amount of other innexin mRNA at 72hr. One interpretation of this result is that post-transcriptional regulation of mRNA abundances are achieved through signaling based on the absolute amount of *INX* mRNA present, and increasing one type of innexin feeds back to reduce the other types of innexin. This could take place at the level of reduced transcription of new innexin mRNA or a decrease in the stability of existing mRNA.

This study is not without its limitations. First, while the 3'-UTR alone was insufficient to generate new mRNA correlations, it is entirely possible that a lack of pre-mRNA processing in the nucleus could have left the injected mRNA without tags, such as protein complexes similar to those seen in nonsense-

mediated decay (Lejeune, 2017), that are necessary to give rise to correlations post-transcriptionally. Second, while ActD treatment of STG neurons did disrupt the *Shal-IH* correlation in PD cells, this mechanism might not rely on transcriptional blockade as no reduction in 18s rRNA was observed, obfuscating the mechanism by which this disruption took place. However, these are necessary starting points for tracking down the molecular mechanisms that drive correlated mRNA patterns.
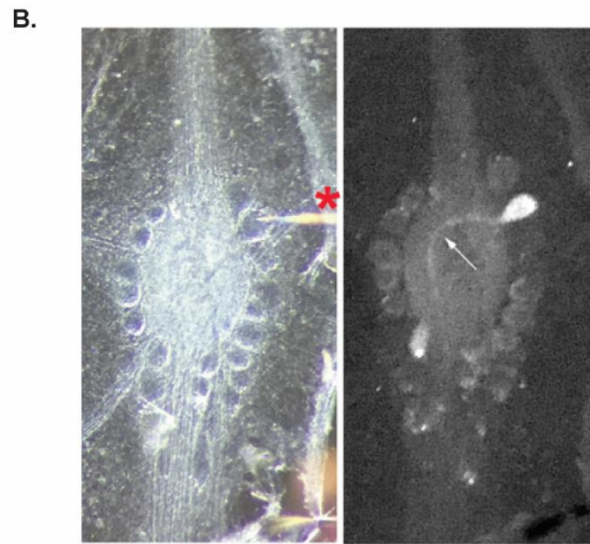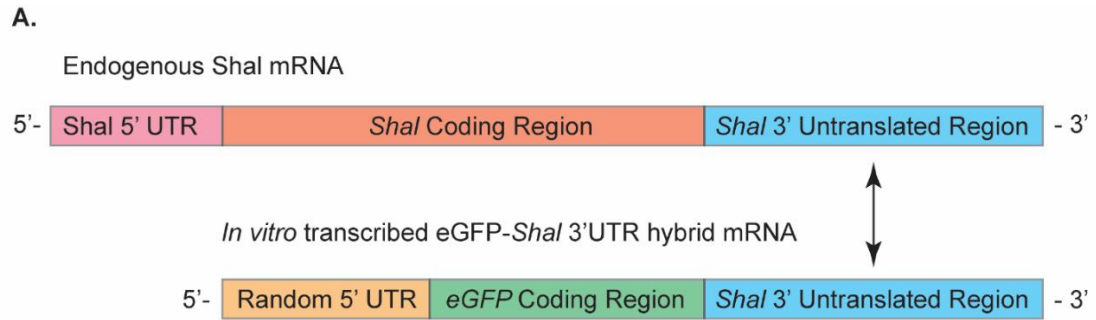
FIGURE 5.1: *Injection of eGFP-Shal 3'-UTR mRNA into PD neurons.* **A.** *In vitro*
transcribed mRNA from a hybrid *eGFP-Shal* 3'-UTR fusion construct is illustrated
showing the only common region between endogenous *Shal* mRNA and the
injected hybrid mRNA is a mutual 3'-UTR sequence. **B.** *eGFP-Shal* 3'-UTR
hybrid mRNA injections yield detectable fluorescence within 24 hours of injection.
The injection electrode (red asterisk, left panel) is used to identify PD neurons
through intracellular activity. Strong eGFP fluorescence (right panel) can be
detected in the cell soma and neurite (white arrow) at 24 hours following

injection. Both PD neurons were injected in this preparation as indicated by the

two glowing neural somata.
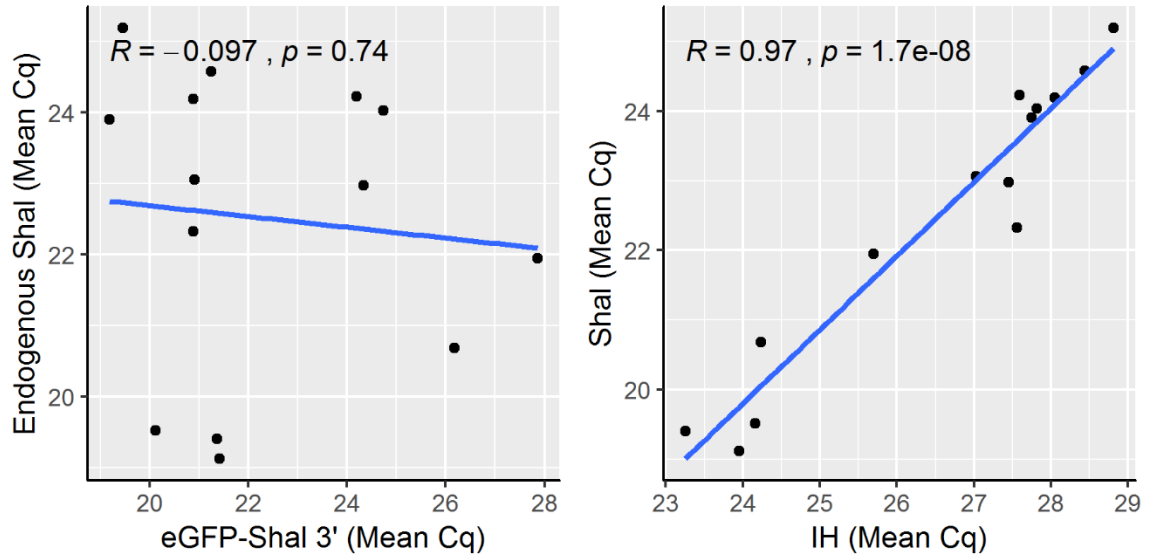
FIGURE 5.2: *mRNA correlations in PD neurons following eGFP-Shal 3'-UTR hybrid mRNA injection.* Pearson correlations of mean Cycle quantity (Cq) values from qRT-PCR measurements from PD neurons (N = 14) 24 hours after injection with hybrid mRNA. *Shal* and *IH* measurements are only for endogenous transcript abundances.

A

CoG   OG   CoG

ActD 5µg/mL

STG   Vaseline
Well

B. PD 8hr ActD

$R = 0.41$, $p = 0.18$

Shal (Mean Cq)

IH (Mean Cq)

C. PD 8hr ActD

$R = 0.93$, $p = 1.5e\text{-}06$

Shaker (Mean Cq)

Shab (Mean Cq)

PD 8hr ActD

$R = 1$, $p = 1.5e\text{-}12$

INX2 (Mean Cq)

INX1 (Mean Cq)

D. ActD Treatement on Whole STG

ns

18s rRNA (Mean Cq)

Condition
ActD
Control

ActD   Control
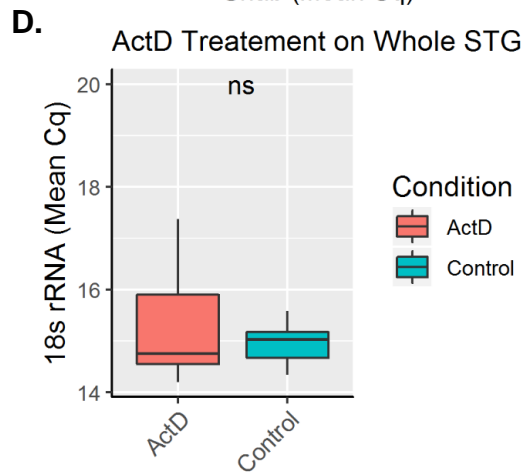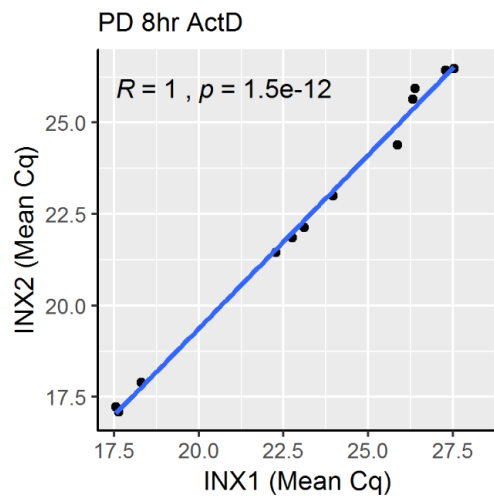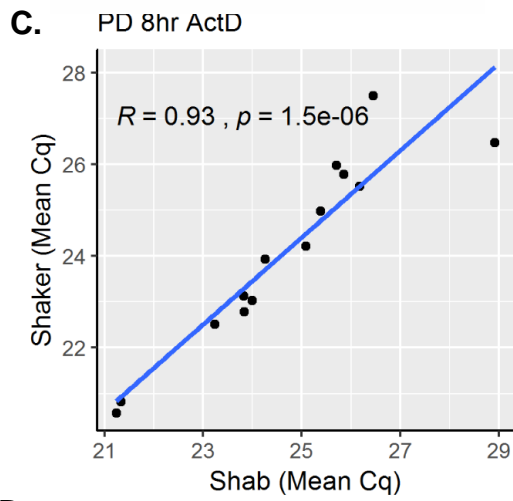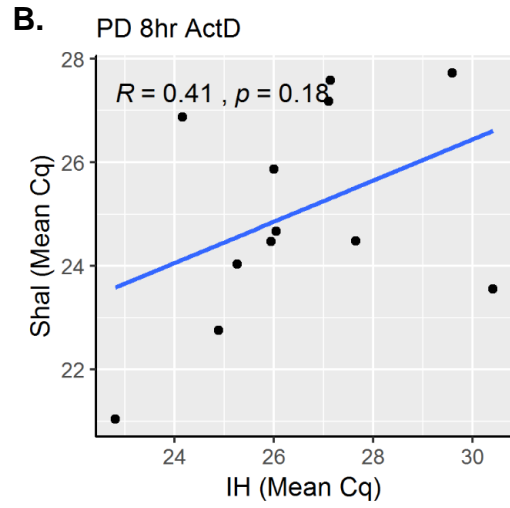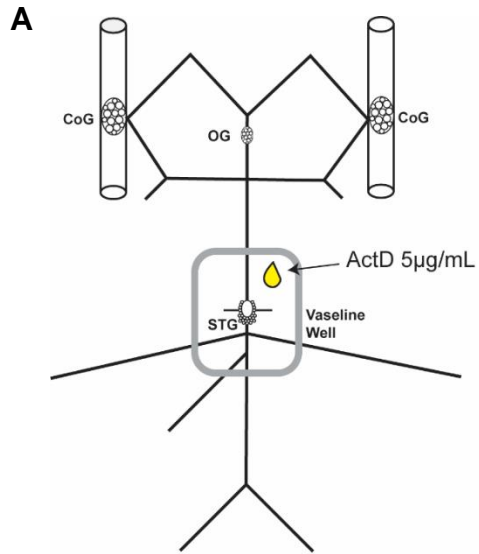
FIGURE 5.3: *Actinomycin-D treatment of STG neurons alters certain mRNA correlations, but did not change total RNA.* **A.** Schematic of the stomatogastric nervous system preparation with stomatogastric ganglion (STG), commissural ganglia (CoGs), and oesophageal ganglion (OG). A Vaseline well is used to isolate the STG for bath application of ActD at a concentration of 5μg/mL. **B.** Disrupted mRNA correlations between *Shal* and *IH* mRNA in PD neurons following bath application of ActD for 8 hours. **C.** mRNA correlations that persisted in PD neurons following ActD treatment. **D.** ActD treatment of whole STGs did not reduce the amount of 18s rRNA present in the ganglion (ns = non-significant).
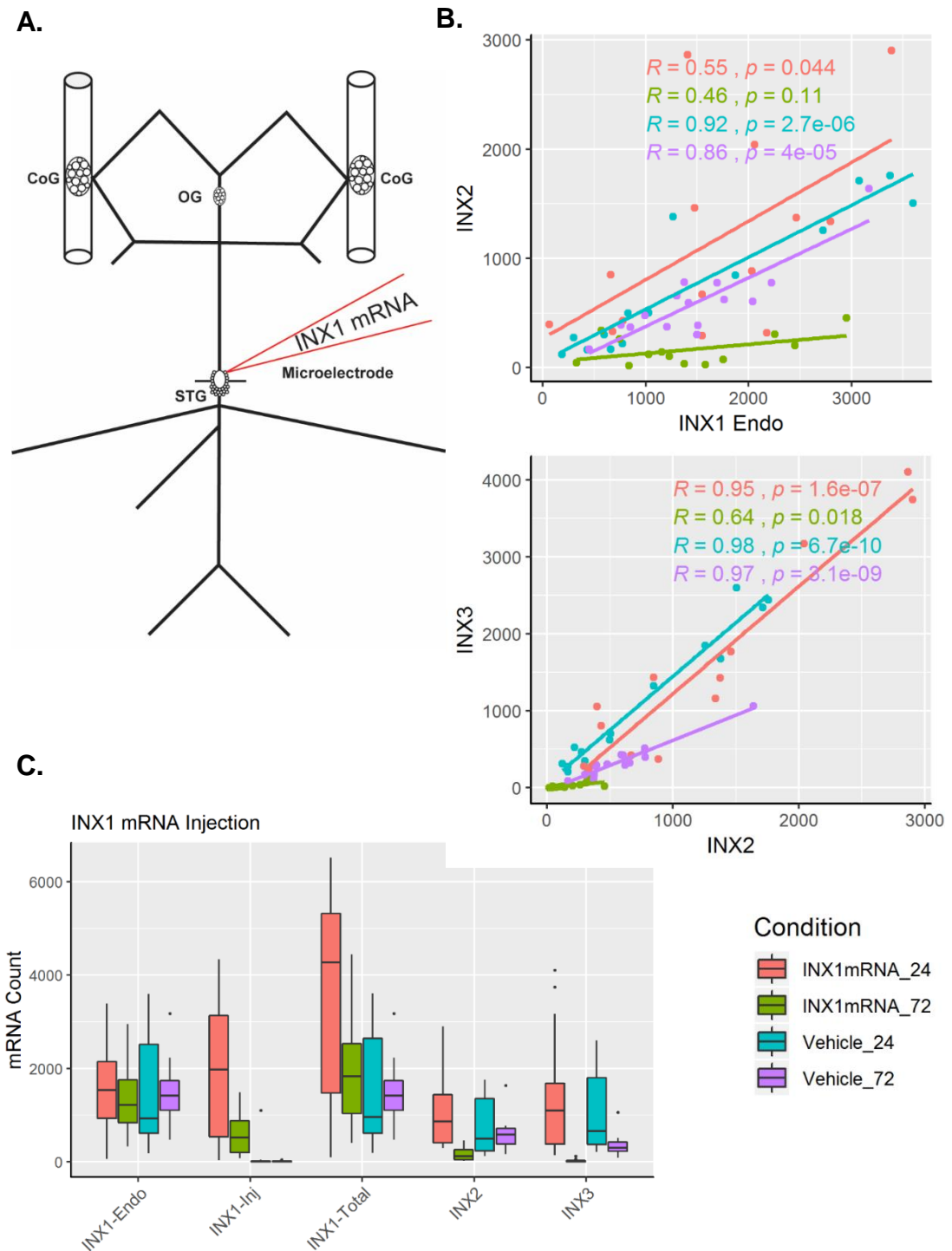
FIGURE 5.4: *Innexin1 mRNA injections into unidentified STG neurons disrupts*

endogenous innexin mRNA correlations. A. A microelectrode is used to inject

barcoded *INX1* mRNA into STG neurons (N = 15) through pressure injection. *B.*

Disrupted innexin mRNA correlations in STG neurons injected with *INX1* mRNA. Endo represents the endogenously expressed *INX1* mRNA. *C.* Absolute innexin mRNA abundances are represented in box plots with "inj" referring to injected barcoded *INX1* and "Total" representing both barcoded and endogenous *INX1* mRNA measured at 24hr and 72hr time points.
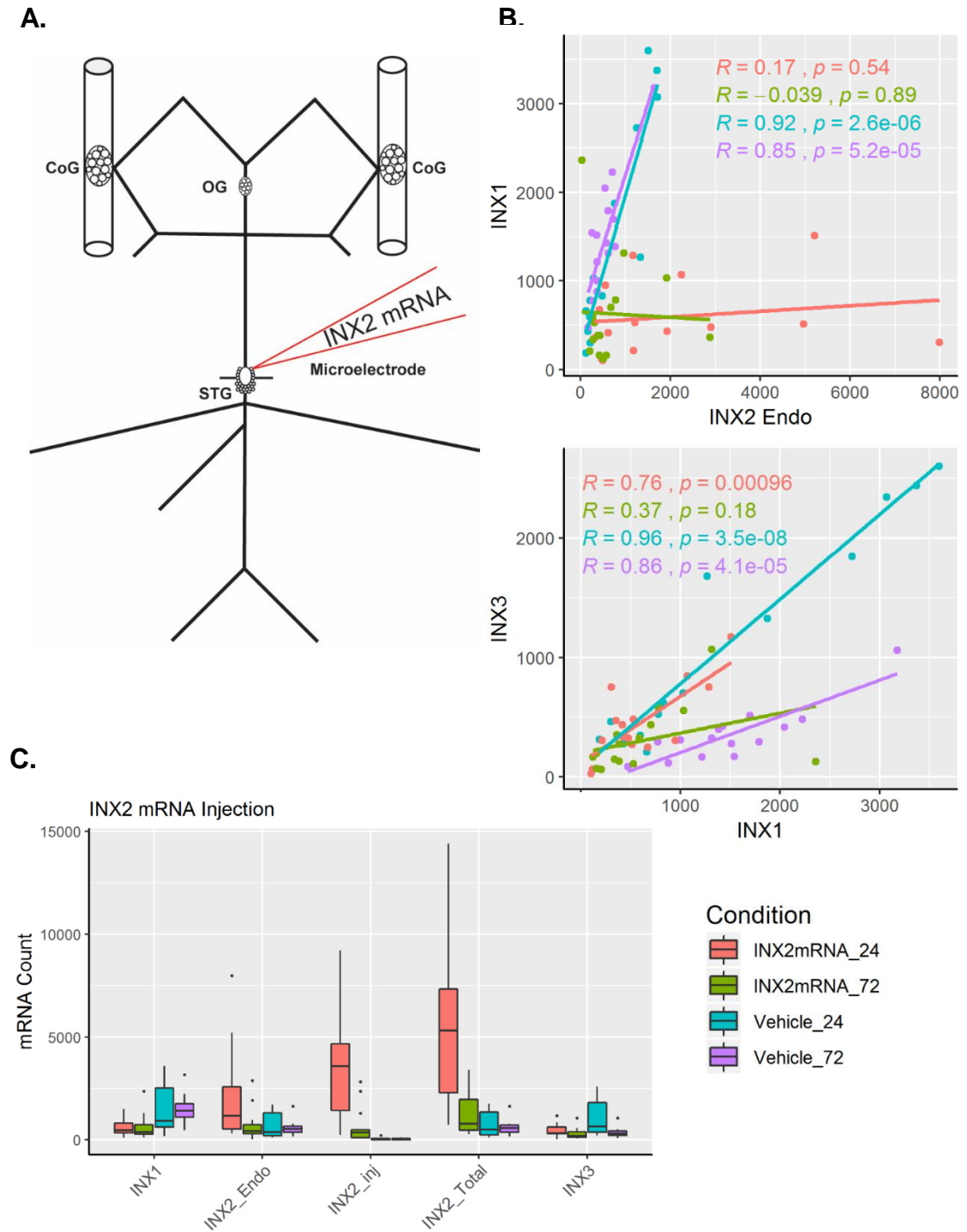
FIGURE 5.5: *Innexin2 mRNA injections into unidentified STG neurons disrupts*

endogenous innexin mRNA correlations. **A.** A microelectrode is used to inject

barcoded *INX2* mRNA into STG neurons (N = 15) through pressure injection. **B.**

Disrupted innexin mRNA correlations in STG neurons injected with *INX2* mRNA.

C. Absolute innexin mRNA abundances are represented in box plots with "inj"

referring to injected barcoded *INX2* and "Total" representing both barcoded and

endogenous *INX2* mRNA measured at 24hr and 72hr time points.

CHAPTER 6

Conclusions and Future Directions:

Advancing genetic manipulation of the stomatogastric neurons system of *Cancer borealis*

This thesis was pursued with the goal of better understanding the molecular mechanisms that give rise to the patterns of mRNA expression that underlie the activity produced in the central pattern generating STG of *C. borealis*. Here in chapter 1, the molecular mechanisms suspected to govern plasticity and homeostasis in central pattern generators was reviewed. The mechanisms examined included activity-dependent, activity-independent, and modulator-dependent pathways for regulating homeostatic plasticity from a molecular standpoint. Activity-dependent pathways utilize a signal mediated through intracellular calcium concentration to regulate the expression of ion channels (Barish, 1998; Mermelstein et al., 2000; Timothy O'Leary et al., 2010; G Turrigiano et al., 1994; Wheeler et al., 2012) and receptors (Greenberg et al., 1986a). Computational models support this signaling pathway with networks can tune their transcription and translation through calcium signals to enable maintenance of homeostatic plasticity for conserving activity (O'Leary et al., 2014). Activity-independent pathways do not utilize a calcium signal, but rather involve more direct regulation of mRNA abundances, such as loss of the A-type

potassium channel *Shal* in *Drosophila* motoneurons resulting in to an increase in

Shaker expression as a compensation (Bergquist et al., 2010). This is also

witnessed in the STG of the spiny lobster where PD neurons were injected with

Shal mRNA, increasing the A-current, but leaving the activity of PD unchanged

through a compensation of hyperpolarization-activated inward current (I_H)

(MacLean et al., 2003). Neuromodulatory signals that can drive homeostatic

plasticity in the STG place tonic dopaminergic signaling as a salient feature that

can increase in fast transient outward K^+ current (I_A) and hyperpolarization-

activated current (I_H) in LP neurons (Rodgers, Fu, et al., 2011; Rodgers, Krenz,

et al., 2011) to maintain activity. This mechanism is dependent on Protein Kinase

A (PKA), RNA polymerase II transcription, mTORC1-dependent translation, and

the RNA interference (RNAi) pathway (W.-D. C. Krenz et al., 2014, 2015).

   In chapter 2, the first whole nervous systems transcriptomes for *Cancer*

borealis and *Homarus americanus* were described. Here the transcriptomes were

thoroughly mined for genes related to the functioning of neural circuits and

modulation, namely the ion channels, receptors, gap junction proteins, and

neurotransmitter-related genes. By examining the transcriptomes of the two

species in parallel, we strengthened the references for each by having a closely

related comparator in discovering crustacean transcripts. Further, these

references have been of significant value to the communities of scientists who

study neural function in these systems. Indeed, we have often received

correspondence regarding the identification of a particular sequence of interest

that an investigator is pursuing. Beyond this, having this transcriptomic reference

set the stage for all the downstream work in my thesis as it allowed for the

quantitation and manipulation of cell-type specific mRNA expression. From this

transcriptomic reference, we have been able to synthesize full length mRNA for

injection and overexpression, dsRNA for targeted knockdown of transcripts to

better understand their function, and realtime PCR assays for measuring the

absolute abundances of mRNAs across tissues and individual cells.

This is only the beginning of improving the genetic knowledge of *C.*

borealis, however, as there are current efforts to generate a complete genome to

better probe the origins of characteristics that give neurons and networks their

functional properties. At present, there is a draft genome reference for *C. borealis*

that was performed with Illumina sequencing technologies to create genomic

scaffolds (Table 6.1). The read information obtained is estimated to currently be

at 100x coverage of the genome, but assembling scaffolds into more complete

chromosomes is proving difficult as the complexity of the genome reduces the

efficacy of short read technology. In future endeavors, combining long-read

technologies with the short-reads already obtained should be able to greatly

reduce the number of scaffolds present in the draft assembly while increasing the

N50 (kb). Of note, the GC content of the nervous system transcriptome (43.4%)

is also similar to that of the draft genome (~45%), as would be expected. With a

more complete genome, investigations into common promoter sites driving

transcription could be informative for the generation of correlated mRNA

expression, as shared transcription factors could drive similar rates of transcription.

TABLE 6.1: *Summary statistics of the draft genome assembly of C. borealis.* Illumina short read sequencing generated paired-end reads that achieved approximately 100x coverage, but further rounds of sequencing were unable to substantially increase the N50 for the scaffolds. Combining long-read technology with the present draft assembly could dramatically reduce the number of scaffolds and increase the N50.

Metric	1 lane, k=71	2 lanes, k=71	2 lanes, k=91
Length	780,361,419	916,882,343	871,927,031
Coverage Estimate	~50X	~100X	~100X
# of Scaffolds	2,499,023	3,004,229	2,173,669
Scaffold N20 (kb)	5,839	6,915	7,308
Scaffold N50 (kb)	1,139	1,194	1,410
% GC Content	46.2%	44.8%	45.2%

In Chapter 3, molecular profiling techniques to recapitulate neuron cell identity were examined using machine learning with qPCR and RNA-seq counts of mRNA abundances. The driving force behind this research is the overarching goal of mapping and defining all the neuron types in the mammalian brain, perhaps one of the loftiest goals in all of science. Many projects currently attempting to describe neuronal cell types begin with the acquisition of molecular

profiles from populations of unidentified neurons (H. Li et al., 2017; Tasic et al., 2016; Usoskin et al., 2015). Here we showed the limitations of unsupervised methods that rely solely on the molecular profile by demonstrating how poorly predicted clusters recapitulated known cell identity. However, once we employed supervised methods incorporating known cell identification and expected cluster numbers the results greatly improved to capturing most of the known identities using only a focal set of genes. Distinguishing STG neurons from one another was only partially successful, reliably clustering 8 of the 11 cell types, and only after data were scaled by gene. LC neurons from the cardiac ganglion of *C. borealis* were arguably the easiest neuron type to be unambiguously separated from the neurons of the STG, suggesting that neurons from different networks have more diverse molecular profiles. Our results argue that many studies would benefit from revisiting previous molecular classifications and incorporating multimodal data alongside the molecular profile to probe more deeply and separate classes with additional granularity.

Taking this analysis further, our qPCR results suggest that if one knows the cell type *a priori*, they should be able to distinguish characteristics of the molecular profile and generate a set of genes that are most useful in separating neurons into discrete cell types. By hand selecting individual transcripts that have strong differential expression across cell types, a panel of genes can be generated that once measured should be able to classify neurons into specific types (Figure 6.1). A panel of 8 genes comprised of *vGluT, GABAB-R1,*

mACHrB, BKKCa, IH, TRP-A1, DAR2, and *Shaw2* has been identified as having

the potential to distinguish the 10 STG neuron types apart. Theoretically, two

multiplex reactions could be run on an individual neuron, and by comparing the

mRNA abundances for this panel of genes against what is known for populations

of identified STG neurons, a confidence value could be generated for how likely it

is that individual neuron belongs to a particular STG cell type. While the

practicality of this is perhaps small since STG neurons are readily identifiable by

simultaneous intracellular and extracellular recording, the practice of using

panels of gene expression as a diagnostic tool is appealing in many systems.

The principles used to create this this diagnostic panel could be more broadly

applicable than the precise panel itself. Particularly, it is interesting that the

genes comprising this panel belong to diverse families, spanning receptor

subtypes, channel subtypes, and neurotransmitter-related genes. That is, it is not

ion channel phenotype or receptor phenotype alone that is sufficient to arrive at

this conclusion. Therefore, the implication is that widely probing the molecular

profile may be more fruitful in generating useful panels than being myopic.
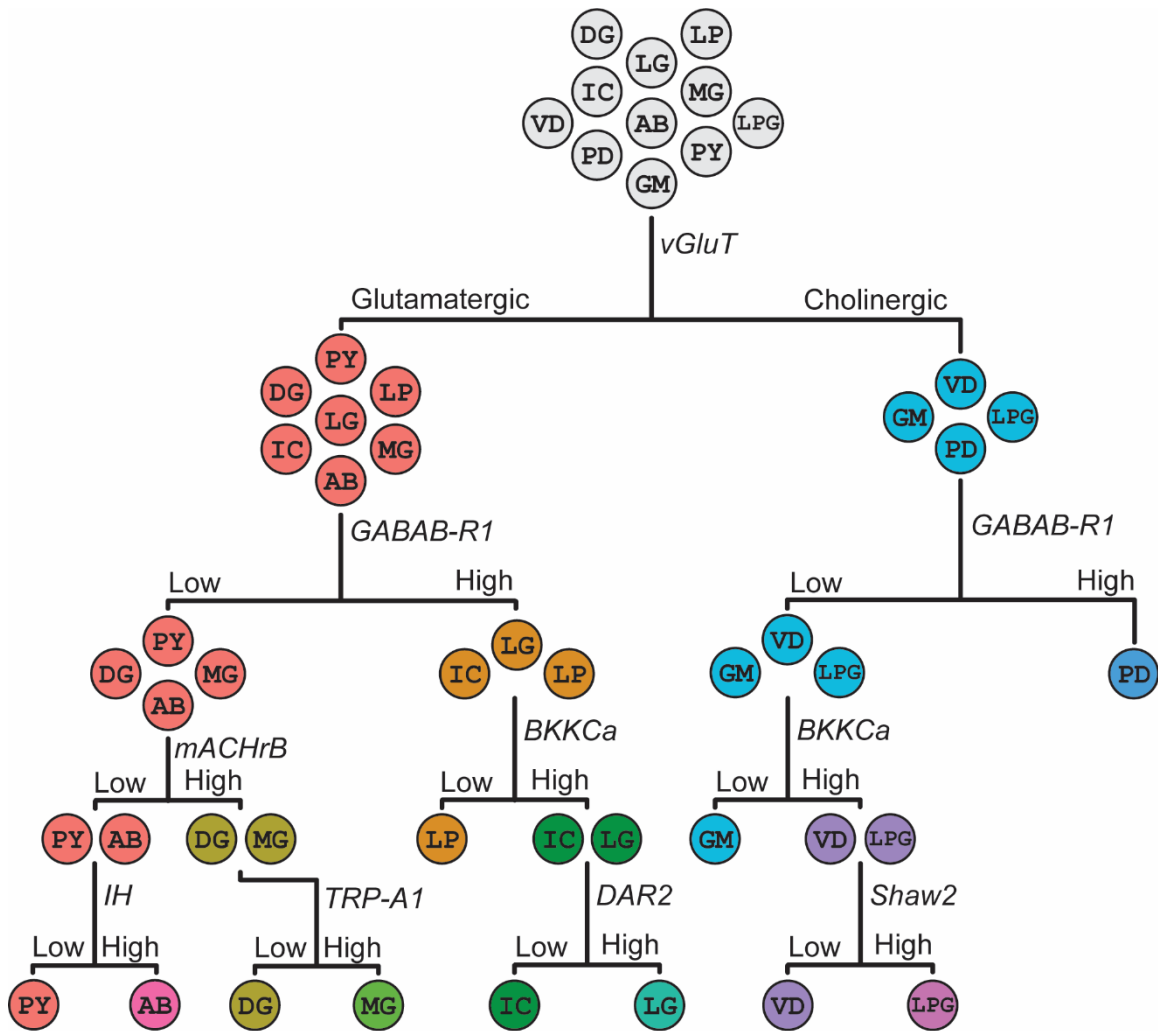
FIGURE 6.1: *Hand-curated separation of STG neuron types by differential expression.* From the qPCR data generated to distinguish neuron cell types, specific transcripts were selected for the degree to which they could separate out neuron types. Shared colors indicate neurons that could not be distinguished from one another at that current level.

In Chapter 4, we examined the expression profile of STG neurons as it relates to their intrinsic electrophysiological properties. Since the STG has a rich literature of over 50 years of STG for network and intrinsic electrophysiological characterization, it is an optimal system to examine relationships between mRNA expression profiles and the known unique properties of each cell type. The molecular profile corroborated many of the expected phenotypes previously examined. The most salient examples include the hyperpolarization-activated inward current (IH) being remarkably high in the AB neuron type compared to all other types as AB is an intrinsically oscillating pacemaking interneuron in the STG that drives the pyloric network, and the I_H produced in AB helps to drive these intrinsic oscillations (Jorge Golowasch et al., 2017). Also, the neurotransmitter phenotypes associated with STG neurons (VD, GM, LPG, and PD being cholinergic with PY, DG, LG, LP, IC, AB, and MG being glutamatergic) was faithfully recapitulated in the mRNA expression if *vGluT, ChAT,* and *vAChT*. Investigations into neuromodulators using CCAP also resulted in a high correspondence in the cell-type specific response to CCAP and the levels of *CCAPr* measured in this study. Beyond considering each mRNA in isolation, cell-type specific correlations previously identified were also bore out in the present investigations with a high rate. Further, the massive scaling of mRNAs measured increased the scope of which cell-type specific correlation patterns can be examined with thousands of pairwise correlations to be considered. This revealed networks of mRNA that are potentially co-regulated in a cell-type specific manner.

Moving this work forward, techniques to better understand the relationship between mRNA and network activity are currently being carried out. With the transcriptomic reference and mRNA synthesis techniques, the generation of dsRNA to greatly knockdown the mRNA abundance in particular cell types has become feasible. This can be used to simultaneously investigate how the knockdown of a targeted mRNA can lead to changes in other co-regulated mRNAs as well as how network activity relies on the expression patterns of specific mRNAs. The first investigations into this involve using *INX1* dsRNA injected directly into STG neurons in order to reduce the levels of *INX1* mRNA through the RNAi silencing pathway. Following injection of *INX1* dsRNA, the triphasic rhythm of the pyloric network was disrupted (Figure 6.2). This is presumably due to a reduction in the efficacy of gap junctions that rely on innexin 1 protein to stably transmit electrical signals directly between coupled neurons. Further investigations using dsRNA to target other transcripts could better reveal the specific relationships between mRNA and neuronal and network activity.

Control rhythm before injection

Immediately after dsRNA injection

24 hours after dsRNA injection

2 s
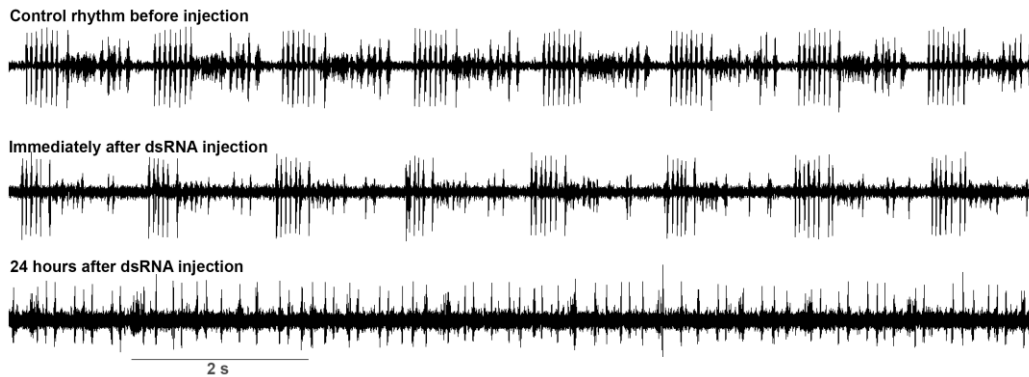
FIGURE 6.2: *Network activity before and after dsRNA injection into STG*

neurons. Top trace shows the control network activity with the characteristic

triphasic rhythm of the pyloric network. *Middle trace* shows network activity

directly after injection of dsRNA into STG neurons. This preparation remained

rhythmic for at least 2 hours after this injection. *Bottom trace* shows the network

activity has lost its rhythm 24 hours after dsRNA injection.

In Chapter 5, the molecular basis of mRNA correlations in STG neurons was directly probed by manipulations to alter the relationships between mRNAs that are known to be correlated in a cell-type specific manner. It was identified that mRNAs sharing a common 3'-UTR alone is insufficient to drive a new correlation in mRNA abundance. This was carried out through the generation of hybrid constructs of eGFP with the 3'-UTR of the A-type potassium channel *Shal* and injection of this hybrid mRNA into PD neurons, which is known to have many correlations with *Shal* mRNA and other ion channels. This injection also did not disrupt the endogenous correlations present between *Shal* and *IH.* To examine the role of transcription on mRNA correlations, STG neurons were treated with actinomycin-D to block transcription of mRNA and rRNA. This was carried out through bath application and direct injection into STG neurons across multiple time points and concentrations. Even though it has been claimed in the literature multiple times that ActD blocks transcription in crustaceans, no significant reduction in mRNA or rRNA was observed with any treatment. This observation warrants further investigation to discern the root of the differences in effects seen across studies. Injection of *INX1* mRNA lead to a decrease in the abundance of *INX2* and *INX3* after 72 hours, but surprisingly did not lead to a reduction in *INX1.* Disruption of endogenous correlations between *INX1* vs *INX2* and *INX1* vs *INX3* were also observed, but this could simply be due to a reduction in the abundance in *INX2* and *INX3* and not a direct change in the mechanism by which correlations are achieved. Further, injection of *INX2* resulted in similar results where *INX1* mRNA was reduced and correlations were disrupted. All together

these results suggest that mRNA correlations are driven by multiple mechanisms and are subject to different levels of regulation.

Beyond experiments to manipulate the mRNA correlations present in STG neurons, techniques are being developed to streamline the process of inducing changes in expression across crustacean systems. In a single preparation, neurons can be inserted with vectors of varying concentrations in a cell-type dependent manner (Figure 6.4A), giving strong experimental control over the system. This can be coupled with intracellular recordings (Figure 6.4B) for identification of neuron types, and subsequent expression of the vector can be visualized in that same neuron. This is not limited to the STG either as manipulations in the cardiac ganglion have seen similar success (Figure 6.4C). Insertion of plasmids with a gene for *eGFP* driven by a cytomegalovirus (CMV) promoter has had success in producing green fluorescence in STG neurons (Figure 6.4D). Co-injection of eGFP mRNA with other vectors can be carried out to ensure successful translation following insertion of desired material (Figure 6.4E).
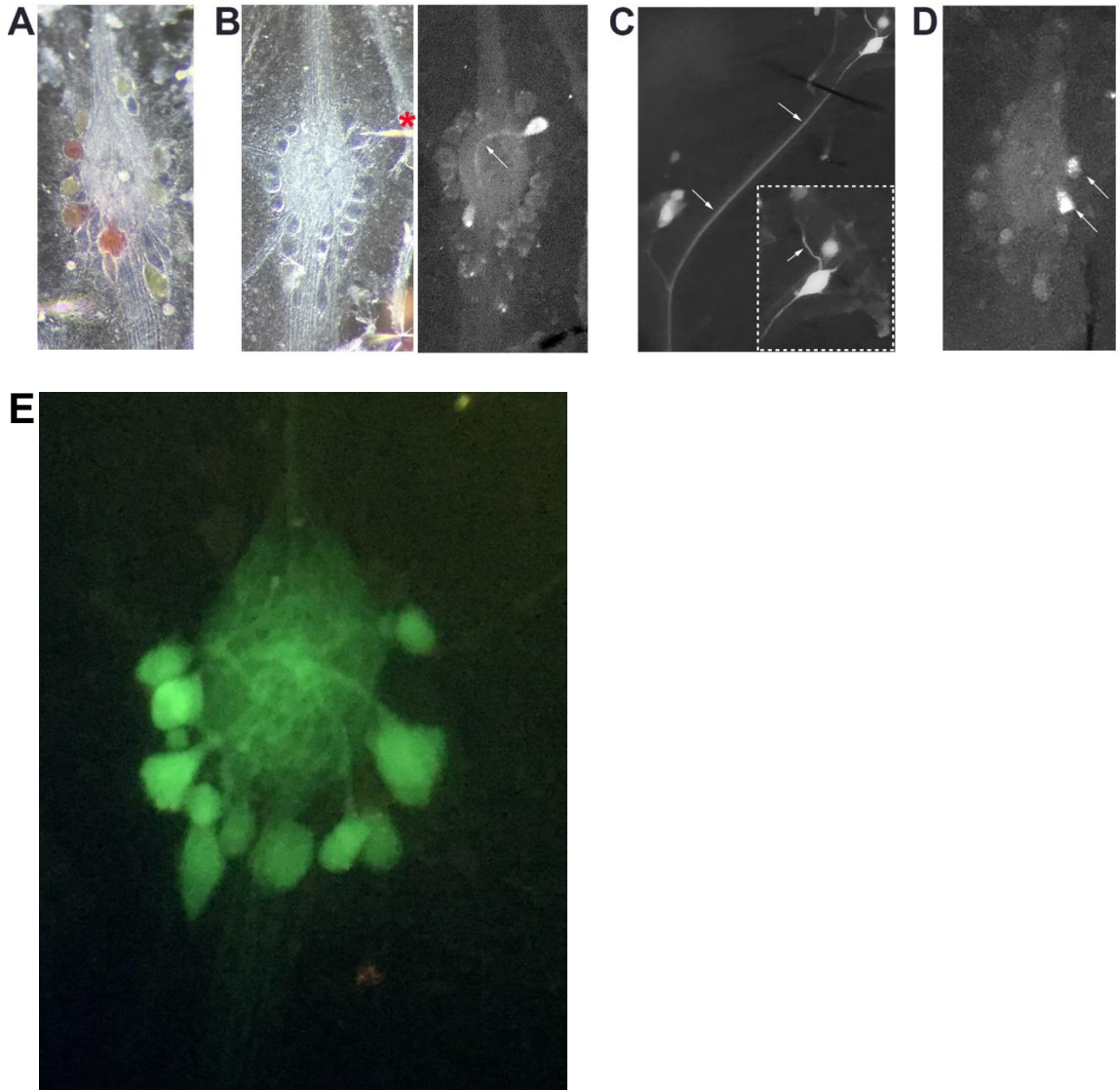
FIGURE 6.3: *Transfection of genetic material and subsequent expression in STG and CG neurons.* **A.** Phenol red solution as a vehicle in the electrode solution allows for visualization of successful injection into STG neurons. **B.** Injections can be coupled with intracellular recording (red asterisk, left panel) to identify neuron types. eGFP mRNA injections yield detectable fluorescence in the soma and neurite (white arrow) within 24 hours of injection. **C.** The crab cardiac ganglion is another system that is amenable to this mRNA injection approach, as

eGFP mRNA are effective at labeling the somata and neurites (white arrows). **D.**

Two cells were injected with plasmid (white arrows) with *eGFP* driven with a

CMV promoter also result in detectable fluorescence within 24 hours. Neurites

are not yet producing visible fluorescence, perhaps due to the time required for

both transcription and translation to take place.

To take these procedures further and make them more readily available to the community at large, the creation of a web interface that would allow for uninitiated users to quickly and easily generate ingredients and protocols for the synthesis of sense and antisense mRNA, as well as dsRNA (Figure 6.5). Following input of a gene of interest, primers can be generated with attached barcodes for qPCR measurements as well as promoters for mRNA synthesis. Promoters can be either input as a custom sequence or selected from a list of popular promoters depending on the RNA-polymerase that the user would prefer to use. For the synthesis of dsRNA, the full length sequence is not necessary to even be known since shorter siRNA will be generated from injected dsRNA through the RNAi-silencing pathway. Only a few hundred nucleotide long sequence is necessary for the generation of dsRNA. Once primers have been generated, the user can also select the preferred methods/kits for carrying out the synthesis and isolation of molecular product. This will allow for a custom protocol to be generated that will guide the user from the first steps in RT-PCR amplification of their target, through synthesis and isolation, and finally injection into their desired cell or tissue. Throughout my education and development, I have found that one of my greatest joys in science is developing new tools and resources for other scientists to use in the pursuit of information, and I plan to keep pursuing this goal in my future endeavors.

Target Gene of Interest - Lookup by name, Accession Number, or direct entry of sequence in FASTA format

Primer Selection Tool generates automated primers and performs QC, exports to format ready for oligo ordering

RT-PCR on RNA extracted from target tissue of interest generates cDNA and ultimately a dsDNA template with T7 and SP6 adapter sequences

In vitro transcription generates sense, antisense, or dsRNA based on the target gene modification strategy.

Injections into target cells delivers RNA via pressure or iontophoresis, preps are monitored for phenotypic effects of altered expression

FIGURE 6.4: *Workflow for the generation of streamlined, ready-to-use molecular products.* Through a web interface, users would input sequences of interest for

the generation of ready to synthesize sense mRNA, antisense mRNA, or dsRNA. Primers with promoters attached as adapters allow for the direct synthesis of molecular products following RT-PCR of cDNA pools containing the target of interest. Protocols will be generated alongside primer outputs to allow a user friendly pathway from synthesis to injection of targets.

BIBLIOGRAPHY

Aalto, A. P., & Pasquinelli, A. E. (2012). Small non-coding RNAs mount a silent revolution in gene expression. *Current Opinion in Cell Biology*, *24*(3), 333–340. http://doi.org/10.1016/J.CEB.2012.03.006

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., … Yamamoto, M.-T. (2000). The genome sequence of *Drosophila melanogaster. Science (New York, N.Y.)*, *287*(5461), 2185–95. http://doi.org/10.1126/science.287.5461.2185

Allocco, D. J., Kohane, I. S., & Butte, A. J. (2004). Quantifying the relationship between co-expression, co-regulation and gene function. *BMC Bioinformatics*, *5*, 18. http://doi.org/10.1186/1471-2105-5-18

Amunts, K., Ebell, C., Muller, J., Telefont, M., Knoll, A., & Lippert, T. (2016). The Human Brain Project: Creating a European Research Infrastructure to Decode the Human Brain. *Neuron*, *92*(3), 574–581. http://doi.org/10.1016/j.neuron.2016.10.046

Armañanzas, R., & Ascoli, G. A. (2015). Towards the automatic classification of neurons. *Trends in Neurosciences*, *38*(5), 307–318. http://doi.org/10.1016/J.TINS.2015.02.004

Ayali, A., Johnson, B. R., & Harris-Warrick, R. M. (1998). Dopamine modulates graded and spike-evoked synaptic inhibition independently at single synapses in pyloric network of lobster. *Journal of Neurophysiology*, *79*, 2063–2069.

Ball, J. M., Franklin, C. C., Tobin, A.-E., Schulz, D. J., & Nair, S. S. (2010). Coregulation of ion channel conductances preserves output in a computational model of a crustacean cardiac motor neuron. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(25), 8637–8649. http://doi.org/10.1523/JNEUROSCI.6435-09.2010

Barish, M. E. (1998). Intracellular calcium regulation of channel and receptor expression in the plasmalemma: potential sites of sensitivity along the pathways linking transcription, translation, and insertion. *Journal of Neurobiology*, *37*(1), 146–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9777738

Barker, D. L., Kushner, P. D., & Hooper, N. K. (1979). Synthesis of dopamine and octopamine in the crustacean stomatogastric nervous system. *Brain Research*, *161*(1), 99–113.

Baro, D. J., Cole, C. L., & Harris-Warrick, R. M. (1996). The lobster shaw gene: Cloning, sequence analysis and comparison to fly shaw. *Gene*, *170*(2), 267–270. http://doi.org/10.1016/0378-1119(95)00876-4

Baro, D. J., Coniglio, L. M., Cole, C. L., Rodriguez, H. E., Lubell, J. K., Kim, M. T., & Harris-Warrick, R. M. (1996). Lobster shal: comparison with *Drosophila* shal and native potassium currents in identified neurons. *Journal of Neuroscience*, *16*(5), 1689–1701.

Baro, D. J., Quiones, L., Lanning, C. C., Harris-Warrick, R. M., & Ruiz, M. (2001). Alternate splicing of the shal gene and the origin of IA diversity among neurons in a dynamic motor network. *Neuroscience*, *106*(2), 419–432. http://doi.org/10.1016/S0306-4522(01)00261-5

Barrett, L. W., Fletcher, S., & Wilton, S. D. (2012). Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cellular and Molecular Life Sciences : CMLS*, *69*(21), 3613–34. http://doi.org/10.1007/s00018-012-0990-9

Baxter, D. A., & Byrne, J. H. (2006). Feeding behavior of *Aplysia*: A model system for comparing cellular mechanisms of classical and operant conditioning. *Learning & Memory*, *13*(6), 669–680. http://doi.org/10.1101/lm.339206

Beg, A. A., & Jorgensen, E. M. (2003). EXP-1 is an excitatory GABA-gated cation channel. *Nature Neuroscience*, *6*(11), 1145–52. http://doi.org/10.1038/nn1136

Beltz, B., Eisen, J. S., Flamm, R., Harris-Warrick, R. M., Hooper, S. L., & Marder, E. (1984). Serotonergic innervation and modulation of the stomatogastric ganglion of three decapod crustaceans (*Panulirus interruptus, Homarus americanus* and *Cancer irroratus*). *The Journal of Experimental Biology*, *109*, 35–54.

Benito, E., & Barco, A. (2015). The Neuronal Activity-Driven Transcriptome. *Molecular Neurobiology*, *51*(3), 1071–1088. http://doi.org/10.1007/s12035-014-8772-z

Bennett, D. J., Gorassini, M., Fouad, K., Sanelli, L., Han, Y., & Cheng, J. (1999). Spasticity in rats with sacral spinal cord injury. *Journal of Neurotrauma*, *16*(1), 69–84. http://doi.org/10.1089/neu.1999.16.69

Bensaude, O. (2011). Inhibiting eukaryotic transcription. Which compound to choose? How to evaluate its activity? *Transcription*, *2*(3), 103–108. http://doi.org/10.4161/trns.2.3.16172

Bergquist, S., Dickman, D. K., & Davis, G. W. (2010). A hierarchy of cell intrinsic and target-derived homeostatic signaling. *Neuron*, *66*(2), 220–34.

http://doi.org/10.1016/j.neuron.2010.03.023

Bertrand, S. S., & Cazalets, J.-R. (2013). Activity-dependent synaptic plasticity and metaplasticity in spinal motor networks. *Current Pharmaceutical Design*, *19*(24), 4498–508. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23360279

Bobkov, Y. V, & Ache, B. W. (2005). Pharmacological properties and functional role of a TRP-related ion channel in lobster olfactory receptor neurons. *Journal of Neurophysiology*, *93*(3), 1372–80. http://doi.org/10.1152/jn.00990.2004

Boldog, E., Bakken, T. E., Hodge, R. D., Novotny, M., Aevermann, B. D., Baka, J., … Tamás, G. (2018). Transcriptomic and morphophysiological evidence for a specialized human cortical GABAergic cell type. *Nature Neuroscience*, *21*(9), 1185–1195. http://doi.org/10.1038/s41593-018-0205-2

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, *34*(5), 525–527. http://doi.org/10.1038/nbt.3519

Brennecke, P., Anders, S., Kim, J. K., Kołodziejczyk, A. A., Zhang, X., Proserpio, V., … Heisler, M. G. (2013). Accounting for technical noise in single-cell RNA-seq experiments. *Nature Methods*, *10*(11), 1093–1095. http://doi.org/10.1038/nmeth.2645

Buckley, S. J., Fitzgibbon, Q. P., Smith, G. G., & Ventura, T. (2016). In silico prediction of the G-protein coupled receptors expressed during the metamorphic molt of *Sagmariasus verreauxi* (Crustacea: Decapoda) by mining transcriptomic data: RNA-seq to repertoire. *General and Comparative Endocrinology*, *228*, 111–127. http://doi.org/10.1016/j.ygcen.2016.02.001

Cadwell, C. R., Palasantza, A., Jiang, X., Berens, P., Deng, Q., Yilmaz, M., … Tolias, A. S. (2015). Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq. *Nature Biotechnology*, *34*(2), 199–203. http://doi.org/10.1038/nbt.3445

Castellucci, V., & Kandel, E. (1976). Presynaptic facilitation as a mechanism for behavioral sensitization in *Aplysia*. *Science*, *194*(4270), 1176–1178. http://doi.org/10.1126/science.11560

Cembrowski, M. S., & Menon, V. (2018). Continuous Variation within Cell Types of the Nervous System. *Trends in Neurosciences*, *41*(6), 337–348. http://doi.org/10.1016/j.tins.2018.02.010

Cembrowski, M. S., Wang, L., Sugino, K., Shields, B. C., & Spruston, N. (2016). Hipposeq: a comprehensive RNA-seq database of gene expression in hippocampal principal neurons. *ELife*, *5*, e14997.

http://doi.org/10.7554/eLife.14997

Christie, A. E. (2016). Prediction of *Scylla olivacea* (Crustacea; Brachyura) peptide hormones using publicly accessible transcriptome shotgun assembly (TSA) sequences. *General and Comparative Endocrinology*, *230–231*, 1–16. http://doi.org/10.1016/j.ygcen.2016.03.008

Christie, A. E., & Chi, M. (2015). Prediction of the neuropeptidomes of members of the Astacidea (Crustacea, Decapoda) using publicly accessible transcriptome shotgun assembly (TSA) sequence data. *General and Comparative Endocrinology*, *224*, 38–60. http://doi.org/10.1016/j.ygcen.2015.06.001

Christie, A. E., Chi, M., Lameyer, T. J., Pascual, M. G., Shea, D. N., Stanhope, M. E., … Dickinson, P. S. (2015). Neuropeptidergic Signaling in the American Lobster *Homarus americanus*: New Insights from High-Throughput Nucleotide Sequencing. *Plos One*, *10*(12), e0145964. http://doi.org/10.1371/journal.pone.0145964

Chung, S., Weber, F., Zhong, P., Tan, C. L., Nguyen, T. N., Beier, K. T., … Dan, Y. (2017). Identification of preoptic sleep neurons using retrograde labelling and gene profiling. *Nature*, *545*(7655), 477–481. http://doi.org/10.1038/nature22350

Claiborne, B. J., & Selverston, A. I. (1984). Histamine as a neurotransmitter in the stomatogastric nervous system of the spiny lobster. *The Journal of Neuroscience*, *4*(3), 708–721.

Clark, M. C., & Baro, D. J. (2006). Molecular cloning and characterization of crustacean type-one dopamine receptors: D1alphaPan and D1betaPan. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, *143*(3), 294–301. http://doi.org/10.1016/j.cbpb.2005.11.017

Clark, M. C., Dever, T. E., Dever, J. J., Xu, P., Rehder, V., Sosa, M. a, & Baro, D. J. (2004). Arthropod 5-HT2 receptors: a neurohormonal receptor in decapod crustaceans that displays agonist independent activity resulting from an evolutionary alteration to the DRY motif. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *24*(13), 3421–35. http://doi.org/10.1523/JNEUROSCI.0062-04.2004

Clark, M. C., Khan, R., & Baro, D. J. (2008). Crustacean dopamine receptors: Localization and G protein coupling in the stomatogastric ganglion. *Journal of Neurochemistry*, *104*(4), 1006–1019. http://doi.org/10.1111/j.1471-4159.2007.05029.x

Cleland, T. A., & Selverston, A. I. (1995). Glutamate-gated inhibitory currents of central pattern generator neurons in the lobster stomatogastric ganglion. *The Journal of Neuroscience : The Official Journal of the Society for*

Neuroscience, *15*(10), 6631–9. Retrieved from
http://www.ncbi.nlm.nih.gov/pubmed/7472424

Cleland, T. A., & Selverston, A. I. (1997). Dopaminergic modulation of inhibitory
glutamate receptors in the lobster stomatogastric ganglion. *Journal of
Neurophysiology*, *78*(1 mM), 3450–3452.

Colbourne, J. K., Pfrender, M. E., Gilbert, D., Thomas, W. K., Tucker, A., Oakley,
T. H., … Boore, J. L. (2011). The ecoresponsive genome of *Daphnia pulex*.
Science (New York, N.Y.), *331*(6017), 555–561.
http://doi.org/10.1126/science.1197761

Collin, C., Hauser, F., Gonzalez de Valdivia, E., de Valdivia, E. G., Li, S.,
Reisenberger, J., … Grimmelikhuijzen, C. J. P. (2013). Two types of
muscarinic acetylcholine receptors in *Drosophila* and other arthropods.
Cellular and Molecular Life Sciences : CMLS, *70*(17), 3231–42.
http://doi.org/10.1007/s00018-013-1334-0

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M.
(2005). Blast2GO: A universal tool for annotation, visualization and analysis
in functional genomics research. *Bioinformatics*, *21*(18), 3674–3676.
http://doi.org/10.1093/bioinformatics/bti610

Connor, J. A. (1975). Neural repetitive firing: a comparative study of membrane
properties of crustacean walking leg axons. *Journal of Neurophysiology*,
38(4), 922–32.

Connor, J. A., Walter, D., & McKown, R. (1977). Neural repetitive firing:
modifications of the Hodgkin-Huxley axon suggested by experimental results
from crustacean axons. *Biophysical Journal*, *18*(1), 81–102.
http://doi.org/10.1016/S0006-3495(77)85598-7

Contractor, A., Mulle, C., & Swanson, G. T. (2011). Kainate receptors coming of
age: milestones of two decades of research. *Trends in Neurosciences*,
34(3), 154–63. http://doi.org/10.1016/j.tins.2010.12.002

Crocker, A., Guan, X.-J., Murphy, C. T., & Murthy, M. (2016). Cell-Type-Specific
Transcriptome Analysis in the Drosophila Mushroom Body Reveals Memory-
Related Changes in Gene Expression. *Cell Reports*, *15*(7), 1580–1596.
http://doi.org/10.1016/j.celrep.2016.04.046

Cuevas-Diaz Duran, R., Wei, H., & Wu, J. Q. (2017). Single-cell RNA-sequencing
of the brain. *Clinical and Translational Medicine*, *6*(1), 20.
http://doi.org/10.1186/s40169-017-0150-9

Cull-Candy, S. G. (1976). Two types of extrajunctional L-glutamate receptors in
locust muscle fibres. *The Journal of Physiology*, *255*(2), 449–64. Retrieved
from http://www.ncbi.nlm.nih.gov/pubmed/1255528

Daur, N., Nadim, F., & Bucher, D. (2016). The complexity of small circuits: the stomatogastric nervous system. *Current Opinion in Neurobiology*, *41*, 1–7. http://doi.org/10.1016/j.conb.2016.07.005

Davie, K., Janssens, J., Koldere, D., De Waegeneer, M., Pech, U., Kreft, Ł., … Aerts, S. (2018). A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain. *Cell*, *174*(4), 982–998.e20. http://doi.org/10.1016/j.cell.2018.05.057

Delcomyn, F. (1980). Neural basis of rhythmic behavior in animals. *Science (New York, N.Y.)*, *210*(4469), 492–8. http://doi.org/10.1126/SCIENCE.7423199

Deneris, E. S., & Hobert, O. (2014). Maintenance of postmitotic neuronal cell identity. *Nature Neuroscience*, *17*(7), 899–907. http://doi.org/10.1038/nn.3731

Diehl, F., White, R. S., Stein, W., & Nusbaum, M. P. (2013). Motor circuit-specific burst patterns drive different muscle and behavior patterns. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(29), 12013–29. http://doi.org/10.1523/JNEUROSCI.1060-13.2013

Doyle, F., & Tenenbaum, S. A. (2014). Trans-regulation of RNA-binding protein motifs by microRNA. *Frontiers in Genetics*, *5*, 79. http://doi.org/10.3389/fgene.2014.00079

Doyle, J. P., Dougherty, J. D., Heiman, M., Schmidt, E. F., Stevens, T. R., Ma, G., … Heintz, N. (2008). Application of a Translational Profiling Approach for the Comparative Analysis of CNS Cell Types. *Cell*, *135*(4), 749–762. http://doi.org/10.1016/J.CELL.2008.10.029

Ducret, E., Alexopoulos, H., Le Feuvre, Y., Davies, J. A., Meyrand, P., Bacon, J. P., & Fénelon, V. S. (2006). Innexins in the lobster stomatogastric nervous system: cloning, phylogenetic analysis, developmental changes and expression within adult identified dye and electrically coupled neurons. *European Journal of Neuroscience*, *24*(11), 3119–3133. http://doi.org/10.1111/j.1460-9568.2006.05209.x

Dudel, J. (1965). Facilitatory effects of 5-hydroxy-tryptamine on the crayfish neuromuscular junction. *Naunyn-Schmiedebergs Archiv Für Experimentelle Pathologie Und Pharmakologie*, *249*(6), 515–528. http://doi.org/10.1007/BF00246558

DUDEL, J., GRYDER, R., KAJI, A., KUFFLER, S. W., & POTTER, D. D. (1963). GAMMA-AMINOBUTYRIC ACID AND OTHER BLOCKING COMPOUNDS IN CRUSTACEA. I. CENTRAL NERVOUS SYSTEM. *Journal of Neurophysiology*, *26*, 721–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14065323

Dudel, J., & KUFFLER, S. W. (1961). The quantal nature of transmission and

spontaneous miniature potentials at the crayfish neuromuscular junction. *The Journal of Physiology*, *155*, 514–29. http://doi.org/10.1113/jphysiol.1961.sp006644

Eads, B. D., & Hand, S. C. (1999). Regulatory features of transcription in isolated mitochondria from *Artemia franciscana* embryos. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *277*(6), R1588–R1597. http://doi.org/10.1152/ajpregu.1999.277.6.R1588

Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., … Coleman, P. (1992). Analysis of gene expression in single live neurons. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(7), 3010–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1557406

Ecker, J. R., Geschwind, D. H., Kriegstein, A. R., Ngai, J., Osten, P., Polioudakis, D., … Zeng, H. (2017). Perspective The BRAIN Initiative Cell Census Consortium: Lessons Learned toward Generating a Comprehensive Brain Cell Atlas. *Neuron*, *96*, 542–557. http://doi.org/10.1016/j.neuron.2017.10.007

Elmer, K. R., Fan, S., Gunter, H. M., Jones, J. C., Boekhoff, S., Kuraku, S., & Meyer, a. (2010). Rapid evolution and selection inferred from the transcriptomes of sympatric crater lake cichlid fishes. *Molecular Ecology*, *19*(SUPPL. 1), 197–211. http://doi.org/10.1111/j.1365-294X.2009.04488.x

Esumi, S., Wu, S.-X., Yanagawa, Y., Obata, K., Sugimoto, Y., & Tamamaki, N. (2008). Method for single-cell microarray analysis and application to gene-expression profiling of GABAergic neuron progenitors. *Neuroscience Research*, *60*(4), 439–451. http://doi.org/10.1016/j.neures.2007.12.011

Eugenia Pedreira, M., Dimant, B., & Maldonado, H. (1996). Inhibitors of protein and RNA synthesis block context memory and long-term habituation in the crab Chasmagnathus. *Pharmacology Biochemistry and Behavior*, *54*(3), 611–617. http://doi.org/10.1016/0091-3057(95)02206-6

Ferraguti, F., & Shigemoto, R. (2006). Metabotropic glutamate receptors. *Cell and Tissue Research*, *326*(2), 483–504. http://doi.org/10.1007/s00441-006-0266-5

Fishell, G., & Heintz, N. (2013). The Neuron Identity Problem: Form Meets Function. *Neuron, 80*(3), 602–612. http://doi.org/10.1016/j.neuron.2013.10.035

Flamm, R. E., & Harris-Warrick, R. M. (1986a). Aminergic modulation in lobster stomatogastric ganglion. I. Effects on motor pattern and activity of neurons within the pyloric circuit. *Journal of Neurophysiology*, *55*(5), 847–865.

Flamm, R. E., & Harris-Warrick, R. M. (1986b). Aminergic modulation in lobster stomatogastric ganglion. II. Target neurons of dopamine, octopamine, and

serotonin within the pyloric circuit. *Journal of Neurophysiology*, *55*(5), 866–881.

FLOREY, E. (1963). ACETYLCHOLINE IN INVERTEBRATE NERVOUS SYSTEMS. *Canadian Journal of Biochemistry and Physiology*, *41*, 2619–26. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14099718

FLOREY, E., & CHAPMAN, D. D. (1961). The non-identity of the transmitter substance of crustacean inhibitory neurons and gamma-aminobutyric acid. *Comparative Biochemistry and Physiology*, *3*, 92–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13700340

Frigon, A., & Rossignol, S. (2006). Functional plasticity following spinal cord lesions. *Progress in Brain Research*, *157*, 231–260. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17167915

Fuccillo, M. V., Földy, C., Gökce, Ö., Rothwell, P. E., Sun, G. L., Malenka, R. C., & Südhof, T. C. (2015). Single-Cell mRNA Profiling Reveals Cell-Type-Specific Expression of Neurexin Isoforms. *Neuron*, *87*(2), 326–340. http://doi.org/10.1016/j.neuron.2015.06.028

Fuchs, P. A., & Getting, P. A. (1980). Ionic basis of presynaptic inhibitory potentials at crayfish claw opener. *Journal of Neurophysiology*, *43*(6), 1547–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7411176

Furshpan, E. J., & Potter, D. D. (1959). Transmission at the giant motor synapses of the crayfish. *The Journal of Physiology*, *145*(2), 289–325. http://doi.org/10.1113/jphysiol.1959.sp006143

Gafurov, B. S., Urazaev, A. K., Grossfeld, R. M., & Lieberman, E. M. (2002). Mechanism of NMDA receptor contribution to axon-to-glia signaling in the crayfish medial giant nerve fiber. *Glia*, *38*(1), 80–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11921205

Galtrey, C. M., & Fawcett, J. W. (2007). The role of chondroitin sulfate proteoglycans in regeneration and plasticity in the central nervous system. *Brain Research Reviews*, *54*(1), 1–18. http://doi.org/10.1016/j.brainresrev.2006.09.006

Getting, P. A. (1983a). Mechanisms of pattern generation underlying swimming in *Tritonia*. II. Network reconstruction. *Journal of Neurophysiology*, *49*(4), 1017–1035.

Getting, P. A. (1983b). Neural control of swimming in Tritonia. *Symposia of the Society for Experimental Biology*, *37*, 89–128. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6679124

Ghazalpour, A., Bennett, B., Petyuk, V. A., Orozco, L., Hagopian, R., Mungrue, I.

N., … Lusis, A. J. (2011). Comparative Analysis of Proteome and Transcriptome Variation in Mouse. *PLoS Genetics*, *7*(6), e1001393. http://doi.org/10.1371/journal.pgen.1001393

Gisselmann, G., Plonka, J., Pusch, H., & Hatt, H. (2004). *Drosophila melanogaster* GRD and LCCH3 subunits form heteromultimeric GABA-gated cation channels. *British Journal of Pharmacology*, *142*(3), 409–13. http://doi.org/10.1038/sj.bjp.0705818

Glanzman, D. L., & Krasne, F. B. (1983). Serotonin and octopamine have opposite modulatory effects on the crayfish's lateral giant escape reaction. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *3*(11), 2263–9.

Glickstein, M. (1991). Cajal on the cerebral cortex: an annotated translation of the complete writings. *Medical History*, *35*(2), 267. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1036352/?page=1

Gokce, O., Stanley, G. M., Treutlein, B., Neff, N. F., Camp, J. G., Malenka, R. C., … Quake, S. R. (2016). Cellular Taxonomy of the Mouse Striatum as Revealed by Single-Cell RNA-Seq. *Cell Reports*, *16*(4), 1126–1137. http://doi.org/10.1016/j.celrep.2016.06.059

Golowasch, J., Bose, A., Guan, Y., Salloum, D., Roeser, A., & Nadim, F. (2017). A balance of outward and linear inward ionic currents is required for generation of slow-wave oscillations. *Journal of Neurophysiology*, *118*(2), 1092–1104. http://doi.org/10.1152/jn.00240.2017

Golowasch, J., Goldman, M. S., Abbott, L. F., & Marder, E. (2002). Failure of averaging in the construction of a conductance-based neuron model. *Journal of Neurophysiology*, *87*(2), 1129–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11826077

Golowasch, J., & Marder, E. (1992). Ionic currents of the lateral pyloric neuron of the stomatogastric ganglion of the crab. *Journal of Neurophysiology*, *67*(2), 318–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1373762

Greenberg, M. E., Ziff, E. B., & Greene, L. A. (1986a). Stimulation of neuronal acetylcholine receptors induces rapid gene transcription. *Science (New York, N.Y.)*, *234*(4772), 80–3. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3749894

Greenberg, M. E., Ziff, E. B., & Greene, L. A. (1986b). Stimulation of neuronal acetylcholine receptors induces rapid gene transcription. *Science (New York, N.Y.)*, *234*(4772), 80–3. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3749894

Greer, P. L., & Greenberg, M. E. (2008). From Synapse to Nucleus: Calcium-

Dependent Gene Transcription in the Control of Synapse Development and Function. *Neuron*, *59*(6), 846–860. http://doi.org/10.1016/j.neuron.2008.09.002

Grundfest, H., Reuben, J. P., & Rickles, W. H. (1959). The electrophysiology and pharmacology of lobster neuromuscular synapses. *The Journal of General Physiology*, *42*(6), 1301–23.

Gu, X. C., Zhang, Y. N., Kang, K., Dong, S. L., & Zhang, L. W. (2015). Antennal transcriptome analysis of odorant reception genes in the red turpentine beetle (RTB), *Dendroctonus valens*. *PLoS ONE*, *10*(5), e0125159. http://doi.org/10.1371/journal.pone.0125159

Guerra, L., McGarry, L. M., Robles, V., Bielza, C., Larrañaga, P., & Yuste, R. (2011). Comparison between supervised and unsupervised classifications of neuronal cell types: A case study. *Developmental Neurobiology*, *71*(1), 71–82. http://doi.org/10.1002/dneu.20809

Guertin, P. A. (2009). The mammalian central pattern generator for locomotion. *Brain Research Reviews*, *62*(1), 45–56. http://doi.org/10.1016/J.BRAINRESREV.2009.08.002

Gutierrez, G. J., O'Leary, T., & Marder, E. (2013). Multiple Mechanisms Switch an Electrically Coupled, Synaptically Inhibited Neuron between Competing Rhythmic Oscillators. *Neuron*, *77*(5), 845–858. http://doi.org/10.1016/j.neuron.2013.01.016

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–512. http://doi.org/10.1038/nprot.2013.084

Hamilton, D. J., Shepherd, G. M., Martone, M. E., & Ascoli, G. A. (2012). An ontological approach to describing neurons and their relationships. *Frontiers in Neuroinformatics*, *6*, 15. http://doi.org/10.3389/fninf.2012.00015

Hamood, A. W., & Marder, E. (2014). Animal-to-animal variability in neuromodulation and circuit function. *Cold Spring Harbor Symposia on Quantitative Biology*, *79*, 21–28. http://doi.org/10.1101/sqb.2014.79.024828

Harms, L., Frickenhaus, S., Schiffer, M., Mark, F. C., Storch, D., Held, C., … Lucassen, M. (2014). Gene expression profiling in gills of the great spider crab *Hyas araneus* in response to ocean acidification and warming. *BMC Genomics*, *15*(1), 789. http://doi.org/10.1186/1471-2164-15-789

Harms, L., Frickenhaus, S., Schiffer, M., Mark, F. C., Storch, D., Pörtner, H. O., … Lucassen, M. (2013). Characterization and analysis of a transcriptome from the boreal spider crab Hyas araneus. *Comparative Biochemistry and*

Physiology - Part D: Genomics and Proteomics, *8*(4), 344–351. http://doi.org/10.1016/j.cbd.2013.09.004

Harris-Warrick, R. M., & Johnson, B. R. (2010). Checks and balances in neuromodulation. *Frontiers in Behavioral Neuroscience*, *4*(July), 1–9. http://doi.org/10.3389/fnbeh.2010.00047

Harris, T. W., Baran, J., Bieri, T., Cabunoc, A., Chan, J., Chen, W. J., … Sternberg, P. W. (2014). WormBase 2014: new views of curated biology. *Nucleic Acids Research*, *42*(D1), D789–D793. http://doi.org/10.1093/nar/gkt1063

Hartline, D. (1967). Impulse identification and axon mapping of the nine neurons in the cardiac ganglion of the lobster *Homarus americanus*. *Journal of Experimental Biology*, *47*(2), 327–341.

Hasan, A., Cotobal, C., Duncan, C. D. S., & Mata, J. (2014). Systematic analysis of the role of RNA-binding proteins in the regulation of RNA stability. *PLoS Genetics*, *10*(11), e1004684. http://doi.org/10.1371/journal.pgen.1004684

Heaton, B., Decker, C., Muhlrad, D., Donahue, J., Jacobson, A., & Parker, R. (1992). Analysis of chimeric mRNAs derived from the STE3 mRNA identifies multiple regions within yeast mRNAs that modulate mRNA decay. *Nucleic Acids Research*, *20*(20), 5365–73. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1437553

Heiman, M., Schaefer, A., Gong, S., Peterson, J. D., Day, M., Ramsey, K. E., … Heintz, N. (2008). A Translational Profiling Approach for the Molecular Characterization of CNS Cell Types. *Cell*, *135*(4), 738–748. http://doi.org/10.1016/J.CELL.2008.10.028

Heitler, W. J. (1978). Coupled motoneurones are part of the crayfish swimmeret central oscillator. *Nature*, *275*(5677), 231–4.

Herrick, D., Parker, R., & Jacobson, A. (1990). Identification and comparison of stable and unstable mRNAs in Saccharomyces cerevisiae. *Molecular and Cellular Biology*, *10*(5), 2269–84. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2183028

Hildebrand, J. G., Townsel, J. G., & Kravitz, E. A. (1974). Distribution of acetylcholine, choline, choline acetyltransferase and acetylcholinesterase in regions and single identified axons of the lobster nervous system. *Journal of Neurochemistry*, *23*(5), 951–63. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4215870

Hirono, C., Ito, I., Yamagishi, S., & Sugiyama, H. (1988). Characterization of glutamate receptors induced in Xenopus oocytes after injection of rat brain mRNA. *Neuroscience Research*, *6*(2), 106–14. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/2905791

Ho, H., Both, M. De, Siniard, A., Sharma, S., Notwell, J. H., Wallace, M., … Portmann, T. (2018). A Guide to Single-Cell Transcriptomics in Adult Rodent Brain: The Medium Spiny Neuron Transcriptome Revisited. *Frontiers in Cellular Neuroscience*, *12*, 159. http://doi.org/10.3389/fncel.2018.00159

Hobert, O., Carrera, I., & Stefanakis, N. (2010). The molecular and gene regulatory signature of a neuron. *Trends in Neurosciences*, *33*(10), 435–45. http://doi.org/10.1016/j.tins.2010.05.006

Hobert, O., Glenwinkel, L., & White, J. (2016). Revisiting Neuronal Cell Type Classification in Caenorhabditis elegans. *Current Biology : CB*, *26*(22), R1197–R1203. http://doi.org/10.1016/j.cub.2016.10.027

Hooper, S. L., O'Neil, M. B., Wagner, R., Ewer, J., Golowasch, J., & Marder, E. (1986). The innervation of the pyloric region of the crab, Cancer borealis: homologous muscles in decapod species are differently innervated. *Journal of Comparative Physiology. A, Sensory, Neural, and Behavioral Physiology*, *159*(2), 227–40. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2876096

Hoopfer, E. D., Jung, Y., Inagaki, H. K., Rubin, G. M., & Anderson, D. J. (2015). P1 interneurons promote a persistent internal state that enhances inter-male aggression in Drosophila. *ELife*, *4*. http://doi.org/10.7554/eLife.11346

Hudson, A. E., Gollnick, C., Gourdine, J.-P., & Prinz, A. A. (2015). Degradation of extracellular chondroitin sulfate delays recovery of network activity after perturbation. *Journal of Neurophysiology*, *114*(2), 1346–52. http://doi.org/10.1152/jn.00455.2015

Hultborn, H. (2003). Changes in neuronal properties and spinal reflexes during development of spasticity following spinal cord lesions and stroke: studies in animal models and patients. *Journal of Rehabilitation Medicine*, (41 Suppl), 46–55. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12817657

Ito, H., Fujitani, K., Usui, K., Shimizu-Nishikawa, K., Tanaka, S., & Yamamoto, D. (1996). Sexual orientation in Drosophila is altered by the satori mutation in the sex-determination gene fruitless that encodes a zinc finger protein with a BTB domain. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(18), 9687–92. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8790392

Ito, H., Sato, K., Koganezawa, M., Ote, M., Matsumoto, K., Hama, C., & Yamamoto, D. (2012). Fruitless Recruits Two Antagonistic Chromatin Factors to Establish Single-Neuron Sexual Dimorphism. *Cell*, *149*(6), 1327–1338. http://doi.org/10.1016/j.cell.2012.04.025

Iversen, L. L., Kravitz, E. A., & Otsuka, M. (1967). Release of gamma-aminobutyric acid (GABA) from lobster inhibitory neurones. *The Journal of Physiology*, *188*(2), 21P–22P.

Jezzini, S. H., Reyes-Colón, D., & Sosa, M. A. (2014). Characterization of a prawn OA/TA receptor in *Xenopus* oocytes suggests functional selectivity between octopamine and tyramine. *PloS One*, *9*(10), e111314. http://doi.org/10.1371/journal.pone.0111314

Jiménez-Vázquez, E. N., Díaz-Velásquez, C. E., Uribe, R. M., Arias, J. M., & García, U. (2016). Molecular cloning and expression of a GABA receptor subunit from the crayfish *Procambarus clarkii*. *Journal of Neuroscience Research*, *94*(2), 190–203. http://doi.org/10.1002/jnr.23695

Johnson, B. R., Kloppenburg, P., & Harris-Warrick, R. M. (2003). Dopamine modulation of calcium currents in pyloric neurons of the lobster stomatogastric ganglion. *Journal of Neurophysiology*, *90*(2), 631–643. http://doi.org/10.1152/jn.00037.2003

Johnson, B. R., Schneider, L. R., Nadim, F., & Harris-Warrick, R. M. (2005). Dopamine modulation of phasing of activity in a rhythmic motor network: contribution of synaptic and intrinsic modulatory actions. *Journal of Neurophysiology*, *94*(5), 3101–3111. http://doi.org/10.1152/jn.00440.2005

Kadiri, L. R., Kwan, A. C., Webb, W. W., & Harris-Warrick, R. M. (2011). Dopamine-induced oscillations of the pyloric pacemaker neuron rely on release of calcium from intracellular stores. *Journal of Neurophysiology*, *106*(3). Retrieved from http://jn.physiology.org/content/106/3/1288

Katz, P. S. (2016). Evolution of central pattern generators and rhythmic behaviours. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1685), 20150057. http://doi.org/10.1098/rstb.2015.0057

Kelley, K. W., Nakao-Inoue, H., Molofsky, A. V., & Oldham, M. C. (2018). Variation among intact tissue samples reveals the core transcriptional features of human CNS cell classes. *Nature Neuroscience*, *21*(9), 1171–1184. http://doi.org/10.1038/s41593-018-0216-z

Kiehn, O., & Harris-Warrick, R. M. (1992). 5-HT modulation of hyperpolarization-activated inward current and calcium-dependent outward current in a crustacean motor neuron. *Journal of Neurophysiology*, *68*(2), 496–508.

Kilman, V. L., & Marder, E. (1996). Ultrastructure of the stomatogastric ganglion neuropil of the crab, Cancer borealis. *The Journal of Comparative Neurology*, *374*(3), 362–75. http://doi.org/10.1002/(SICI)1096-9861(19961021)374:3<362::AID-CNE5>3.0.CO;2-#

Kim, B. G., Dai, H.-N., Lynskey, J. V, McAtee, M., & Bregman, B. S. (2006). Degradation of chondroitin sulfate proteoglycans potentiates transplant-mediated axonal remodeling and functional recovery after spinal cord injury in adult rats. *The Journal of Comparative Neurology*, *497*(2), 182–98. http://doi.org/10.1002/cne.20980

Kim, M., Baro, D. J., Lanning, C. C., Doshi, M., Moskowitz, H. S., Farnham, J., & Harris-Warrick, R. M. (1998). Expression of *Panulirus* shaker potassium channel splice variants. *Receptors and Channels*, *5*(5), 291–304.

King, A. E., & Wheal, H. V. (1984). The excitatory actions of kainic acid and some derivatives at the crab neuromuscular junction. *European Journal of Pharmacology*, *102*(1), 129–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6148242

Kodama, T., Guerrero, S., Shin, M., Moghadam, S., Faulstich, M., & du Lac, S. (2012). Neuronal classification and marker gene identification via single-cell expression profiling of brainstem vestibular neurons subserving cerebellar learning. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(23), 7819–31. http://doi.org/10.1523/JNEUROSCI.0543-12.2012

Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C., & Teichmann, S. A. (2015). The Technology and Biology of Single-Cell RNA Sequencing. *Molecular Cell*, *58*(4), 610–620. http://doi.org/10.1016/J.MOLCEL.2015.04.005

Krasne, F. B., & Edwards, D. H. (2002). Modulation of the crayfish escape reflex-physiology and neuroethology. *Integrative and Comparative Biology*, *42*(4), 705–15. http://doi.org/10.1093/icb/42.4.705

Kravitz, E. A. (2000). Serotonin and aggression: insights gained from a lobster model system and speculations on the role of amine neurons in a complex behavior. *Journal of Comparative Physiology. A, Sensory, Neural, and Behavioral Physiology*, *186*(3), 221–38.

Kravitz, E., Potter, D. D., & Van Gelder, N. (1962). Gamma-aminobutyric acid and other blocking substances extracted from crab muscle. *Nature*, *194*, 382–3. http://doi.org/10.1038/194382b0

Krenz, W.-D. C., Parker, A. R., Rodgers, E. W., & Baro, D. J. (2014). Dopaminergic tone persistently regulates voltage-gated ion current densities through the D1R-PKA axis, RNA polymerase II transcription, RNAi, mTORC1, and translation. *Frontiers in Cellular Neuroscience*, *8*, 39. http://doi.org/10.3389/fncel.2014.00039

Krenz, W.-D. C., Rodgers, E. W., Baro, D. J., Selverston, A., & Liu, H. (2015). Tonic 5nM DA Stabilizes Neuronal Output by Enabling Bidirectional Activity-

Dependent Regulation of the Hyperpolarization Activated Current via PKA and Calcineurin. *PLOS ONE, 10*(2), e0117965. http://doi.org/10.1371/journal.pone.0117965

Krenz, W.-D., Parker, A. R., Rodgers, E., & Baro, D. J. (2015). Monoaminergic tone supports conductance correlations and stabilizes activity features in pattern generating neurons of the lobster, *Panulirus interruptus. Frontiers in Neural Circuits*, *9*, 63. http://doi.org/10.3389/fncir.2015.00063

Krenz, W. D., Nguyen, D., Pérez-Acevedo, N. L., & Selverston, A. I. (2000). Group I, II, and III mGluR compounds affect rhythm generation in the gastric circuit of the crustacean stomatogastric ganglion. *Journal of Neurophysiology*, *83*(3), 1188–201. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10712449

Kristensen, A. R., Gsponer, J., & Foster, L. J. (2014). Protein synthesis rate is the predominant regulator of protein expression during differentiation. *Molecular Systems Biology*, *9*(1), 689–689. http://doi.org/10.1038/msb.2013.47

Lai, W. S., Arvola, R. M., & Blackshear, P. J. (2019). Inhibiting transcription in cultured metazoan cells with actinomycin D to monitor mRNA turnover. *Methods*, *155*, 77–87. http://doi.org/10.1016/J.YMETH.2019.01.003

Lee, E. K., & Gorospe, M. (2011). Coding region: the neglected post-transcriptional code. *RNA Biology*, *8*(1), 44–8. http://doi.org/10.4161/RNA.8.1.13863

Lee, Y., Lee, Y., Lee, J., Bang, S., Hyun, S., Kang, J., … Kim, J. (2005). Pyrexia is a new thermal transient receptor potential channel endowing tolerance to high temperatures in Drosophila melanogaster. *Nature Genetics*, *37*(3), 305–10. http://doi.org/10.1038/ng1513

Lejeune, F. (2017). Nonsense-mediated mRNA decay at the crossroads of many cellular pathways. *BMB Reports*, *50*(4), 175–185. http://doi.org/10.5483/BMBREP.2017.50.4.015

Leppek, K., Das, R., & Barna, M. (2017). Functional 5′ UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nature Reviews Molecular Cell Biology*, *19*(3), 158–174. http://doi.org/10.1038/nrm.2017.103

Leu, J.-H., Kuo, Y.-C., Kou, G.-H., & Lo, C.-F. (2008). Molecular cloning and characterization of an inhibitor of apoptosis protein (IAP) from the tiger shrimp, Penaeus monodon. *Developmental & Comparative Immunology*, *32*(2), 121–133. http://doi.org/10.1016/j.dci.2007.05.005

Li, G.-W., & Xie, X. S. (2011). Central dogma at the single-molecule level in living cells. *Nature, 475*(7356), 308–315. http://doi.org/10.1038/nature10315

Li, H., Horns, F., Wu, B., Xie, Q., Li, J., Li, T., … Luo, L. (2017). Classifying Drosophila Olfactory Projection Neuron Subtypes by Single-Cell RNA Sequencing. *Cell*, *171*(5), 1206–1220.e22. http://doi.org/10.1016/j.cell.2017.10.019

Li, J. J., & Biggin, M. D. (2015). Gene expression. Statistics requantitates the central dogma. *Science (New York, N.Y.)*, *347*(6226), 1066–7. http://doi.org/10.1126/science.aaa8332

Lind, U., Alm Rosenblad, M., Hasselberg Frank, L., Falkbring, S., Brive, L., Laurila, J. M., … Blomberg, A. (2010). Octopamine receptors from the barnacle *Balanus improvisus* are activated by the alpha2-adrenoceptor agonist medetomidine. *Molecular Pharmacology*, *78*(2), 237–248. http://doi.org/10.1124/mol.110.063594

Liu, Y., Beyer, A., & Aebersold, R. (2016). On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell*, *165*(3), 535–550. http://doi.org/10.1016/j.cell.2016.03.014

Livingstone, M. S., Harris-Warrick, R. M., & Kravitz, E. A. (1980). Serotonin and octopamine produce opposite postures in lobsters. *Science (New York, N.Y.)*, *208*(4439), 76–9. http://doi.org/10.1126/science.208.4439.76

Llorens-Bobadilla, E., Zhao, S., Baser, A., Saiz-Castro, G., Zwadlo, K., & Martin-Villalba, A. (2015). Single-Cell Transcriptomics Reveals a Population of Dormant Neural Stem Cells that Become Activated upon Brain Injury. *Cell Stem Cell*, *17*(3), 329–340. http://doi.org/10.1016/j.stem.2015.07.002

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. http://doi.org/10.1186/s13059-014-0550-8

Lu, B., Su, Y., Das, S., Liu, J., Xia, J., & Ren, D. (2007). The neuronal channel NALCN contributes resting sodium permeability and is required for normal respiratory rhythm. *Cell*, *129*(2), 371–83. http://doi.org/10.1016/j.cell.2007.02.041

Luo, L., Callaway, E. M., & Svoboda, K. (2018). Genetic Dissection of Neural Circuits: A Decade of Progress. *Neuron*, *98*(2), 256–281. http://doi.org/10.1016/J.NEURON.2018.03.040

MacLean, J. N., Zhang, Y., Johnson, B. R., & Harris-Warrick, R. M. (2003). Activity-independent homeostasis in rhythmically active neurons. *Neuron*, *37*(1), 109–120. http://doi.org/10.1016/S0896-6273(02)01104-2

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., … McCarroll, S. A. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*, *161*(5), 1202–1214.

http://doi.org/10.1016/J.CELL.2015.05.002

Marder, E. (1976). Cholinergic motor neurones in the stomatogastric system of the lobster. *The Journal of Physiology*, *257*(1), 63–86.

Marder, E. (2011). Variability, compensation, and modulation in neurons and circuits. *Proceedings of the National Academy of Sciences of the United States of America*, *108 Suppl*, 15542–8. http://doi.org/10.1073/pnas.1010674108

Marder, E. (2012). Neuromodulation of neuronal circuits: Back to the future. *Neuron, 76*(1), 1–11. http://doi.org/10.1016/j.neuron.2012.09.010

Marder, E., & Bucher, D. (2001). Central pattern generators and the control of rhythmic movements. *Current Biology : CB*, *11*(23), R986-96. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11728329

Marder, E., & Bucher, D. (2007). Understanding circuit dynamics using the stomatogastric nervous system of lobsters and crabs. *Annual Review of Physiology*, *69*, 291–316. http://doi.org/10.1146/annurev.physiol.69.031905.161516

Marder, E., & Eisen, J. S. (1984). Transmitter identification of pyloric neurons: electrically coupled neurons use different transmitters. *J Neurophysiol*, *51*(6), 1345–1361.

Marder, E., O'Leary, T., & Shruti, S. (2014). Neuromodulation of circuits with variable parameters: Single neurons and small circuits reveal principles of state-dependent and robust neuromodulation. *Annual Review of Neuroscience*, *37*(1), 329–346. http://doi.org/10.1146/annurev-neuro-071013-013958

Marder, E., & Paupardin-Tritsch, D. (1978). The pharmacological properties of some crustacean neuronal acetylcholine, gamma-aminobutyric acid, and L-glutamate responses. *The Journal of Physiology*, *280*, 213–36. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/211227

Marder, E., & Paupardin-Tritsch, D. (1980). The pharmacological profile of the acetylcholine response of a crustacean muscle. *The Journal of Experimental Biology*, *88*, 147–59.

Masland, R. H. (2004). Neuronal cell types. *Current Biology*, *14*(13), R497–R500. http://doi.org/10.1016/j.cub.2004.06.035

Masuyama, K., Zhang, Y., Rao, Y., & Wang, J. W. (2012). Mapping neural circuits with activity-dependent nuclear import of a transcription factor. *Journal of Neurogenetics*, *26*(1), 89–102. http://doi.org/10.3109/01677063.2011.642910

Maynard, D. M. (1972). Simpler networks. *Annals of the New York Academy of Sciences*, *193*, 59–72. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4564740

Maynard, D. M., & Dando, M. R. (1974). The Structure of the Stomatogastric Neuromuscular System in Callinectes sapidus, Homarus americanus and Panulirus argus (Decapoda Crustacea). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *268*(892), 161–220. http://doi.org/10.1098/rstb.1974.0024

McClintock, T. S., & Ache, B. W. (1989). Histamine directly gates a chloride channel in lobster olfactory receptor neurons. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(20), 8137–41.

McClintock, T. S., Ache, B. W., & Derby, C. D. (2006). Lobster olfactory genomics. *Integrative and Comparative Biology*, *46*(6), 940–7. http://doi.org/10.1093/icb/icj050

McGrath, L. L., Vollmer, S. V., Kaluziak, S. T., & Ayers, J. (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. http://doi.org/10.1186/s12864-016-2373-3

Mehta, P., Kreeger, L., Wylie, D. C., Pattadkal, J. J., Lusignan, T., Davis, M. J., … Zemelman, B. V. (2019). Functional Access to Neuron Subclasses in Rodent and Primate Forebrain. *Cell Reports*, *26*(10), 2818–2832.e8. http://doi.org/10.1016/J.CELREP.2019.02.011

Mermelstein, P. G., Bito, H., Deisseroth, K., & Tsien, R. W. (2000). Critical dependence of cAMP response element-binding protein phosphorylation on L-type calcium channels supports a selective response to EPSPs in preference to action potentials. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *20*(1), 266–73. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10627604

Mitchell, G. S., & Johnson, S. M. (2003). Invited Review: Neuroplasticity in respiratory motor control. *Journal of Applied Physiology*, *94*(1), 358–374. http://doi.org/10.1152/japplphysiol.00523.2002

Mito, M., Kadota, M., Tanaka, K., Furuta, Y., Abe, K., Iwasaki, S., & Nakagawa, S. (2018). Cell Type-Specific Survey of Epigenetic Modifications by Tandem Chromatin Immunoprecipitation Sequencing. *Scientific Reports*, *8*(1), 1143. http://doi.org/10.1038/s41598-018-19494-9

Moore, M. J. (2005). From Birth to Death: The Complex Lives of Eukaryotic mRNAs. *Science*, *309*(5740), 1514–1518. http://doi.org/10.1126/science.1111443

Moroz, L. L., Edwards, J. R., Puthanveettil, S. V, Kohn, A. B., Ha, T., Heyland, A., … Kandel, E. R. (2006). Neuronal transcriptome of *Aplysia*: Neuronal compartments and circuitry. *Cell, 127*(7), 1453–1467. http://doi.org/10.1016/j.cell.2006.09.052

Moroz, L. L., & Kohn, A. B. (2013). Single-Neuron Transcriptome and Methylome Sequencing for Epigenomic Analysis of Aging. In *Methods in molecular biology (Clifton, N.J.)* (Vol. 1048, pp. 323–352). http://doi.org/10.1007/978-1-62703-556-9_21

Mulloney, B., & Selverston, A. I. (1974). Organization of the stomatogastric ganglion of the spiny lobster - III. Coordination of the two subsets of the gastric system. *Journal of Comparative Physiology*, *91*(1), 53–78. http://doi.org/10.1007/BF00696156

Neville, M. C., Nojima, T., Ashley, E., Parker, D. J., Walker, J., Southall, T., … Goodwin, S. F. (2014). Male-specific fruitless isoforms target neurodevelopmental genes to specify a sexually dimorphic nervous system. *Current Biology : CB*, *24*(3), 229–41. http://doi.org/10.1016/j.cub.2013.11.035

Northcutt, A. J., Fischer, E. K., Puhl, J. G., Mesce, K. A., & Schulz, D. J. (2018). An annotated CNS transcriptome of the medicinal leech, Hirudo verbana: De novo sequencing to characterize genes associated with nervous system activity. *PLoS ONE*, *13*(7). http://doi.org/10.1371/journal.pone.0201206

Northcutt, A. J., Lett, K. M., Garcia, V. B., Diester, C. M., Lane, B. J., Marder, E., & Schulz, D. J. (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. http://doi.org/10.1186/s12864-016-3215-z

O'Leary, T., van Rossum, M. C. W., & Wyllie, D. J. A. (2010). Homeostasis of intrinsic excitability in hippocampal neurones: dynamics and mechanism of the response to chronic depolarization. *The Journal of Physiology*, *588*(Pt 1), 157–70. http://doi.org/10.1113/jphysiol.2009.181024

O'Leary, T., Williams, A. H., Caplan, J. S., & Marder, E. (2013). Correlations in ion channel expression emerge from homeostatic tuning rules. *Proceedings of the National Academy of Sciences*, *110*(28), E2645–E2654. http://doi.org/10.1073/pnas.1309966110

O'Leary, T., Williams, A. H., Franci, A., & Marder, E. (2014). Cell Types, Network Homeostasis, and Pathological Compensation from a Biologically Plausible Ion Channel Expression Model. *Neuron*, *82*(4), 809–821. http://doi.org/10.1016/j.neuron.2014.04.002

O'Leary, T., Williams, A. H., Franci, A., & Marder, E. (2014). Cell Types, Network

Homeostasis, and Pathological Compensation from a Biologically Plausible Ion Channel Expression Model. *Neuron*, *82*(4), 809–821. http://doi.org/10.1016/j.neuron.2014.04.002

Olivera-Martinez, I., Schurch, N., Li, R. A., Song, J., Halley, P. A., Das, R. M., … Storey, K. G. (2014). Major transcriptome re-organisation and abrupt changes in signalling, cell cycle and chromatin regulation at neural differentiation in vivo. *Development*, *141*(16), 3266–3276. http://doi.org/10.1242/dev.112623

Otopalik, A. G., Lane, B., Schulz, D. J., & Marder, E. (2017). Innexin expression in electrically coupled motor circuits. *Neuroscience Letters*. http://doi.org/10.1016/j.neulet.2017.07.016

Otopalik, A. G., Sutton, A. C., Banghart, M., & Marder, E. (2017). When complex neuronal structures may not matter. *ELife*, *6*. http://doi.org/10.7554/eLife.23508

Otsuka, M., Iversen, L. L., Hall, Z. W., & Kravitz, E. A. (1966). Release of gamma-aminobutyric acid from inhibitory nerves of lobster. *Proceedings of the National Academy of Sciences of the United States of America*, *56*(4), 1110–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5230136

Otsuka, M., Kravitz, E. A., & Potter, D. D. (1967). Physiological and chemical architecture of a lobster ganglion with particular reference to gamma-aminobutyrate and glutamate. *Journal of Neurophysiology*, *30*(4), 725–52.

Ouyang, Q., Goeritz, M., & Harris-Warrick, R. M. (2007). *Panulirus interruptus* Ih-channel gene PIIH: modification of channel properties by alternative splicing and role in rhythmic activity. *Journal of Neurophysiology, 97*(April 2007), 3880–3892. http://doi.org/10.1152/jn.00246.2007

Padgett, C. L., & Slesinger, P. A. (2010). GABAB receptor coupling to G-proteins and ion channels. *Advances in Pharmacology (San Diego, Calif.)*, *58*, 123–47. http://doi.org/10.1016/S1054-3589(10)58006-2

Pan, Y., & Baker, B. S. (2014). Genetic identification and separation of innate and experience-dependent courtship behaviors in Drosophila. *Cell*, *156*(1–2), 236–48. http://doi.org/10.1016/j.cell.2013.11.041

Parmhans, N., Sajgo, S., Niu, J., Luo, W., & Badea, T. C. (2018). Characterization of retinal ganglion cell, horizontal cell, and amacrine cell types expressing the neurotrophic receptor tyrosine kinase Ret. *Journal of Comparative Neurology*, *526*(4), 742–766. http://doi.org/10.1002/cne.24367

Parrish, J. Z., Kim, C. C., Tang, L., Bergquist, S., Wang, T., Derisi, J. L., … Davis, G. W. (2014). Krüppel mediates the selective rebalancing of ion channel expression. *Neuron*, *82*(3), 537–44.

http://doi.org/10.1016/j.neuron.2014.03.015

Paul, A., Crow, M., Raudales, R., He, M., Gillis, J., & Huang, Z. J. (2017). Transcriptional Architecture of Synaptic Communication Delineates GABAergic Neuron Identity. *Cell*, *171*(3), 522–539.e20. http://doi.org/10.1016/J.CELL.2017.08.032

Pavlou, H. J., & Goodwin, S. F. (2013). Courtship behavior in Drosophila melanogaster: towards a "courtship connectome". *Current Opinion in Neurobiology*, *23*(1), 76–83. http://doi.org/10.1016/j.conb.2012.09.002

Peng, Y.-R., Shekhar, K., Yan, W., Herrmann, D., Sappington, A., Bryman, G. S., … Sanes, J. R. (2019). Molecular Classification and Comparative Taxonomics of Foveal and Peripheral Cells in Primate Retina. *Cell*, *176*(5), 1222–1237.e22. http://doi.org/10.1016/j.cell.2019.01.004

Pérez-Acevedo, N. L., & Krenz, W. D. (2005). Metabotropic glutamate receptor agonists modify the pyloric output of the crustacean stomatogastric ganglion. *Brain Research*, *1062*(1–2), 1–8. http://doi.org/10.1016/j.brainres.2005.06.061

Perez-Ortin, J. E. (2008). Genomics of mRNA turnover. *Briefings in Functional Genomics and Proteomics*, *6*(4), 282–291. http://doi.org/10.1093/bfgp/elm029

Pérez-Ortín, J. E., Alepuz, P., Chávez, S., & Choder, M. (2013). Eukaryotic mRNA Decay: Methodologies, Pathways, and Links to Other Stages of Gene Expression. *Journal of Molecular Biology*, *425*(20), 3750–3775. http://doi.org/10.1016/j.jmb.2013.02.029

Perry, R. P., & Kelley, D. E. (1970). Inhibition of RNA synthesis by actinomycin D: Characteristic dose-response of different RNA species. *Journal of Cellular Physiology*, *76*(2), 127–139. http://doi.org/10.1002/jcp.1040760202

Pfeiffer-Linn, C., & Glantz, R. M. (1991). An arthropod NMDA receptor. *Synapse (New York, N.Y.)*, *9*(1), 35–42. http://doi.org/10.1002/syn.890090106

Phelan, P. (2005). Innexins: Members of an evolutionarily conserved family of gap-junction proteins. *Biochimica et Biophysica Acta - Biomembranes*, *1711*(2 SPEC. ISS.), 225–245. http://doi.org/10.1016/j.bbamem.2004.10.004

Pichon, X., Wilson, L. A., Stoneley, M., Bastide, A., King, H. A., Somers, J., & Willis, A. E. E. (2012). RNA binding protein/RNA element interactions and the control of translation. *Current Protein & Peptide Science*, *13*(4), 294–304. http://doi.org/10.2174/138920312801619475

Podda, M. V., & Grassi, C. (2014). New perspectives in cyclic nucleotide-mediated functions in the CNS: the emerging role of cyclic nucleotide-gated

(CNG) channels. *Pflügers Archiv : European Journal of Physiology*, *466*(7), 1241–57. http://doi.org/10.1007/s00424-013-1373-2

Poulin, J.-F., Tasic, B., Hjerling-Leffler, J., Trimarchi, J. M., & Awatramani, R. (2016). Disentangling neural cell diversity using single-cell transcriptomics. *Nature Neuroscience*, *19*(9), 1131–41. http://doi.org/10.1038/nn.4366

Raj, A., & van Oudenaarden, A. (2008). Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell*, *135*(2), 216–226. http://doi.org/10.1016/j.cell.2008.09.050

Ransdell, J. L., Faust, T. B., & Schulz, D. J. (2010). Correlated Levels of mRNA and Soma Size in Single Identified Neurons: Evidence for Compartment-specific Regulation of Gene Expression. *Frontiers in Molecular Neuroscience*, *3*, 116. http://doi.org/10.3389/fnmol.2010.00116

Ransdell, J. L., Temporal, S., West, N. L., Leyrer, M. L., & Schulz, D. J. (2013). Characterization of inward currents and channels underlying burst activity in motoneurons of crab cardiac ganglion. *Journal of Neurophysiology*, *110*(Berlind 1993), 42–54. http://doi.org/10.1152/jn.00009.2013

Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., … Human Cell Atlas Meeting Participants, H. C. A. M. (2017). The Human Cell Atlas. *ELife*, *6*. http://doi.org/10.7554/eLife.27041

Remler, M., Selverston, A., & Kennedy, D. (1968). Lateral giant fibers of crayfish: location of somata by dye injection. *Science (New York, N.Y.)*, *162*(3850), 281–3.

Ren, G. R., Folke, J., Hauser, F., Li, S., & Grimmelikhuijzen, C. J. P. (2015). The A- and B-type muscarinic acetylcholine receptors from *Drosophila melanogaster* couple to different second messenger pathways. *Biochemical and Biophysical Research Communications*, *462*(4), 358–64. http://doi.org/10.1016/j.bbrc.2015.04.141

Reuter, J. A., Spacek, D. V., & Snyder, M. P. (2015). High-Throughput Sequencing Technologies. *Molecular Cell*, *58*(4), 586–597. http://doi.org/10.1016/J.MOLCEL.2015.05.004

Reyes-Colón, D., Vázquez-Acevedo, N., Rivera, N. M., Jezzini, S. H., Rosenthal, J., Ruiz-Rodríguez, E. A., … Sosa, M. A. (2010). Cloning and distribution of a putative octopamine/tyramine receptor in the central nervous system of the freshwater prawn *Macrobrachium rosenbergii*. *Brain Research*, *1348*, 42–54. http://doi.org/10.1016/j.brainres.2010.06.021

Richardson, G. M., Lannigan, J., & Macara, I. G. (2015). Does FACS perturb gene expression? *Cytometry Part A*, *87*(2), 166–175. http://doi.org/10.1002/cyto.a.22608

Robbins, J. (2001). KCNQ potassium channels: physiology, pathophysiology, and pharmacology. *Pharmacology & Therapeutics*, *90*(1), 1–19.

Rodgers, E. W., Fu, J. J., Krenz, W.-D. C., & Baro, D. J. (2011). Tonic Nanomolar Dopamine Enables an Activity-Dependent Phase Recovery Mechanism That Persistently Alters the Maximal Conductance of the Hyperpolarization-Activated Current in a Rhythmically Active Neuron. *Journal of Neuroscience*, *31*(45). Retrieved from http://www.jneurosci.org/content/31/45/16387.long

Rodgers, E. W., Krenz, W.-D. C., & Baro, D. J. (2011). Tonic Dopamine Induces Persistent Changes in the Transient Potassium Current through Translational Regulation. *Journal of Neuroscience*, *31*(37). Retrieved from http://www.jneurosci.org/content/31/37/13046.long

Roeder, T. (2003). Metabotropic histamine receptors—nothing for invertebrates? *European Journal of Pharmacology*, *466*(1–2), 85–90. http://doi.org/10.1016/S0014-2999(03)01553-X

Rosenberg, A. B., Roco, C. M., Muscat, R. A., Kuchina, A., Sample, P., Yao, Z., … Seelig, G. (2018). Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science*, *360*(6385), 176–182. http://doi.org/10.1126/science.aam8999

Rydqvist, B., & Purali, N. (1993). Transducer properties of the rapidly adapting stretch receptor neurone in the crayfish (Pacifastacus leniusculus). *The Journal of Physiology*, *469*, 193–211. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8271197

Ryge, J., Winther, O., Wienecke, J., Sandelin, A., Westerdahl, A.-C., Hultborn, H., & Kiehn, O. (2010). Transcriptional regulation of gene expression clusters in motor neurons following spinal cord injury. *BMC Genomics*, *11*(1), 365. http://doi.org/10.1186/1471-2164-11-365

Sachs, A. B. (1993). Messenger RNA degradation in eukaryotes. *Cell*, *74*(3), 413–21. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7688664

Sanes, J. R., & Masland, R. H. (2015). The Types of Retinal Ganglion Cells: Current Status and Implications for Neuronal Classification. *Annual Review of Neuroscience*, *38*(1), 221–246. http://doi.org/10.1146/annurev-neuro-071714-034120

Schulz, D. J., Goaillard, J.-M., & Marder, E. (2006). Variable channel expression in identified single and electrically coupled neurons in different animals. *Nature Neuroscience*, *9*(3), 356–62. http://doi.org/10.1038/nn1639

Schulz, D. J., Goaillard, J.-M., & Marder, E. E. (2007). Quantitative expression profiling of identified neurons reveals cell-specific constraints on highly

variable levels of gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(32), 13187–91. http://doi.org/10.1073/pnas.0705827104

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., … Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature*, *473*(7347), 337–342. http://doi.org/10.1038/nature10098

Selverston, A. I., & Moulins, M. (1987). *The Crustacean Stomatogastric System: A Model for the Study of Central Nervous Systems*. Springer Berlin Heidelberg.

Semenov, E. P., & Pak, W. L. (1999). Diversification of Drosophila chloride channel gene by multiple posttranscriptional mRNA modifications. *Journal of Neurochemistry*, *72*(1), 66–72. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9886055

Senatore, A., Edirisinghe, N., & Katz, P. S. (2015). Deep mRNA sequencing of the *Tritonia diomedea* brain transcriptome provides access to gene homologues for neuronal excitability, synaptic transmission and peptidergic signalling. *PLoS ONE*, *10*(2), e0118321. http://doi.org/10.1371/journal.pone.0118321

Sha, Y., Phan, J. H., & Wang, M. D. (2015). Effect of low-expression gene filtering on detection of differentially expressed genes in RNA-seq data. *Conference Proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, *2015*, 6461–4. http://doi.org/10.1109/EMBC.2015.7319872

Sheng, K., Cao, W., Niu, Y., Deng, Q., & Zong, C. (2017). Effective detection of variation in single-cell transcriptomes using MATQ-seq. *Nature Methods*, *14*(3), 267–270. http://doi.org/10.1038/nmeth.4145

Shin, J., Ming, G., & Song, H. (2014). Decoding neural transcriptomes and epigenomes via high-throughput sequencing. *Nature Neuroscience*, *17*(11), 1463–75. http://doi.org/10.1038/nn.3814

Shinozaki, H., & Ishida, M. (1981). Electrophysiological studies of kainate, quisqualate, and ibotenate action on the crayfish neuromuscular junction. *Advances in Biochemical Psychopharmacology*, *27*, 327–36. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7004118

Shrestha, B. R., Chia, C., Wu, L., Kujawa, S. G., Liberman, M. C., & Goodrich, L. V. (2018). Sensory Neuron Diversity in the Inner Ear Is Shaped by Activity. *Cell*, *174*(5), 1229–1246.e17. http://doi.org/10.1016/J.CELL.2018.07.007

Shruti, S., Schulz, D. J., Lett, K. M., & Marder, E. (2014). Electrical coupling and

innexin expression in the stomatogastric ganglion of the crab Cancer borealis. *Journal of Neurophysiology*, *112*(11), 2946–58. http://doi.org/10.1152/jn.00536.2014

Siebert, J. R., Conta Steencken, A., & Osterhout, D. J. (2014). Chondroitin Sulfate Proteoglycans in the Nervous System: Inhibitors to Repair. *BioMed Research International*, *2014*, 1–15. http://doi.org/10.1155/2014/845323

Siebert, J. R., & Osterhout, D. J. (2011). The inhibitory effects of chondroitin sulfate proteoglycans on oligodendrocytes. *Journal of Neurochemistry*, *119*(1), 176–188. http://doi.org/10.1111/j.1471-4159.2011.07370.x

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics (Oxford, England)*, *31*(19), 3210–2. http://doi.org/10.1093/bioinformatics/btv351

Smith, J. J., Timoshevskaya, N., Ye, C., Holt, C., Keinath, M. C., Parker, H. J., … Amemiya, C. T. (2018). The sea lamprey germline genome provides insights into programmed genome rearrangement and vertebrate evolution. *Nature Genetics*, *50*(2), 270–277. http://doi.org/10.1038/s41588-017-0036-1

Sobell, H. M. (1985). Actinomycin and DNA transcription. *Proceedings of the National Academy of Sciences of the United States of America*, *82*(16), 5328–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2410919%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC390561

Södersten, E., Toskas, K., Rraklli, V., Tiklova, K., Björklund, Å. K., Ringnér, M., … Holmberg, J. (2018). A comprehensive map coupling histone modifications with gene regulation in adult dopaminergic and serotonergic neurons. *Nature Communications*, *9*(1), 1226. http://doi.org/10.1038/s41467-018-03538-9

Sosa, M. A., Spitzer, N., Edwards, D. H., & Baro, D. J. (2004). A crustacean serotonin receptor: Cloning and distribution in the thoracic ganglia of crayfish and freshwater prawn. *Journal of Comparative Neurology*, *473*(4), 526–537. http://doi.org/10.1002/cne.20092

Stebbings, L. A., Todman, M. G., Phelan, P., Bacon, J. P., & Davies, J. A. (2000). Two *Drosophila* Innexins Are Expressed in Overlapping Domains and Cooperate to Form Gap-Junction Channels. *Molecular Biology of the Cell*, *11*(7), 2459–2470. http://doi.org/10.1091/mbc.11.7.2459

Stepanyan, R., Day, K., Urban, J., Hardin, D. L., Shetty, R. S., Derby, C. D., … McClintock, T. S. (2006). Gene expression and specificity in the mature zone of the lobster olfactory organ. *Physiological Genomics*, *25*(2), 224–33. http://doi.org/10.1152/physiolgenomics.00276.2005

Stevens, C. F. (1998). Neuronal diversity: too many cell types for comfort? *Current Biology : CB*, *8*(20), R708-10. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9778523

Stretton, A. O., & Kravitz, E. A. (1968). Neuronal geometry: determination with a technique of intracellular dye injection. *Science (New York, N.Y.)*, *162*(3849), 132–4.

Subramaniam, S. (1998). The Biology Workbench--a seamless database and analysis environment for the biologist. *Proteins*, *32*(1), 1–2. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9672036

Sunkin, S. M., Ng, L., Lau, C., Dolbeare, T., Gilbert, T. L., Thompson, C. L., … Dang, C. (2012). Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic Acids Research*, *41*(D1), D996–D1008. http://doi.org/10.1093/nar/gks1042

Swensen, A. M., Golowasch, J., Christie, A. E., Coleman, M. J., Nusbaum, M. P., & Marder, E. (2000). GABA and responses to GABA in the stomatogastric ganglion of the crab Cancer borealis. *The Journal of Experimental Biology*, *203*(Pt 14), 2075–2092.

TAKEUCHI, A., & TAKEUCHI, N. (1964). THE EFFECT ON CRAYFISH MUSCLE OF IONTOPHORETICALLY APPLIED GLUTAMATE. *The Journal of Physiology*, *170*, 296–317. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14165167

Tamvacakis, A. N., Senatore, A., & Katz, P. S. (2018). Single neuron serotonin receptor subtype gene expression correlates with behaviour within and across three molluscan species. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1885), 20180791. http://doi.org/10.1098/rspb.2018.0791

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., … Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods*, *6*(5), 377–382. http://doi.org/10.1038/nmeth.1315

Tasic, B. (2018). Single cell transcriptomics in neuroscience: cell classification and beyond. *Current Opinion in Neurobiology*, *50*, 242–249. http://doi.org/10.1016/J.CONB.2018.04.021

Tasic, B., Levi, B. P., & Menon, V. (2017). Single-Cell Transcriptomic Characterization of Vertebrate Brain Composition, Development, and Function. In *Decoding Neural Circuit Structure and Function* (pp. 437–468). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-319-57363-2_18

Tasic, B., Menon, V., Nguyen, T. N., Kim, T. K., Jarsky, T., Yao, Z., … Zeng, H.

(2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nature Neuroscience*, *19*(2), 335–46. http://doi.org/10.1038/nn.4216

Tasic, B., Yao, Z., Graybuck, L. T., Smith, K. A., Nguyen, T. N., Bertagnolli, D., … Zeng, H. (2018). Shared and distinct transcriptomic cell types across neocortical areas. *Nature*, *563*(7729), 72–78. http://doi.org/10.1038/s41586-018-0654-5

Temporal, S., Lett, K. M., & Schulz, D. J. (2014a). Activity-dependent feedback regulates correlated ion channel mRNA levels in single identified motor neurons. *Current Biology*, *24*(16), 1899–1904. http://doi.org/10.1016/j.cub.2014.06.067

Temporal, S., Lett, K. M., & Schulz, D. J. (2014b). Activity-dependent feedback regulates correlated ion channel mRNA levels in single identified motor neurons. *Current Biology*, *24*(16), 1899–1904. http://doi.org/10.1016/j.cub.2014.06.067

Tenenbaum, S. A., Christiansen, J., & Nielsen, H. (2011). The Post-transcriptional Operon. In *Methods in molecular biology (Clifton, N.J.)* (Vol. 703, pp. 237–245). http://doi.org/10.1007/978-1-59745-248-9_16

Tessmar-Raible, K., Raible, F., Christodoulou, F., Guy, K., Rembold, M., Hausen, H., & Arendt, D. (2007). Conserved sensory-neurosecretory cell types in annelid and fish forebrain: insights into hypothalamus evolution. *Cell*, *129*(7), 1389–400. http://doi.org/10.1016/j.cell.2007.04.041

Theissinger, K., Falckenhayn, C., Blande, D., Toljamo, A., Gutekunst, J., Makkonen, J., … Kokko, H. (2016). De Novo assembly and annotation of the freshwater crayfish Astacus astacus transcriptome. *Marine Genomics*, *28*, 7–10. http://doi.org/10.1016/j.margen.2016.02.006

Thoby-Brisson, M., & Simmers, J. (1998). Neuromodulatory inputs maintain expression of a lobster motor pattern-generating network in a modulation-dependent state: evidence from long-term decentralization in vitro. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *18*(6), 2212–25. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9482805

Thoby-Brisson, M., & Simmers, J. (2000). Transition to Endogenous Bursting After Long-Term Decentralization Requires De Novo Transcription in a Critical Time Window. *Journal of Neurophysiology*, *84*(1), 596–599. http://doi.org/10.1152/jn.2000.84.1.596

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic*

Acids Research, *22*(22), 4673–80.

Tietjen, I., Rihel, J. M., Cao, Y., Koentges, G., Zakhary, L., & Dulac, C. (2003). Single-cell transcriptional analysis of neuronal progenitors. *Neuron*, *38*(2), 161–75. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12718852

Tobin, A.-E., Cruz-Bermúdez, N. D., Marder, E., & Schulz, D. J. (2009). Correlations in ion channel mRNA in rhythmically active neurons. *PloS One*, *4*(8), e6742. http://doi.org/10.1371/journal.pone.0006742

Tobin, A. E., Cruz-Bermúdez, N. D., Marder, E., & Schulz, D. J. (2009). Correlations in ion channel mRNA in rhythmically active neurons. *PLoS ONE*, *4*(8), 2–9. http://doi.org/10.1371/journal.pone.0006742

Tomer, R., Denes, A. S., Tessmar-Raible, K., & Arendt, D. (2010). Profiling by Image Registration Reveals Common Origin of Annelid Mushroom Bodies and Vertebrate Pallium. *Cell*, *142*(5), 800–809. http://doi.org/10.1016/j.cell.2010.07.043

Touroutine, D., Fox, R. M., Von Stetina, S. E., Burdina, A., Miller, D. M., & Richmond, J. E. (2005). acr-16 encodes an essential subunit of the levamisole-resistant nicotinic receptor at the Caenorhabditis elegans neuromuscular junction. *The Journal of Biological Chemistry*, *280*(29), 27013–21. http://doi.org/10.1074/jbc.M502818200

Towle, D. W., & Smith, C. M. (2006). Gene discovery in *Carcinus maenas* and *Homarus americanus* via expressed sequence tags. In *Integrative and Comparative Biology* (Vol. 46, pp. 912–918). http://doi.org/10.1093/icb/icl002

Trapnell, C. (2015). Defining cell types and states with single-cell genomics. *Genome Research*, *25*(10), 1491–1498. http://doi.org/10.1101/gr.190595.115

Tripathy, S. J., Toker, L., Li, B., Crichlow, C.-L., Tebaykin, D., Mancarci, B. O., & Pavlidis, P. (2017). Transcriptomic correlates of neuron electrophysiological diversity. *PLoS Computational Biology*, *13*(10), e1005814. http://doi.org/10.1371/journal.pcbi.1005814

Troncoso, J., & Maldonado, H. (2002). Two related forms of memory in the crab Chasmagnathus are differentially affected by NMDA receptor antagonists. *Pharmacology, Biochemistry, and Behavior*, *72*(1–2), 251–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11900795

Turrigiano, G. (2012). Homeostatic Synaptic Plasticity: Local and Global Mechanisms for Stabilizing Neuronal Function. *Cold Spring Harbor Perspectives in Biology*, *4*(1), a005736. http://doi.org/10.1101/CSHPERSPECT.A005736

Turrigiano, G., Abbott, L. F., & Marder, E. (1994). Activity-dependent changes in the intrinsic properties of cultured neurons. *Science (New York, N.Y.)*, *264*(5161), 974–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8178157

Turrigiano, G. G., & Nelson, S. B. (2004). Homeostatic plasticity in the developing nervous system. *Nature Reviews Neuroscience*, *5*(2), 97–107. http://doi.org/10.1038/nrn1327

Tyssowski, K. M., DeStefino, N. R., Cho, J.-H., Dunn, C. J., Poston, R. G., Carty, C. E., … Gray, J. M. (2018). Different Neuronal Activity Patterns Induce Different Gene Expression Programs. *Neuron*, *98*(3), 530–546.e11. http://doi.org/10.1016/J.NEURON.2018.04.001

Usoskin, D., Furlan, A., Islam, S., Abdo, H., Lönnerberg, P., Lou, D., … Ernfors, P. (2015). Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nature Neuroscience*, *18*(1), 145–53. http://doi.org/10.1038/nn.3881

van den Brink, S. C., Sage, F., Vértesy, Á., Spanjaard, B., Peterson-Maduro, J., Baron, C. S., … van Oudenaarden, A. (2017). Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nature Methods*, *14*(10), 935–936. http://doi.org/10.1038/nmeth.4437

van Nierop, P., Bertrand, S., Munno, D. W., Gouwenberg, Y., van Minnen, J., Spafford, J. D., … Smit, A. B. (2006). Identification and functional expression of a family of nicotinic acetylcholine receptor subunits in the central nervous system of the mollusc *Lymnaea stagnalis*. *The Journal of Biological Chemistry*, *281*(3), 1680–91. http://doi.org/10.1074/jbc.M508571200

Ventura, T., Cummins, S. F., Fitzgibbon, Q., Battaglene, S., & Elizur, A. (2014). Analysis of the central nervous system transcriptome of the Eastern rock lobster *Sagmariasus verreauxi* reveals its putative neuropeptidome. *PLoS ONE*, *9*(5), e97323. http://doi.org/10.1371/journal.pone.0097323

Ventura, T., Fitzgibbon, Q. P., Battaglene, S. C., Elizur, A., Medina, M., Anger, K., … Ventura, T. (2015). Redefining metamorphosis in spiny lobsters: molecular analysis of the phyllosoma to puerulus transition in Sagmariasus verreauxi. *Scientific Reports*, *5*, 13537. http://doi.org/10.1038/srep13537

Vogel, C., & Marcotte, E. M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics*, *13*(4), 227–232. http://doi.org/10.1038/nrg3185

Wagner, A., Regev, A., & Yosef, N. (2016). Revealing the vectors of cellular identity with single-cell genomics. *Nature Biotechnology*, *34*(11), 1145–1160. http://doi.org/10.1038/nbt.3711

Wang, Y., Ni, T., Wang, W., & Liu, F. (2018). Gene transcription in bursting: a unified mode for realizing accuracy and stochasticity. *Biological Reviews*. http://doi.org/10.1111/brv.12452

Weimann, J. M., Meyrand, P., & Marder, E. (1991). Neurons that form multiple pattern generators: identification and multiple activity patterns of gastric/pyloric neurons in the crab stomatogastric system. *Journal of Neurophysiology*, *65*(1), 111–22. http://doi.org/10.1152/jn.1991.65.1.111

West, A. E., Chen, W. G., Dalva, M. B., Dolmetsch, R. E., Kornhauser, J. M., Shaywitz, A. J., … Greenberg, M. E. (2001). Calcium regulation of neuronal gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(20), 11024–31. http://doi.org/10.1073/pnas.191352298

Wheeler, D. G., Groth, R. D., Ma, H., Barrett, C. F., Owen, S. F., Safa, P., & Tsien, R. W. (2012). Ca(V)1 and Ca(V)2 channels engage distinct modes of Ca(2+) signaling to control CREB-dependent gene expression. *Cell*, *149*(5), 1112–24. http://doi.org/10.1016/j.cell.2012.03.041

Whitaker, K. W., Neumeister, H., Huffman, L. S., Kidd, C. E., Preuss, T., & Hofmann, H. A. (2011). Serotonergic modulation of startle-escape plasticity in an African cichlid fish: a single-cell molecular and physiological analysis of a vital neural circuit. *Journal of Neurophysiology*, *106*(1), 127–37. http://doi.org/10.1152/jn.01126.2010

Wiersma, C. A. ., & Ikeda, K. (1964). Interneurons commanding swimmeret movements in the crayfish, *Procambarus clarkii* (girard). *Comparative Biochemistry and Physiology*, *12*, 509–525. http://doi.org/10.1016/0010-406X(64)90153-7

Williams, A. H., Kwiatkowski, M. A., Mortimer, A. L., Marder, E., Zeeman, M. Lou, & Dickinson, P. S. (2013). Animal-to-animal variability in the phasing of the crustacean cardiac motor pattern: an experimental and computational analysis. *Journal of Neurophysiology*, *109*(10), 2451–65. http://doi.org/10.1152/jn.01010.2012

Wissenbach, U., Niemeyer, B. A., & Flockerzi, V. (2004). TRP channels as potential drug targets. *Biology of the Cell / under the Auspices of the European Cell Biology Organization*, *96*(1), 47–54. http://doi.org/10.1016/j.biolcel.2003.12.003

Xu, C., & Su, Z. (2015). Identification of cell types from single-cell transcriptomes using a novel clustering method. *Bioinformatics*, *31*(12), 1974–1980. http://doi.org/10.1093/bioinformatics/btv088

Xu, F., Hollins, B., Landers, T. M., & McClintock, T. S. (1998). Molecular cloning of a lobster Gbeta subunit and Gbeta expression in olfactory receptor neuron

dendrites and brain neuropil. *Journal of Neurobiology, 36*(4), 525–36.

Yamamoto, D., & Koganezawa, M. (2013). Genes and circuits of courtship behaviour in Drosophila males. *Nature Reviews Neuroscience, 14*(10), 681–692. http://doi.org/10.1038/nrn3567

Yamamoto, D., & Kohatsu, S. (2017). What does the fruitless gene tell us about nature vs. nurture in the sex life of Drosophila? *Fly, 11*(2), 139–147. http://doi.org/10.1080/19336934.2016.1263778

Zahavi, T., Stelzer, G., Strauss, L., Salmon, A. Y., & Salmon-Divon, M. (2015). VennBLAST—whole transcriptome comparison and visualization tool. *Genomics, 105*(3), 131–6. http://doi.org/10.1016/j.ygeno.2014.12.004

Zeisel, A., Muñoz-Manchado, A. B., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., … Linnarsson, S. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science (New York, N.Y.), 347*(6226), 1138–42. http://doi.org/10.1126/science.aaa1934

Zeng, H., & Sanes, J. R. (2017). Neuronal cell-type classification: challenges, opportunities and the path forward. *Nature Reviews Neuroscience, 18*(9), 530–546. http://doi.org/10.1038/nrn.2017.85

Zeng, V., Villanueva, K. E., Ewen-Campen, B. S., Alwes, F., Browne, W. E., & Extavour, C. G. (2011). De novo assembly and characterization of a maternal and developmental transcriptome for the emerging model crustacean *Parhyale hawaiensis*. *BMC Genomics, 12*, 581. http://doi.org/10.1186/1471-2164-12-581

Zhang, B., Wang, J., Wang, X., Zhu, J., Liu, Q., Shi, Z., … NCI CPTAC. (2014). Proteogenomic characterization of human colon and rectal cancer. *Nature, 513*(7518), 382–387. http://doi.org/10.1038/nature13438

Zhang, Y., Oliva, R., Gisselmann, G., Hatt, H., Guckenheimer, J., & Harris-Warrick, R. M. (2003). Overexpression of a hyperpolarization-activated cation current (Ih) channel gene modifies the firing activity of identified motor neurons in a small neural network. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 23*(27), 9059–67. http://doi.org/10.1523/JNEUROSCI.23-27-09059.2003

Zhou, T., & Matsunami, H. (2018). Lessons from single-cell transcriptome analysis of oxygen-sensing cells. *Cell and Tissue Research, 372*(2), 403–415. http://doi.org/10.1007/s00441-017-2682-0

Zmora, N., Trant, J., Zohar, Y., & Chung, J. S. (2009). Molt-inhibiting hormone stimulates vitellogenesis at advanced ovarian developmental stages in the female blue crab, Callinectes sapidus 1: an ovarian stage dependent

involvement. *Saline Systems*, *5,* 7. http://doi.org/10.1186/1746-1448-5-7

VITA

I was born in De Soto, Missouri to parents Joyce and Alvin Northcutt and grew up with my brother Kyle. I attended De Soto High School from which I graduated at the top of my class with a full ride scholarship to Southeast Missouri State University (SEMO). My original drive was to become a medical doctor and help people through direct intervention, but after a few hundred hours of shadowing practitioners, I decided that I was more interested in mechanisms and treatments of disease than patient care. I also had been a part of multiple undergraduate research opportunities that instilled in me an appreciation for the scientific process. My freshman I work with Dr. Matthew Fasnacht in his analytical chemistry laboratory examining the cation exchange capacity for nutrient partitioning across different pastures. From this, I transitioned into the lab of Dr. Santaneel Ghosh and into the realm of nanoparticle research. My project sought to generate a device capable of delivering a magnetic field and increased temperature to stimulate gold nanoparticles to disrupt cultured neuroblastoma cell lines. I also had a summer research opportunity through Washington University in St. Louis coupled with Barnes Jewish Hospital where I worked in the mass spectrometry lab of Dr. Reid Townsend. My project was to generate standard curves for mass spec analysis that were used in the analysis of protein composition of various cancer biopsies.

Following graduating from SEMO with honors with a Bachelor's of Science in Biology, a minor in physics, and a minor in chemistry, I decided to pursue higher education in neuroscience. I was always, and still am, fascinated at how humans could improve themselves through a better understanding of the mechanisms by which our nervous systems function. I was accepted into the PhD graduate program at Mizzou in the department of biological sciences with both a NIH T32 Trainee Grant and a McNair Fellowship that have allowed me to be fully funded throughout my program. I rotated and landed in the lab of Dr. David Schulz after finding much common ground in our research interests, and it was easily one of the best decisions I have ever made. Here I was able to expand my interests into the realm of molecular biology and bioinformatics to ask and answer questions about fundamental neuronal regulation and expression of RNA. Following completion of my dissertation, I have accepted a position as a post-doctoral researcher in the Schulz Lab to continue my work in this area. Further, I will become a father this August with the new addition of a son to my family. After completion of my post doc, I plan on pursuing a career in industry or government.